DOI:10.15774/PPKE.ITK.2020.004

# Modular analysis of brain and free word association networks

PhD Dissertation

Bálint File

Scientific advisor:
Dr István Ulbert DSc

Roska Tamás Doctoral School of Sciences and Technology

Faculty of Information Technology and Bionics

Pázmány Péter Catholic University

Budapest, 2019

# TABLE OF CONTENTS

# List of abbreviations

Q: modularity value

fMRI: functional magnetic resonance imaging

N: Sample size

SD: standard deviation

TR: repetition time

TE: echo time

FA: fractional anisotropy

FWHM: Gaussian smoothing kernel with a full width at half maximum

ROI: region of interest

FSL: comprehensive library of analysis tools for FMRI

MIn: normalized mutual information

DMN: default mode network

dQ: delta modularity value

dMIn: delta normalized mutual information

r: Pearson's correlation coefficient

p: probability value

CoOp networks: networks of co-occurring opinions

QAP: quadratic assignment procedure

LLR: Log-likelihood Ratio

POT: Perceived Outgroup Threat

GM: Group Malleability

SDO: Social Dominance Orientation

M: mean

WAM: Weighted attitude score means

WAV: Weighted attitude score variance

$r_s$: Spearman's correlation coefficient

PANAS: Positive and Negative Affect Schedule

DOI:10.15774/PPKE.ITK.2020.004

# List of figures

significantly higher than the modularity values of the null models referring to multiple socio-cognitive categories.

Figure 3.3. The social representations of Hero (A) and Everyday Hero (B). The association networks are visualized with the ForceAtlas 2 layout. The size of a node denotes the node strength and the thickness of an edge refers to the edge weight. The networks were thresholded (edges below the value of 1 were deleted) for a better visualization. In case of Everyday Hero, nodes with the same color belong to the same module. The hierarchy and descriptive labels of the modules are presented on the dendrogram.

Figure 3.4. Common associations in the social representations of Hero (A) and Everyday Hero (B). The associations are arranged in a circular alphabetic order. The size of a node denotes the node strength and the thickness of an edge refers to the edge weight. The network was thresholded (edges below the value of 1 were deleted) for a better visualization.

Figure 4.1. Correlation between affective similarity values and co-occurrence connections (A) and correlation between the identical co-occurrence connections in Sample1 and Sample 2 (B). (A) The x-axis shows the co-occurrence connection values and the y-axis shows the affective similarity values. The x-coordinates of the data points represent the averages of the co-occurrence connection values in each of the 100 equal intervals. The association pairs were determined in each interval and their affective similarity values were also averaged. The y-coordinates of the data points represent these averaged affective similarity values. (B) The similarity of the co-occurrence connections between identical associations in the two samples was measured by Spearman's correlation. The x-axis shows the co-occurrence connections in Sample 1 and the y-axis shows the co-occurrence connections in Sample 2.

Figure 4.2. The modules of the CoOp networks. Each module is visualized with different colors. The size of a node and its label is proportional to the frequency of the given association. An edge means that two associations fall into a common module in the consensus partitioning procedure at least 40%. Both sample is displayed by "Yifan Hu Proportional" layout algorithm

DOI:10.15774/PPKE.ITK.2020.004



# List of tables

DOI:10.15774/PPKE.ITK.2020.004

Table 3.2. Modular hubs of Everyday Hero network. Normalized node strength > 2.5 refers to the modular hubs of the association network.

# 1 ABSTRACT

Networks are used for exploring underlying relations in various datasets by defining nodes and edges corresponding to a certain logic of the observed system. In the recent decades, numerous parameters were developed to describe different properties of the networks. Modular organization is an important topological property, in which certain group of nodes have denser connectivity within themselves than between other regions of the network. Also the nodes constituting a given module probably share similar properties regarding the analyzed phenomenon. Modularity maximization is among the most popular network modularization algorithms, hence the number and size of the modules are not predefined. In the present thesis 1) I demonstrated a modularity value based improvement for the statistical evaluation of brain networks and 2) I developed and validated a process for defining polarized opinions based on the modules of free word association networks.

The statistical evaluation of brain networks is based on measuring the decrease of modularity caused by shifting a node from its own module to another module. On one hand, local modularity measures the influence of a node in the formation of the modular organization. On the other hand, we evaluated differences in the community membership of the nodes between the modular structures of two groups. The method was tested on resting-state fMRI data of 20 young and 20 elderly subjects. The findings indicate that applying the modularity on a local scale is a promising biomarker for detecting differences between subpopulations.

We developed a method which can identify polarized public opinions by finding the maximal modularity partitioning in a network of statistically-related free word associations. The co-occurrence based relations of free word associations reflected emotional similarity and the modules of the association network were validated on well-established measures. The results were relatively consistent in two independent samples. We demonstrated that analyzing the modular organization of association networks can be a tool for identifying the most important dimensions of public opinions about a relevant social issue without using pre-defined constructs.

# 2 MEASURING THE INFLUENCE OF A NODE IN THE FORMATION OF THE FUNCTIONAL BRAIN MODULAR ORGANIZATION

## 2.1 INTRODUCTION

Graph theoretical analyses of complex functional networks, obtained using fMRI and MEG (for review see [4]), has demonstrated that brain functional networks have a modular (sub-network) structure, where a module is defined as a highly integrated sub-network consisting of regions with much denser connectivity among themselves than between those regions and the rest of the brain. Importantly, it has been shown that these modules are mainly composed of brain regions which are functionally and anatomically related to each other [5]. In general, integration within a module allows efficient local processing, while segregation between modules may be necessary to avoid interference from internal or external noise that could interfere with function of the modules neural activity. Therefore, studies of brain modules enable the identification of groups of brain regions that may serve common cognitive functions of neural activity [6]. Alterations of the global functional modular structure and decreased efficiency of the modular hub regions are observed in clinical and non-clinical states [7]–[9].

Although there is not a universal way for detecting the community structure in complex networks, the modularity (Q) value optimizing algorithms [10] became the gold standard for identifying the functional brain modules. The modular partition of a network with the highest Q value comprises many within-module links and as few as possible between-module links. Modules of within a complex network can be organized in various ways. A homogenous degree distribution among a module provides equivalent importance of the nodes, while a hierarchical structure of modules establishes nodes with a higher centrality than others [11]. The aim of the current study was to present a method to evaluate differences in the community membership of the nodes between representative modular structures.

Representative community structures are derived from multiple subjects of a single clinical sample, or condition, thus it gives a unique, characteristic modular organization. The visual demonstration of different representative modular structures is hardly accompanied with their statistical evaluation. On one hand, the comparison of two, unique networks are not trivial, on the other hand, the non-linear nature of the community detection algorithms makes extremely difficult to statistically evaluate certain visual differences. For example, how can we prove that the merging of two modules have a higher impact on the differences between two representative modular structures than the altered community assignment of one single node? Considerable visual differences between representative modular structures may originate in the altered connections of a single node, which restructure the whole community structure.

During the recent years graph theoretical measurements were developed to define the differences in community assignments of brain regions between the two subpopulations [12], [13]. The consistency of the modular structure across subjects is measured by the scaled inclusivity [13], and the modular classification of a particular brain region can be detected by Pearson's phi [12]. These methods work with the labels of the modular assignment corresponding to the maximal modularity partitioning. The label assignment is a binary process: a node is included or excluded from a community, however it is clear, that changes of the partition label of certain brain regions influence the modular organization more, than other do.

Our method is based on the quantification of the decrease of the modularity caused by shifting a node from its own module to another module. Measuring the extent of the modularity decrease caused by changing the community membership for every node, highlights nodes which are playing more important role in the maintenance of the modular structure. We also applied the method to identify significant differences in the modular assignment of the nodes between the representative community structures of two samples. We defined nodes whose community membership changes approximate the modular structure of one sample towards the other. A node with a community membership change was considered as approximating one community structure toward the other if the distance between the original community structures decreased by changing the community membership of the node. The modularity decrease caused by the approximation process was evaluated with a permutation framework. Applying the technique, we could pinpoint certain nodes (brain regions), for which community membership is prominently different in the community structure in one sample, compared to that seen in the other.

We tested the method on resting-state fMRI data of young and elderly subjects. The elderly functional network was characterized by altered community structure [14], [15] reduced modularity [14] and generally more between-module and less within-module connections than the young network [16]. We expected that our method can detect the prominent regions whose modular memberships are crucial for the formation of the young and elderly brain networks.

## 2.2 METHODS

### 2.2.1 Participants

Data of healthy young (19-21 years; N = 20; SD= ±1; 9 women) and elderly (67-85 years; N = 20; SD= ±6; 10 women) individuals was analyzed in this study. fMRI data was obtained from the 'INDI NKI/Rockland Sample'. fMRI time series from each participant were acquired in eyes open resting state condition during a 11 minutes period.

*fMRI recording, preprocessing and functional connectivity assessment*

All subjects were scanned with the same scanner (MRC35390 SIEMENS TrioTim 3T, TR=2500=ms, TE=30 ms, FA=80°, 3x3x3 mm voxels, 260 frames).

Preprocessing steps were carried out by using SPM12 and Conn 15d toolboxes. Default preprocessing steps were applied with default parameters in Conn: (1) realignment, (2) slice-timing correction, (3) segmentation and normalization, (4) ART-based scrubbing, (5) smoothing using a 8 mm full-width half-maximum (FWHM) Gaussian kernel. After the preprocessing steps the fMRI time series were band-pass filtered ([0.008 0.09] Hz), white matter and cerebrospinal fluid time series were regressed out. 95 ROIs (cortical areas and the hippocampus) were determined applying the FSL Harvard-Oxford Atlas parcellation scheme [17]. Regional average time courses of the ROIs were extracted for each individual.

Pairwise temporal correlations between all ROIs' time series were calculated, and used as measures of connectivity strengths. Correlation coefficients were converted into z-values using Fisher's transformation. Because of the ambiguity regarding the meaning of negative correlations [18], negative z-values were set to zero in the connectivity matrix. The connectivity strengths of every participant was normalized between 0 and 1. Normalization was done by a linear function, which does not affect individual network properties, but avoiding the possible bias of the inter-subject connectivity strength variance [19].

Every subject was characterized by a weighted, undirected network, where the ROIs represented the nodes and the connectivity strengths defined the weights of edges. The representative modular structures were derived from the average connectivity network [15].

### 2.2.2 Modularity and partition distance

In order to determine the modular structure, smaller functional subgraphs or modules were decomposed from the entire resting state network. The modularity ($Q$) of a graph describes the possible formation of communities in the network:

$$Q = \sum_{s=1}^{N} \left[ \frac{k_s}{L} - \left( \frac{d_s}{2L} \right)^2 \right],$$

where $N$ is the number of modules, $L$ is the total sum of all edge weights in the network, $k_s$ is the sum of all weights in module $s$, and $d_s$ is the sum of the strength of nodes (the sum of edge weights of a certain node) in module $s$ [10]. The Louvain algorithm [20] was applied to identify modular partition with high modularity. The representative modular structure was determined by applying the modularity algorithm on the young and elderly subjects' average connectivity matrices respectively.

The distance between different partition representations of networks with identical nodes can be determined by the normalized mutual information (MIn):

$$MIn = 2 * \frac{H(Y) + H(E) - H(Y,E)}{H(Y) + H(E)}$$

where H(Y) and H(E) is the entropy of the young and elderly partitions respectively and H(Y,E) is the joint entropy of the two partitions [21].

*Local modularity and approximation node shifts*

The relative importance of each region in the maintenance of the modular organization was measured by shifting each brain region to all possible extraneous modules. Shifting a node with an unstable community membership has less effect on the modularity value than shifting a node from its unique group [19]. Each transformation can be characterized by the change of the modularity value:

$$dQ_i = Q_{before\ transformation\ of\ node\ i} - Q_{after\ transformation\ of\ node\ i}$$

The average value of $dQ_i$ is the local modularity for node *i*. It defines how strongly the node is connected to its own module. The local modularity value can also be interpreted as the correspondence of a given node to the modular organization of a network.

Beside the calculation of the local modularity we can mark certain node shifts between modules, which approximate one partition towards the other (approximation node shifts). The changes of the MIn (dMIn) can detect these node shifts:

$$dMIn = MIn(young, elderly) - MIn(young', elderly),$$

in which dMIn values with negative sign denote node shifts, which approximate the young partition towards the elderly.

It is important to emphasize that the presented analysis is not symmetric, thus the two age groups have different representative partitions. Therefore, it is necessary to perform it on the calculated networks both for the young and the elderly separately.

### 2.2.3 Comparison of the local modularity to the within-module connectivity and participation coefficient value.

In order to test the relation between the local modularity and other well-known network measures random networks were generated. Small-world networks (binary, undirected) were generated with different edge density (from 0.1 to 0.7 with 0.1 steps) and with different node numbers (100 to 500 with 50 steps). The rewiring probability was set to 0.2 and 0.8 in order to generate lattice like random networks (rewiring probability = 0.2) and random networks close to the Erdős-Rényi graphs (rewiring probability = 0.8) [22]. 100, independent, random networks were generated with each parameter set. The resulted networks modularized in the same way than described for the biological networks. For every network the modular role of each node was characterized by the local modularity and two additional measures: within-module connectivity and participation coefficient.

Within-module connectivity (Z) measures the overall connectivity of the node (its strength) within the module compared to that of the other nodes in the same module:

$$Z_i = \frac{K_i - \langle K_{S_i} \rangle}{\sigma_{S_i}^{K}},$$

where $K_i$ is the within-module strength of node i (sum of all edge weights between node i and all the other nodes in its own module, $S_i$), $\langle K_{S_i} \rangle$ is the average of the within-module strength for all nodes in module $S_i$, and $\sigma_{S_i}^{K}$ is the standard deviation of K in module $S_i$ [11].

The participation coefficient (PC) of a node refers to the level of "between-modular" connectivity strength expresses how strongly a node is connected to other modules and defined as:

$$PC_i = 1 - \sum_{s=1}^{N} \left( \frac{W_{is}}{W_i} \right)^2$$

where N is the number of modules, $W_{is}$ the summation of edge weights of node i to module s and $W_i$ is the weighted degree of node i (Guimera and Amaral, 2005).

For each network Pearson's correlation was calculated between the local modularity and participation coefficient and between the local modularity and within-module connectivity values of the nodes.

The local modularity values of the representative brain networks of the two age groups were also compared to the PC and Z values.

#### 2.2.4 Statistical evaluation

Modularity values of the two age groups were compared using a permutation procedure. In each step an average network of mixed group was created by randomly exchanging the membership of young and elderly subjects, then the maximal modularity value for this mixed group was calculated. Repeating the procedure 5000 times, we could fit the original young and elderly modularity values to the distribution of the mixed groups' maximal modularity.

The level of significance of the local modularity of every node and the approximation node shifts were also determined with the distribution provided by mixed groups. The young and elderly modular partitions showed a different pattern of node assignment to modules, thus we tested them separately.

For describing the process, we chose the young group as a reference. The nodes of the mixed group average network were assigned to the same modules as the young representative partition. The local modularity of each node and the dQ caused by the approximation node shifts were determined. Since the mixed group was divided into the same modules as the young group, the same approximation node shifts were applied to the mixed group as to the reference group.

Repeating the procedure for 5000 times we got distributions for the local modularity and for the dQ of every approximation node shifts. The null hypothesis was rejected if the local modularity (or dQ of a particular approximation node shift) of the reference group was lower than 95% of the corresponding value of the mixed group. The null hypothesis was tested for every local modularity value and for the dQ of every approximation node shifts separately.

## 2.3 RESULTS

### 2.3.1 Modularity and partition distance

The modularity of the young representative network ($Q_{young}$=0.25) was significantly higher than that of the mixed group (p=0.0036, 5000 permutations), while the modularity of the elderly representative network ($Q_{elderly}$=0.21) was significantly lower than that of the mixed group (p=0.016, 5000 permutations).

The modularity algorithm detected 4 functional modules in the young and 3 modules in the elderly group (Figure 2.1.). The occipital module showed the highest overlap between the two age groups, while the fronto-temporal and the default mode network (DMN) were merged to one single cluster in the elderly group.

DOI:10.15774/PPKE.ITK.2020.004

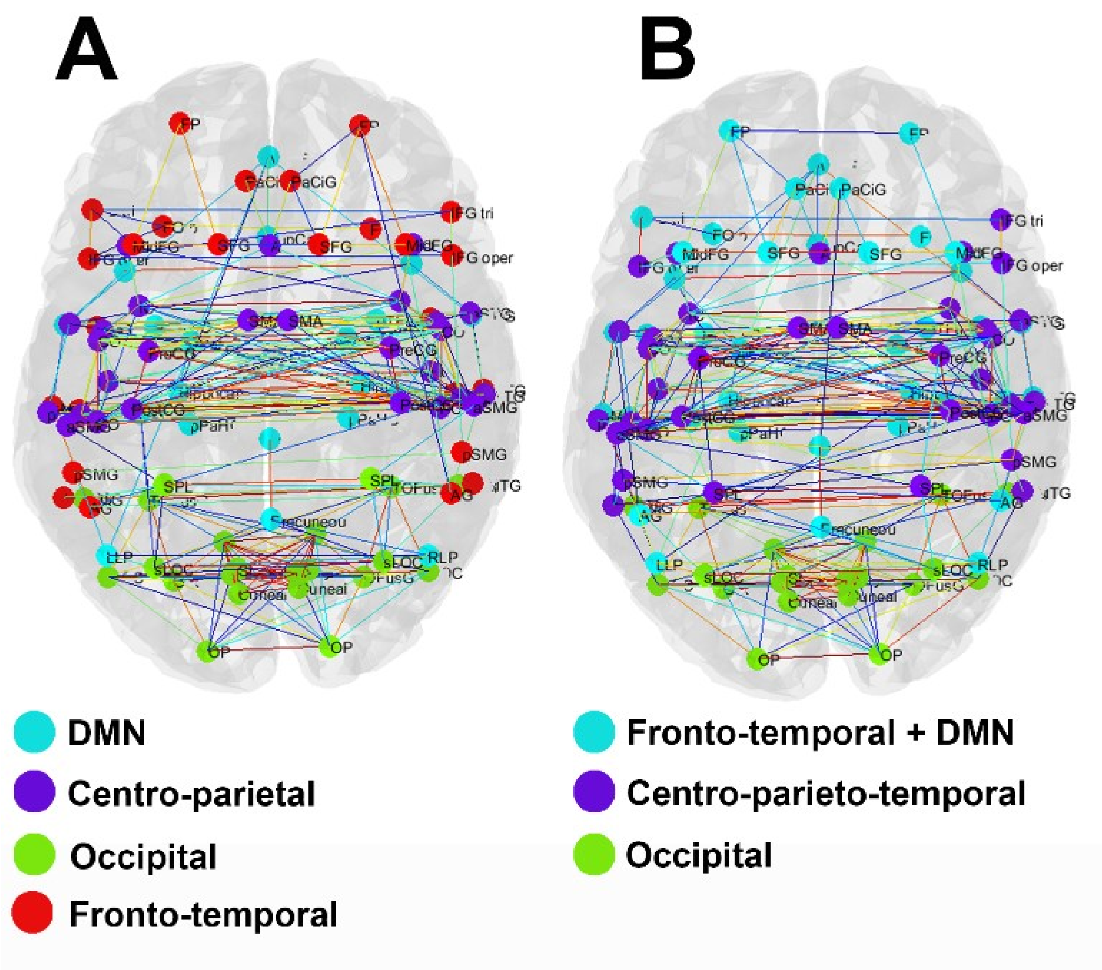



Figure 2.1. *Representative modular organization of young (A) and elderly (B) groups.* Different node colors denote different modules. The coloring of the modules was optimized by the Hungarian algorithm [1] using Jaccard-similarity between the modules as a cost.

We checked whether the modularity differences can be related to the different network density of the two age group. We have not found any significant differences in the density: $\text{Density}_{\text{Young}}$=.84±.05; $\text{Density}_{\text{Eldelry}}$=.85±.05; t(38)=-0.66, p=0.5.

### 2.3.2 Age-related local modularity changes of brain networks

As the young group was characterized by a significantly higher modularity than the elderly group, the higher local modularity values were also found in this group (Figure 2.2.). The most prominent increase of the local modularity in the young group compared to the mixed group was observed in the occipital regions (15 from the 18 occipital regions showed significantly higher local modularity). On the contrary, the local modularity of the structures of the default mode network was only slightly affected by aging. Increased local modularity was found only for 4 (right/left hippocampus, left parahippocampal gyrus, left angular gyrus) of the 13 regions of the DMN.

In contrast to the young group only a few regions of increased local modularity were observed in the elderly group. The medial frontal cortex ($p$=0.03), paracingulate gyri ($p_{right}$=0.03; $p_{left}$=0.001), superior parietal lobules ($p_{right}$=0.009; $p_{left}$=0.01) and inferior temporal gyri ($p_{right}$=0.004; $p_{left}$=0.05) showed an increased local modularity compared to the mixed group.

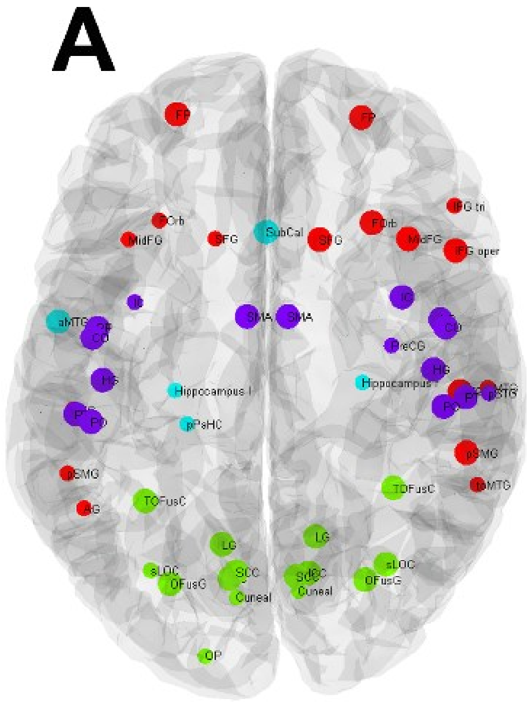



Figure 2.2. *Age-related differences in local modularity.* The local modularity determines the correspondence of a given node to the modular organization of a network. The spheres denote brain regions which show a significantly higher ($p$<0.05) local modularity (segregation) in the young (A) and elderly (B) representative community structure. Spheres with larger radius mark $p$<0.01 level of significance. Color marks the modules in a similar fashion as in Figure 2.1.

Local modularity values were compared to the Z and PC values of simulated small-world networks with different parameters and also to the young and elderly networks, For the biological networks, the local modularity value showed a moderate to strong negative correlation with the PC values ($PC_{Young}$= -0.93, $PC_{Eldelry}$= -0.65) and the same strength but different direction with the Z values ($Z_{Young}$= 0.74, $Z_{Eldelry}$= 0.9). We obtained a similar result for the simulated small-world networks (for detailed results, please see Figure 2.3).

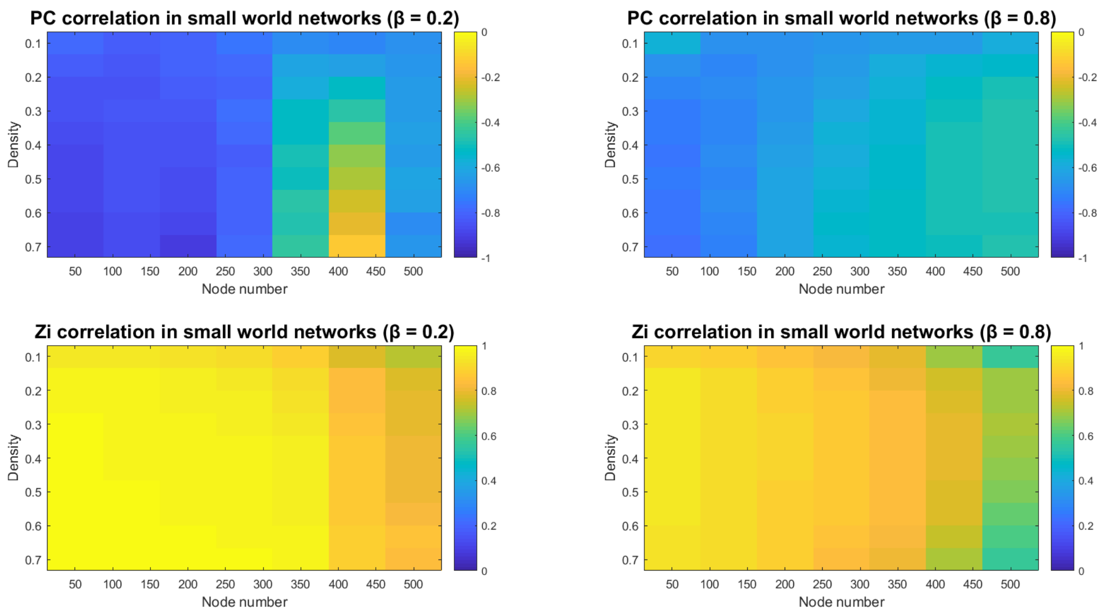



*Figure 2.3. Correlation between „local modularity" and „participation coefficient (PC)/within module connectivity (Z)" on simulated networks.* Small-world networks were generated with different node numbers and density. The rewiring probability (β) was set to 0.2 for a lattice like network and 0.8 for simulating random networks. The heatmaps represent the correlation values. The correlation values were the following: $PC_{lattice}$(-0.7047±0.1891) $Z_{lattice}$ (0.9282±0.07); $PC_{random}$ (-0.6055±0.09), $Z_{random}$(0.8316±0.09).

### 2.3.3 dQ values of the approximation node shifts

There were 28 approximation node shifts from the young group toward the elderly, and 9 approximation node shifts from the elderly group toward the young. In case of the young group node transformations indicated shifts mainly from the fronto-temporal module, since this module was absorbed to other modules in the elderly group (see Figure 2.1.). The majority of the approximation node shifts caused a significant modularity decrease compared to the mixed group.

In case of the elderly, transformation of the bilateral middle temporal gyri ($p_{right}$=0.05; $p_{left}$=0.05) and the right supramarginal gyrus (p=0.045) from the 'centro-parieto-temporal' module to the 'fronto-temporal + DMN' module resulted in a significant modularity decrease compared to the mixed age group. Changing of the modular assignment of the bilateral superior parietal lobules ($p_{right}$=0.02; $p_{left}$=0.04) from the elderly group's 'centro-parieto-temporal' to the young group's 'occipital' module also resulted in a significant modularity decrease compared to the mixed age group.

## 2.4 DISCUSSION

This work introduced a novel method for investigating modularity on a local scale. The method is based on the quantification of the modularity decrease caused by changing the community membership of the nodes. Our novel measure, local modularity, reveals the influence of a node's community membership in the modular organization of a network. We also defined nodes whose community membership changes approximate the modular structure of one sample towards the other. Measuring the modularity decrease caused by these node shifts pinpointed certain nodes community membership of which were prominently different in the community structure in one sample than in the other. The resulting modularity of the node shifting was statistically evaluated with a permutation framework.

We tested the method on resting state fMRI data of young and elderly subjects. As far as we know, this is the first attempt to apply modularity in a nodal scale for investigating the differences between functional brain networks. The local modularity is positively related to the within module connectivity and negatively to the between module connectivity thus nodes with high local modularity value can be related to *local hubs* (nodes with high within module connectivity and low participation coefficient) of the network[23].

We identified 4 modules in the young and 3 in the elderly group. The same number of resting state functional modules were described previously, however more sparse networks can result in more modules [15]. Similarly to other findings [14], [16] we found increased modularity in the young group. The highest local modularity was observed in the occipital module of the young group. This observation is in line with the increased segregation of the visual cortex in the young [14]. Increased local modularity in the elderly was mainly observed within the regions of the dorsal attention network (prefrontal cortex, anterior cingulate, posterior parietal lobule), indicating that these regions may preserve the modular organization of the brain network in advanced age. In a previous study it was found that the long range connections within this system showed an age-related decrease [24]. Probably the disconnection of the network resulted in the increase of local modularity.

We found the merging of the DMN to the 'fronto-temporal' module in the elderly group. Our method showed that the changes of the community assignment of these regions have a significant impact on the modularity. Geerlings and his colleagues found using the labels of the partitions [14], that the DMN only create a separate module in the young group, but merges with the fronto-parietal control network with advanced age. Also a very recent study confirmed that the between module connections of the fronto-parietal control network increased with age [25].The present findings indicate, that modularity estimated on a local scale can be a promising parameter for measuring the influence of a node in the formation of the modular organization of a network and to detect changes of the community memberships between different groups. Future studies may extend the application of the *local modularity* and *approximation node shifts. Local modularity* can be applied on any network to measure the relative importance of each node. *Approximation node shifts* has three criteria for application: 1) comparison of two, or more conditions, 2) more observation within homogenous groups 3) networks with equivalent nodes. Beside the comparison of brain networks, approximation node shift may be used for filtering noise from multiple observations.

As a limitation of this study we have to refer to the thresholding problem of functional brain networks. We used a biologically relevant cut-off values [18], [26] and from a practical view, the positive edges ensure a better partitioning of the functional network than the negative edges [19]. Moreover, age-related changes of the fMRI modular organization investigated only positive edge weight networks, thus for a better comparison to the current literature [14], [15], [27], [28] we applied the common practice. Besides this, we have to acknowledge that choosing a threshold is always an arbitrary choice [29]. Similarly, choosing the number of nodes in our network (i.e. the number of brain regions in the Harvard-Oxford parcellation scheme) can have an effect on the modular structure of the network [30]. Changing the community memberships of a partition was limited to individual node shifts between communities in this study, however applying more sophisticated transformation rules (e.g: merging, splitting partitions) may improve the sensitivity of the method in the future.

# 3 SCALE FREE AND MODULAR NETWORK PROPERTIES CAN DESCRIBE THE ORGANIZATION OF A FREE WORD ASSOCIATION NETWORK

## 3.1 INTRODUCTION

The present study maps the social representations of Hero and Everyday Hero in Hungary by representing them as networks constructed from free associations. We identify modules of the networks and categorize the associations based on their topological positions in the association networks. In order to do that, we define global hubs as the most dominant associations of the whole social representation and modular hubs as the characteristic associations in the different modules. After assessing the two social representations, we analyze the overlapping set of associations.

We emphasize that our aim is not to rigorously define concepts like Hero and Everyday Hero (for such work see Kinsella et al.[31]). What we are after is to observe how the perception of heroism changes when we make this distinction. In order to do that, we combine three theoretical perspectives: heroism as a social construct, social representation theory and network theory. In the following three subsections, we provide a short overview of these approaches.

### 3.1.1 Heroism

As we investigate heroism inductively, we concentrate on studies with a similar approach [31]–[34]. However, inductive studies are exposed to several biases. The concept of heroism is shaped by larger cultural and historical contexts as well [35]. Furthermore, different social groups can have different heroes even in the same culture [54]. In addition to the cultural and social relativity, Sullivan and Venter [34], [36] showed evidence that study participants relate differently to heroes identified by themselves than to heroes identified by others, which highlights the functional significance of the “my hero” concept instead of heroes in general. Kinsella et al. [31] collected several hero-related concepts that are often merged with heroic narrative in the general discourse: leaders, role models, sport stars and celebrities. Franco, Blau and Zimbardo [35] also argued that the meaning of heroism might be overloaded with political and media influences [32], [35].

Farley [37] suggested a distinction between Big H Heroism and Small h Heroism on a theoretical basis. Big H heroism refers to outstanding acts that display prototypical heroism. Small h Heroism does not necessarily imply grand or exceptional moral character or abilities. It usually happens in everyday circumstances and goes unnoticed by the public. Thus, the possibility of experiencing such situations is much higher.

### 3.1.2 Social representations

Social representations are ideas, opinions and attitudes shared by a social group regarding a social object [38], [39]. Inductive social representation studies frequently apply free associations [40], [41]. Moscovici [38], [39] identified the figurative core of a social representation, based on which Abric [41]–[43] developed the central core vs. periphery hypothesis.

The central core of a social representation consists of only a few and relatively abstract associations and has a pervasively influential role by defining the meaning of the whole social representation. The central core has three main functions, namely generating the meaning of the representation, influencing connections between other less important associations and stabilizing the representation under altering environments. Furthermore, the central core provides relevant norms, behavioral action plans and stereotypes in certain situations. Two representations differ if their central cores contain different associations.

DOI:10.15774/PPKE.ITK.2020.004

In contrast with the central core, periphery associations constitute the largest part of the representation. Their meanings are relatively concrete. The periphery operates as an interface between the environment and the central core. The periphery is responsible for the concretization of the representation and gradual changes of the social representation start on the periphery.

### 3.1.3 Networks

Networks are used for exploring underlying relations in various datasets (e.g., innovation processes [44], metabolic relations [45], brain functional interactions [46]). Every network consists of a set of objects, in which some pairs of objects are connected to each other. The objects are called nodes and the link between two nodes is called an edge. A network is undirected if the edges represent symmetric relations between the nodes. A network is weighted if values are assigned to the edges. The weight or even the existence of an edge between two nodes is determined by a predefined logical system [46], [47]. The node strength is the sum of weights attached to the edges of a given node.

Steyvers and Tenenebaum [48] showed that large semantic and association networks are scale-free. Scale-free networks have a small number of hubs (we refer to these hubs as global hubs). Hubs are nodes with outstanding number of edges in the network. Hubs are often defined based on an arbitrarily chosen threshold value considering the order of magnitude of node strengths in the given network [11], [49], [50]. The rest of the nodes are peripheral nodes with significantly lower number of edges.

Many real world networks can be divided into modules. Modules are subunits of the system with much denser connectivity within themselves than between other regions of the network. The elements constituting a given module probably share similar properties regarding the analyzed phenomenon [11], [51], [52]. Palla et al. [53] provided an example for a modular network of word associations starting from the word “bright”. The network was divided into four modules: Intelligence, Astronomy, Light, and Colors. The word “bright” was connected to all of them but the modules revealed alternative meanings.

DOI:10.15774/PPKE.ITK.2020.004

## 3.2 Material and Methods

### 3.2.1 Participants

This research employed two nationally representative probability samples of 506 (in case of Hero) and 503 (in case of Everyday Hero) Hungarians aged between 15 and 75 years. The participants were selected randomly from an internet-enabled panel including 15,000 members with the help of a research market company in March 2014. For the preparation of the sample, a multiple-step, proportionally stratified, probabilistic sampling method was employed.

Members of this panel used the Internet at least once a week. The panel demography is permanently filtered. More specifically, individuals are removed from the panel if they give responses too quickly (i.e., without paying attention to their response,) and/or have fake (or not used) e-mail addresses. The sample is nationally representative in terms of gender, age, level of education, and type of residence for those Hungarians who use the Internet at least once a week.

The final samples comprised $N_H$ = 502 and $N_{EH}$ = 502 respondents who gave valid answers ($M_H$ = 239, $F_H$ = 263; $M_{EH}$ = 238, $F_{EH}$ = 264) aged between 15 and 75 years ($M_{Hage}$ = 44.4 years; $SD_{Hage}$ = 16.2 years; $M_{EHage}$ = 44.0 years; $SD_{EHage}$ = 16.2 years). Regarding the highest completed level of education, 22.9%/23.1% (Hero/Everyday Hero) of the respondents had primary level of education, 24.9%/24.9% had vocational school degree, 31.5%/30.7% graduated from high school and 20.1%/21.3% had higher education degree. Regarding the place of residence, 18.9%/18.7% of the respondents lived in the capital city, 19.5%/18.5% lived in the county capitals, 31.7%/32.2% lived in towns and 29.9%/30.6% lived in villages.

### 3.2.2 Measures

The Research Ethics Committee of the Faculty of Education and Psychology of Eötvös Loránd University approved this study. All participants provided their written informed consent to participate in this study through a check-box on the online platform. In case of underage participants, parents (passive consent) were informed about the topic of the research. The ethics committee approved this consent procedure. Respondents volunteered for the study and they did not receive compensation for the participation. Furthermore, they were assured of their anonymity. Data was collected via an online questionnaire. Participants were informed about the content of the questionnaire (e.g., Hero, Everyday Hero).

We used an associative task based on Abric's [41], [43] theoretical underpinnings and on Vergès' [54] methodological (data gathering) assumptions. A respondent had to associate five words or expressions to one of the cues resulting in an individual representation. The cues were Hero or Everyday Hero.

The instruction was: "Please, write 5 words which first come into your mind about Hero/Everyday hero. Evaluate them on the following scale: negative, neutral, positive". The associations were not categorized. We followed Flament and Rouquette's [55] lemmatization criteria.

### 3.2.3 Setting up the networks

We algorithmically set up two networks which stand for the social representations of Hero and Everyday Hero in Hungary. To create such networks, we had to determine the nodes and the edges. We listed the different associations from the total set of associations to a given cue. The nodes represented these different associations. There was an edge between two associations if they were mentioned together by at least one study participant. The weight of an edge between two associations was equal to the number of times they were mentioned together. Therefore, the construction of networks was only directed by the co-occurrences of associations in the individual representations. More sophisticated methods are also available besides this relatively simple procedure. For example, it is possible to consider the rank order of the associations for each participant [41], [56]. In the present case, edge weights based on rank order would result in an arbitrary effect on our networks.

The method is similar to item-based recommendation algorithms [57], in which an item (product, movie, book, etc.) is recommended to a user based on the general pattern of other users' preferences. When a user buys an item, the algorithm recommends other products that were purchased by previous users who were also interested in the same item [58]. Therefore,

DOI:10.15774/PPKE.ITK.2020.004

products frequently purchased together are strongly linked and often recommended, while weakly tied items are not. In our networks, the associations played the role of products.

The construction of the networks can be summarized in three steps:

1) Participants who mentioned the same association at least twice were deleted.

2) We determined the nodes. We ignored associations which occurred only once. According to Abric [42], [56] they belong to the far periphery and are not necessarily stable parts of the social representation. However, these elements constitute the major part of the representation. From the network perspective, these associations typically have only one connection. These sparsely connected nodes can easily result in disconnected subnetworks which make the modular analysis more difficult. The removal of these nodes has no effect on the scale-free properties of the networks.

3) We determined the edge weight between every pair of associations, which was equal to the number of times the two associations were mentioned together. A strong edge between two associations meant that they were frequently mentioned together in the individual representations, while the absence of an edge referred to the complete separation of the two associations on the individual level.

The above-described process was applied to the Hero and Everyday Hero associations separately, thereby resulting in two weighted and undirected networks. After removing subjects mentioning the same association more than once, the number of subjects was 474 in case of Hero and 481 in case of Everyday Hero. After removing associations that occurred only once, the number of different associations was 222 in case of Hero and 210 in case of Everyday Hero. The total number of associations was 2006 in case of Hero and 1899 in the case of Everyday Hero. Further analyses (calculating scale-free properties, calculating modularity, finding global and modular hubs) were carried out on these reduced datasets.

We constructed common association networks for the two social representations. In this case, the nodes are the associations present in both the Hero and Everyday Hero networks. The edges and the edge weights are determined with the same method as in case of the social representation networks. Therefore, the common association networks are subnetworks extracted from the social representation networks.

### 3.2.4 Scale-free topology and modularity

The scale-free topology of a network refers to the power-law function that the probability distribution function ($P(x)$) of the node strength ($x$) follows:

$$P(x) \sim x^{-\alpha},$$

where $\alpha$ is the scaling parameter [59]. The scaling parameter typically lies in the range $2 < \alpha < 3$ [60].

The power-law distribution of the normalized node strengths was tested separately for the Hero and Everyday Hero networks. The Maximum Likelihood Estimation fitting model determined the scaling parameter ($\alpha$) of the power-law function and the minimum node strength ($X_{\min}$) for which the power law holds. For statistical comparison, datasets were generated with the same parameters ($X_{\min}$ and $\alpha$) as the empirical datasets. According to the null hypothesis of the Kolmogorov-Smirnoff test, the generated dataset has the same distribution as the empirical dataset. Following Clauset et al. [60] we determined the significance level as 0.1. This means that we considered our networks scale-free if p>0.1 (for the applied toolbox and a more detailed description see: http://tuvalu.santafe.edu/~aaronc/powerlaws/ [60]).

We investigated the modular organization of the association networks. In order to do that, smaller subnetworks (modules) were decomposed from the entire networks and the modularity value (Q) was calculated [61]. A modular structure of a network with a high value of Q must comprise many within-module links and as few as possible between-module links. The Louvain algorithm [20] with fine-tuning [62] was applied to identify the modular partition with the highest possible modularity. The resulting modular structure can change run by run [20]. Therefore, we applied the algorithm for 10,000 independent iterations and we chose the partition with the highest modularity value.

We examined the hierarchical relationship between the resulting modules by applying a hierarchical agglomerative clustering technique. Two clusters are merged in each iteration based on the maximal modularity criteria between the $i^{th}$ and $(i-1)^{th}$ community structure of the network (for details see [61]). The construction of the complete dendrogram can mark the cohesive modules of the social representation even if the difference between the modularity values of the $i^{th}$ and $(i-1)^{th}$ partitions is negative.

Degree-, weight-, and strength-preserving randomization [48] was applied to generate 4999 independent null models (random networks) for the social representations of both Hero and Everyday Hero. The modular organizations of the two social representation networks were tested by comparing their maximal modularity values to the corresponding random networks. We applied a nonparametric statistics (one-sided) to test whether the modularity value of the social representation networks differed from that of the random networks (for detailed description see: [63]). The significance level was defined strictly, which means we rejected the null hypothesis if the social representation network's modularity value was always higher than the corresponding random networks' modularity value.

### 3.2.5 Normalization of node strengths

Normalized node strengths and normalized intramodular node strengths [11] were calculated. These characterize the importance of each node in the whole network and within its module, respectively. The normalized node strength of node $i$ is determined as:

$$Normalized\ Node\ Strength_i = \frac{K_i}{\overline{K}},$$

where $K_i$ is the node strength of node $i$, $\overline{K}$ is the average node strength in the network. Nodes with normalized node strength > 2.5 were classified as global hubs of the network.

The normalized intramodular node strength of node $i$ is:

$$Normalized\ Intramodular\ Node\ Strength_i = \frac{K_i}{\overline{K_{S_i}}},$$

where $K_i$ is the intramodular node strength of node $i$ (sum of all edge weights between node $i$ and all the other nodes in its own module, S), $\overline{K_{S_i}}$ is the average intramodular node strengths of all nodes in the module. Nodes with normalized intramodular node strength > 2.5 were classified as modular hubs of the network.

The network construction and analysis were carried out in Matlab 7.9.1 software. All of the applied network parameters are available at https://sites.google.com/site/bctnet/. ForceAtlas2 layout algorithm [64] (Implemented in Gephi 0.8.2) was used for visualizing the networks.

## 3.3 Results

The number of negative associations in both social representations was negligible. It was 100 out of 2510 in case of Hero and 81 out of 2510 in case of Everyday Hero. As most of them occurred only once, 64 from the 2006 analyzed associations in case of the Hero and 23 from the 1899 analyzed associations in case of the Everyday Hero had negative valence. Considering the low number of negative valence scores we ignored the valences of the associations.

Scale-free properties (scaling parameter ($\alpha$), minimal normalized node strength ($X_{\min}$), p-value of the line fitting) were determined for the Hero and Everyday Hero networks. In case of Hero, we found $\alpha$=2.15 from $X_{\min}$=.312. In case of Everyday Hero, we found $\alpha$=2.21 from $X_{\min}$=.8. In the range determined by $X_{\min}$, the normalized node strength distributions showed a power law distribution (p(Hero)=.11, p(Everyday Hero)=.5). The log-log plots of the scale-free properties can be seen in Figure 3.1.

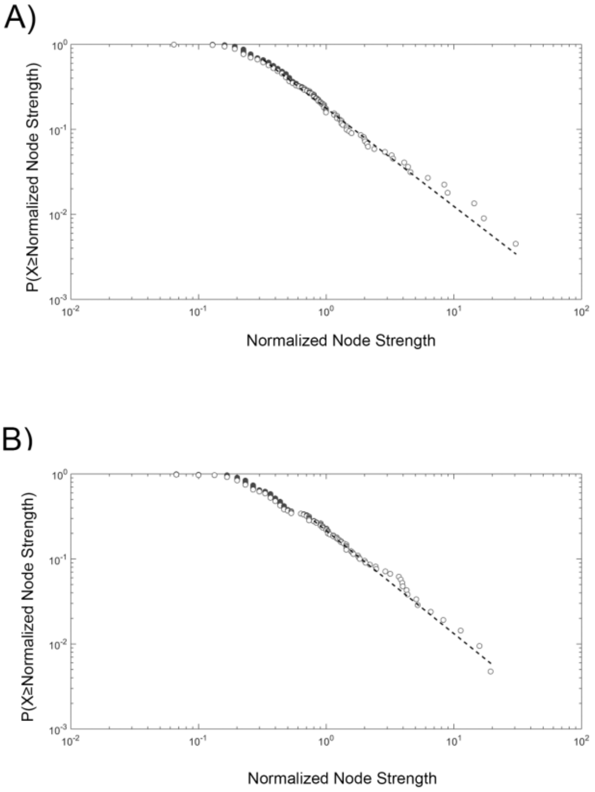



Figure 3.1. *Scale-free properties of the Hero (A) and Everyday Hero (B) networks.* The plots show the cumulative distribution functions of the normalized node strengths on log-log scales. The dashed, straight lines represent the Maximum Likelihood Estimation fitting of the data points. The power law exponents ($\alpha$) for the Hero and Everyday Hero networks are 2.15 and 2.21, respectively.

The modularity value of the Hero network (Q=.19) was not significantly higher than the corresponding modularity values of the null (random) models (p=0.19; mean=.17; standard deviation=0.027). In case of Everyday Hero, the modularity value of every (4999) independent null model was lower (p<.001; mean(random)=.15; standard deviation(random)=0.013) than the modularity value calculated for the social representation network (Q=.26). These results showed that the Hero network was non-modular and the Everyday Hero network was modular (Figure 3.2).

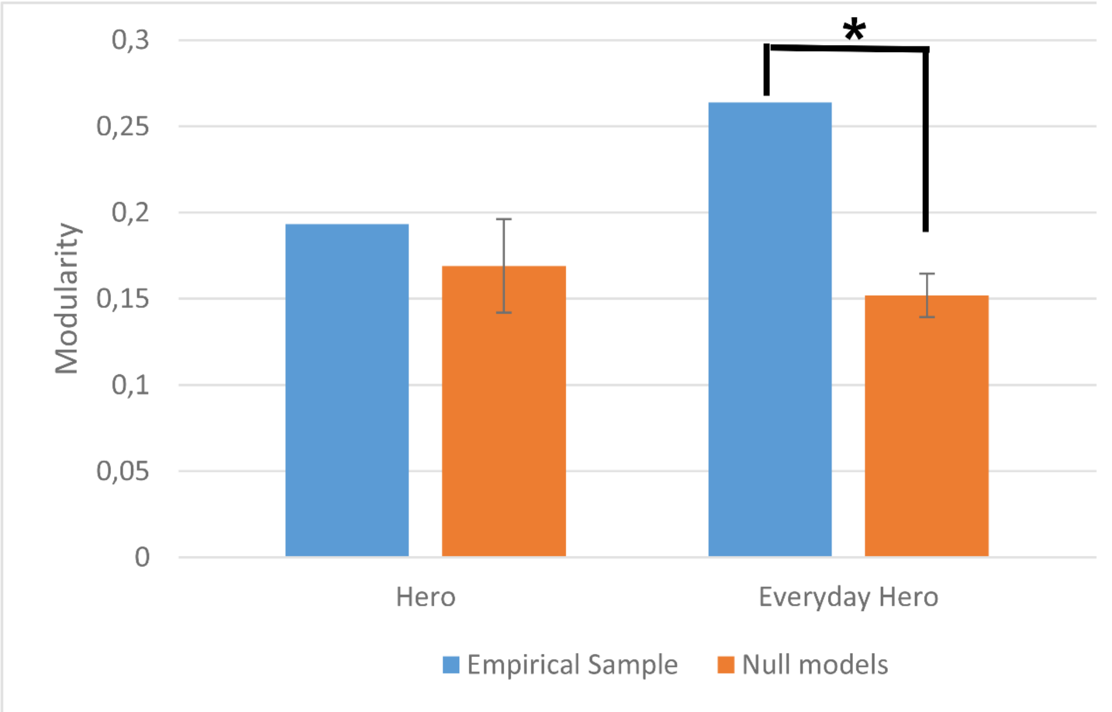



Figure 3.2. *The modularity values of the Hero (left) and Everyday Hero (right) networks comparing to the null models*. The error bars refer to the standard deviation of the modularity value derived from the null models. The modularity value of the Hero network (Q=.19) was not significantly higher than the corresponding modularity values of the null models suggesting a cohesive representation. The modularity value of the Everyday Hero network (Q=.27) was significantly higher than the modularity values of the null models referring to multiple socio-cognitive categories.

The visualization of the networks can be seen in Figure 3.3.

DOI:10.15774/PPKE.ITK.2020.004

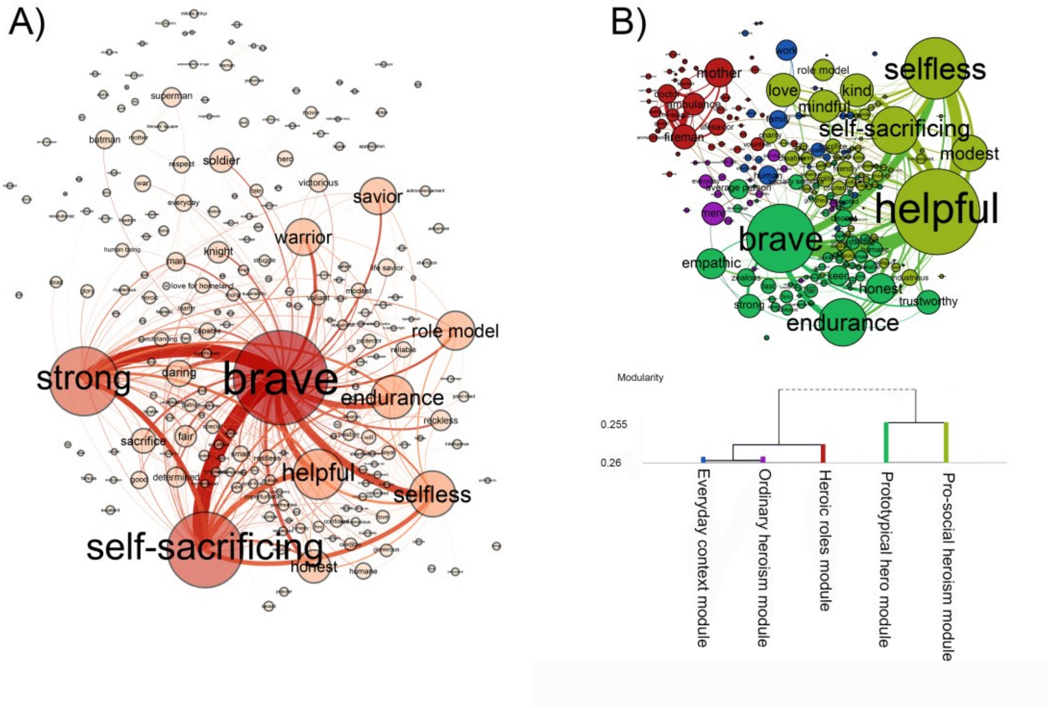



Figure 3.3. *The social representations of Hero (A) and Everyday Hero (B).* The association networks are visualized with the ForceAtlas 2 layout. The size of a node denotes the node strength and the thickness of an edge refers to the edge weight. The networks were thresholded (edges below the value of 1 were deleted) for a better visualization. In case of Everyday Hero, nodes with the same color belong to the same module. The hierarchy and descriptive labels of the modules are presented on the dendrogram.

We identified the global hubs of the Hero and Everyday Hero networks (Table 3.1.). Hero is a network in which “brave” has an outstanding number of connections and it is followed by a couple of weaker global hubs. The global hubs of Hero are predominantly abstract values (“brave”, “self-sacrificing”, “strong”, “helpful”, “selfless”, “endurance”, “honest”, “daring” and “sacrifice”). Among the global hubs of Hero, three concrete nodes appear: “warrior”, “role-model” and “savior”. Everyday Hero also has both abstract and concrete global hubs. The concrete global hubs (“fireman”, “ambulance man”, “mother” and “doctor”) are roles and occupations associated with heroism. The abstract global hubs (“helpful”, “brave”, “selfless”, “self-sacrificing”, “endurance”, “modest”, “modest”, “honest”, “mindful”, “love”, “kind” and “emphatic”) are associations expressing heroic values.

| Hero | Normalized node strength>2.5 | Everyday Hero | Normalized node strength>2.5 |
|---|---|---|---|
| brave | 30.45 | helpful | 19.32 |
| self-sacrificing | 17.15 | brave | 15.89 |
| strong | 14.42 | selfless | 11.28 |
| helpful | 8.90 | self-sacrificing | 8.28 |
| selfless | 8.38 | endurance | 6.61 |
| endurance | 6.26 | fireman | 5.21 |
| warrior | 4.56 | modest | 5.07 |
| role model | 4.37 | honest | 4.34 |
| honest | 4.08 | ambulance man | 4.27 |
| daring | 3.34 | mindful | 3.97 |
| savior | 3.24 | love | 3.94 |
| sacrifice | 2.89 | kind | 3.84 |
| | | mother | 3.70 |
| | | doctor | 3.17 |
| | | empathic | 2.90 |

Table 3.1. *Global hubs of the Hero and Everyday Hero networks.* Normalized node strength > 2.5 refers to the global hubs of the association network.

In case of Everyday Hero, we identified five modules. We labeled each of them based on their modular hubs resulting in the following: Prototypical Hero module, Everyday Context module, Pro-social Heroism module, Ordinary Heroism module and Heroic Roles module (Table 3.2.). Prototypical Hero and Pro-social Heroism belong to a superordinate group while Everyday Context, Ordinary Heroism, and Heroic Roles form another group (see the dendogram in Figure 3.3). Ordinary Heroism is a homogenous subnetwork and its nodes are relatively weakly tied. The only association that has node strength close to the threshold is "mere".

| Prototypical Hero | Everyday Context | Pro-social Heroism | Ordinary Heroism | Heroic Roles |
|---|---|---|---|---|
| brave | work | helpful | | fireman |
| endurance | successful | selfless | | ambulance man |
| honest | family | self-sacrificing | | doctor |
| strong | | modest | | mother |
| | | kind | | policeman |
| | | mindful | | |
| | | love | | |

Table 3.2. *Modular hubs of Everyday Hero network.* Normalized node strength > 2.5 refers to the modular hubs of the association network.

We calculated how many concrete social roles and contexts are present in the social representation of Hero. We found that 38 out the 222 nodes (17.1%) were occupations (e.g., doctor, fireman, etc.), social roles (e.g., warrior, savior, etc.) or concrete characters (e.g., superheroes, historical figures, etc.).

### 3.3.1 Structural differences of the common associations

We gathered the common associations of the two social representations and created two networks from them representing either Hero or Everyday Hero (Figure 3.4). The number of common associations was 85. The list of common associations is available in S4 File. A moderate correlation (r < .58, p < .001) was determined for the edge weights connecting the same nodes in the common association networks. In case of Hero, the majority of the edges are connected to “brave” with dominant links to “strong” and “self-sacrificing” (see in Figure 3.4/A). In case of Everyday Hero, the dominant edges are more balanced between “helpful”, “selfless”, “self-sacrificing” and “brave” and even the less important edges seem to be more homogenously distributed (see in Figure 3.4/B). All modules of Everyday Hero were present among the common associations (the quotient of the number of participating nodes from a module and all nodes of the module expressed in percentage) as follows*:* Prototypical Hero module: 55%; Everyday Context module: 32%; Pro-social Heroism module: 40%, Ordinary Heroism module: 40%; Heroic Roles module: 26%. The global hubs of the original networks were among the nodes of the common association networks except for “warrior” in Hero and “empathic” in Everyday Hero. Prototypical Hero module, Pro-social Heroism module and Ordinary Heroism module overlap to the highest degree with the social representation of Hero. These modules contain abstract heroic values and characteristics. Everyday Context module and Heroic Roles module are present in lower proportion among common associations. They are more concrete in terms of content. They contain social roles, occupations and social contexts.

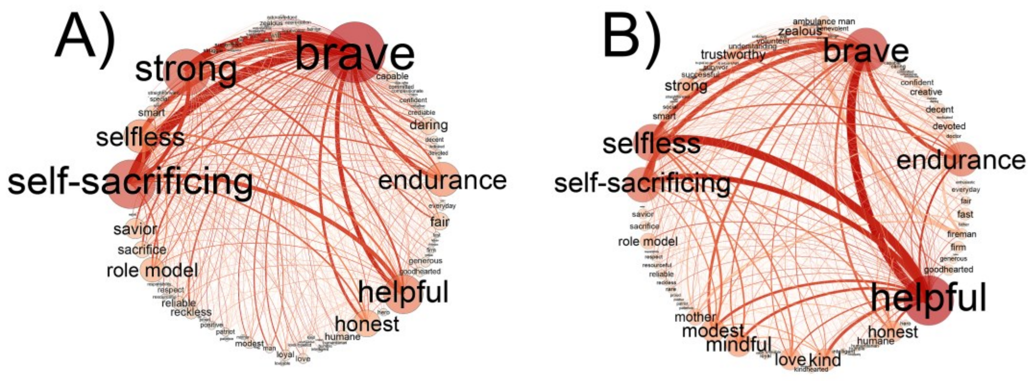



Figure 3.4. *Common associations in the social representations of Hero (A) and Everyday Hero (B).* The associations are arranged in a circular alphabetic order. The size of a node denotes the node strength and the thickness of an edge refers to the edge weight. The network was thresholded (edges below the value of 1 were deleted) for a better visualization.

## 3.4 DISCUSSION

The psychological investigation of heroism is relatively new. At this stage, inductive methods can shed light on its main aspects. Therefore, we examined the social representations of Hero and Everyday Hero by constructing two networks from their free word associations. The results showed that the social representation of Hero is more centralized and it cannot be divided into smaller units. The network of Everyday Hero is divided into five units and the significance moves from abstract hero characteristics to concrete social roles and occupations exhibiting pro-social values. We also created networks from the common associations of Hero and Everyday Hero. The structures of these networks showed a moderate similarity and the connections are more balanced in case of Everyday Hero.

We showed that Abric's [42], [65] theoretical framework about the central core vs. periphery associations is correspond to the scale-free organization of the association networks. Beyond the network interpretation of the classical Abric [65] model (central core of the social representation = the set of global hubs), we determined the modules and their modular hubs, which represent socio-cognitive patterns in the social representations based on Wachelke's theoretical assumptions [66].

Several scholars suggested multiple categories for heroism [32], [35], [37]. Hence, we expected that Hero would have a modular structure that could be interpreted in accordance with prior categorizations. However, the Hero network is non-modular. Contrary to Hero, the social representation of Everyday Hero includes five modules: Pro-social Heroism, Prototypical Hero, Heroic Roles, Everyday Context and Ordinary Heroism (Figure 3.3). In the Hero network, a large proportion of the nodes express social roles, occupations or social contexts. However, they are not organized into modules. This means that the presence of similar elements does not guarantee that they will form coherent units in the structure of the social representation.

Everyday Hero has more concrete contents which are mainly present in Heroic Roles, Ordinary Heroism and Everyday Context modules. In case of both Hero and Everyday Hero, the abstract values and characteristics are in accordance with previous inductive studies [31]–[34]. Changing the topic form “hero” to “everyday hero” resulted in not only a different content but it created an entirely new network structure i.e., different global hubs (central cores), modular organization and more balanced connections of heroic contents. While heroism depicts doing something extraordinary in an abstract manner, everyday heroism implies just doing the right thing and it points out the ordinary roles, occupations and contexts in which the heroic values can be exhibited.

# 4 THE MODULAR INVESTIGATION OF FREE WORD ASSOCIATION NETWORK CAN REVEAL POLARIZED OPINIONS

## 4.1 INTRODUCTION

In the present study, we aim to use one socially prominent issue as a cue (refugee/migrant labelled as migrant) to capture opinions shared by a social group (Hungarians) [65], [67], [68]. As a measure of public opinions, the free association method can be viewed as a semi-structured alternative between traditional questionnaires producing highly structured data and web-mining algorithms collecting large quantities of unstructured data. Hence, the free association method can overcome the predefined scope of questionnaires [69] since respondents can express freely their opinion, yet, it has the advantage of representative samples and fast data processing as opposed to several web-mining methods [70]. Traditionally, free association analysis focuses on consensual meaning (i.e., most frequent words and rankings) regarding a social object [65], [67], [68] and they do not focus on the polarization of opinions [71]–[74].

Different prior word association methods were introduced to distinguish the stable and recurrent associations from peripheral ones. Szalay and Brendt [75] developed the *Associative Group Analysis* approach of free associations. In this method, the early associations in a continued association task were found to have a high probability of being produced again during a retest. Previous studies in *social representation theory* [65], [76] argued that frequent associations are temporally stable and they refer to the consensual meaning regarding a given social object (a.k.a. the central core of the social representation). Alternatively, Kinsella and her coworkers [31] used the *prototype analysis of free associations*, in which most frequent associations (above a threshold) are considered as the consensual prototype of the social object in the perception of the social group.

Despite of the stable core of the representations, social issues can trigger opposite emotions, interpretations, attitudes, ideas and beliefs in a society, which can yield a polarized structure of public opinions. With sufficient data, it is possible to organize free associations not only along a core-periphery dimension, but to identify a more detailed structure with multiple major frames of interpretation in a society. These major frames do not exist only in terms of semantic processes, but also modulated by affective processes, in which emotions can control the recall of information [77]. In alignment, free associations are showed to be sensitive to emotional processes [71]–[74]. Thus, affective information can indicate the polarization of opinions and it helps to interpret association relations beyond lexical distance/semantic similarity.

Prior research used the available up-to-date technology to analyze free associations in relation to ideology [78] and attitude measurement [79]. Furthermore, Szalay and Deese (1978) provided an extensive summary of their pioneering factor analytic method for word associations. Apart from these works, to our best knowledge, no recent data-driven studies focused on the polarization of opinions with free associations. Therefore, we aimed to fill this methodological gap.

#### 4.1.1 Method demonstration: Public opinion of “migrants”

We aimed to demonstrate our method on public opinions about the recent migrant/refugee crisis which had a significant political and social effect in many European countries including Hungary. The increased number of refugees/migrants made migration one of the most prominent political and societal topics in the European Union. Similarly to these countries, in Hungary the intensive political, media and societal discourses merged the terms of migrant, refugee, asylum seeker and generated distinct attitudes toward the refugees/migrants [69], [80], [81]. According to a recent study including 15 European countries, *(i)* humanitarian concerns, *(ii)* anti-Muslim sentiments, and *(iii)* economic reasoning were the key factors in the perception of asylum seekers [69].

By combining affective information on free associations to refugees/migrants (i.e., emotional labels) with traditional attitude measurements such as *Perceived Outgroup Threat* [82]–[84], *Group Malleability* [85], and *Social Dominance Orientation* [86], we aim to demonstrate how free associations can reveal polarized opinions distinguished by their affective content and related attitudes.

### 4.1.2 Research goals and validation process

In this study, we aimed to demonstrate that co-occurrence statistic of associations can identify polarized opinions in the perception of refugees/migrants. For this reason, we constructed networks from free associations, in which associations were considered to reflect opinions and associations were connected base on their statistical co-occurrences (log likelihood ratio, LLR); thus, we refer to our free association networks as networks of co-occurring opinions (CoOp networks). We constructed such CoOp networks from multiple response free associations to the cue “migrant” in case of two independent and comprehensive samples in Hungary. Subsequently, we identified modules (densely connected subnetworks) of the CoOp networks.

We hypothesized that frequently co-occurring associations have higher emotional similarity (Hypothesis 1). To test this, respondents were asked to evaluate their own associations with emotion labels. The emotional similarity for every pair of associations was calculated based on the difference in the empirical distributions of their emotional labels. We calculated the correlation between emotional similarity values and co-occurrence connection values applying a permutation method (quadratic assignment procedure; QAP).

We assumed that the modules of the CoOp network reflect different opinions. Therefore, we statistically compared the attitude values (POT, GM, SDO) of participants whose associations belonged to different modules (Hypothesis 2). We assumed that explicit attitudes toward migrants (POT scores) can differentiate between modules more clearly than abstract construct related to perceived outgroup features (GM and SDO scores).

We tested the robustness of the CoOp networks (Hypothesis 3). The robustness tests were determined between the network parameters of the randomly divided data for Sample 1 and Sample 2 respectively. First, we aimed to test whether the LLR values were correlated between the randomly divided data in both samples (Hypothesis 3a). Second, we aimed to test whether the CoOp networks are more similar to each other—based on normalized mutual information—than a large number of randomized networks (null-models) with similar properties (Hypothesis 3b). Third, we aimed to test whether the exclusion of rare associations increase the robustness of our method due to the lower proportion of peripheral associations and the higher proportion of core associations (Hypothesis 3c).

## 4.2 METHODS

### 4.2.1 Participants and procedure

For our research purposes, two nationally comprehensive samples of Hungarian participants were recruited. The samples were nationally comprehensive in terms of gender, age, level of education, and type of residence for those Hungarians who use the Internet at least once a week. The participants were selected randomly from an Internet-enabled panel including 15,000 members with the help of a market research company in June 2016 (Sample 1) and in October 2016 (Sample 2). The samples were created with a random stratified sampling method among panelists in the online panel of the market research company with the average response rate 25%. Individuals were removed from the panel if they gave responses too quickly (i.e., without paying attention to their response) and/or had fake (or not used) e-mail addresses.

The final samples comprised $N_{S1}$ = 505 and $N_{S2}$ = 505 respondents who gave valid answers ($Male_{S1}$ = 247, $Female_{S1}$ = 258; $Male_{S2}$ = 249, $Female_{S2}$ = 256). Hungarians aged between 18 and 60 years in both samples ($M_{S1}$ = 40.19 years; $SD_{S1}$ = 11.78 years; $M_{S2}$ = 39.24 years; $SD_{S2}$ = 11.9 years). Regarding the highest level of education, 17.62%/17.82% (Sample 1/Sample 2) of the respondents had primary level of education, 0.4%/0.99% studied in secondary school without graduation, 26.14%/25.74% graduated from secondary school, 6.93%/7.13% studied in higher education and 48.91%/48.32% had higher education degree. Regarding the place of residence, 28.51%/28.71% of the respondents lived in villages, 31.49%/31.88% lived in towns, 21.39%/20.79% lived in county capitals and 18.61%/18.61% lived in the capital city.

The Research Ethics Committee of the local university approved this study. Data was collected via an online questionnaire. Participants were informed that the questionnaire was designed for measuring attitudes toward migrants. No other information was provided about the content and respondents could only see the actual task. All participants provided their written informed consent to participate in this study through a check-box on the online platform. The ethics committee approved this consent procedure. Respondents were assured of their anonymity and as a compensation the market research company drew gift cards among those who participated in the study.

### 4.2.2 Measures

**Multiple Response Free Association Task**. In this study an associative task was used, based on Abric's [41], [42] theoretical underpinnings and on Vergès' [54] and on Flament and Rouquette's [55] methodological assumptions. In the most of the social representation studies, a multiple response (a.k.a. continuous association task) response is applied with a limited number (three or five) of required associations. This method can reduce association chaining effects and inhibitory effects [87] that are more prevalent in open-ended association tasks.

Furthermore, open-ended association tasks can generate a lower number of average responses than a task with a pre-defined number of responses [31].

In the present case, the respondent's task was to write five words or expressions that comes into their mind regarding the word "migrant". However, in this study, we did not use the traditional methodology of social representations for identifying the central core and periphery or the density of the representations [41], [55], [88]. Instead, we used a network analytic method. From the perspective of large-scale semantic network studies, multiple response free association tasks generate strong and weak associations as well [87]. Classical social representation studies and network analytic association studies are closely related in terms of data collection procedure. The strong associations can constitute the central core of the representation and weak associations can belong to periphery [65], [87]. This associative task was the first question in the questionnaire for avoiding the influence of prior topic relevant questions.

**Emotional Labelling Task**. After providing all the 5 associations, respondents got back their associations one by one and were asked to provide two emotional labels to each of their own associations. We found that the negative-neutral-positive valence evaluation used in prior similar studies [88] is too constrained. Furthermore, frequently used affect measures as PANAS cannot be effectively used for the present goals as it included several irrelevant items (e.g.: active, strong, alert) and excluded relevant ones (e.g.: antipathy, empathy, anger). For this reason we reviewed basic emotion theories [89]–[92] to identify topic-relevant emotional labels. More precisely, the selection of the emotions was largely built on the ten basic emotions of Izard [92] and the eleven pairs of positive and negative emotion pairs of Robinson [91]. However, in a few cases basic emotions were described with synonyms to fit better to the cue. We used the following 20 emotional labels (differences from the original ones can be seen in parentheses): interest- alarm (anxiety), empathy-contempt, surprise-indifference, hope-fear, gratitude-anger, joy-sadness, calmness- relief (frustration), pride-shame, generosity-envy and love (sympathy)-hate (antipathy). Respondents could choose any two from the 20 emotional labels for each of their own associations (the labels did not appear as opposites).

DOI:10.15774/PPKE.ITK.2020.004

**Perceived outgroup threat (POT).** Perceived threat from refugees/migrant were assessed using seven items (Sample 1 α = .96, Sample 2 α = .96) which were translated from an implementation [82] of the Integrated Threat Theory [83]. The POT scale was translated to Hungarian according to protocol [93] and it was adopted to the contemporary Hungarian context based on a preliminary study. (E.g., “Migrants pose a physical threat to Hungarians”). Responses were made on 5-point Likert-type scales (1 = strongly disagree; 5 = strongly agree). The higher value indicates higher level of perceived threat from migrants.

**Group Malleability (GM).** We adopted a 4-item (Sample 1 α = .95, Sample 2 α = .94) version questionnaire [85] to assess respondents’ implicit assumptions on whether social groups are capable of development. The GM scale was translated to Hungarian according to protocol [93]. (E.g., "Groups can do things differently, but the important parts of who they are can't really be changed"). Respondents indicated their level of agreement using a 6-point Liker-type scale (1 = strongly disagree; 6 = strongly agree). The higher value indicates higher level of agreement with the concept of non-developing groups.

**Social Dominance Orientation (SDO).** The Social Dominance Orientation [86] questionnaire has 8 items (Sample 1 α = .83; Sample 2 α = .83) and it measures respondents’ degree of preference for inequality among social groups. The SDO measure was translated to Hungarian according to protocol [93]. (E.g., “Some groups of people are simply not the equals of others”). Respondents indicated their level of agreement using a 7-point Liker-type scale (1 = strongly disagree; 7 = strongly agree). The higher value indicates higher level of preferred inequality among social groups.

### 4.2.3 Preprocessing of associations

The preprocessing and lemmatization of the associations was carried out by four independent coders. Lemmatization is a linguistic process of grouping inflexions of a word into a single word (lemma) without conjugates. In other words, it is basically the grouping of words with the same stem. In the lemmatization process, two associations were merged in the following cases: *(i)* they had the same lemma (e.g. “refugee” and “refugees” were merged) [55]; *(ii)* they were semantically so close that the English translation could not distinguish between them (e.g. “stain” and “dirt”). Two associations were merged only if the coders could reach to a consensus.

### 4.2.4 CoOp Network Construction

Statistical relations among associations were defined based on their co-occurrences to identify connections. We used log-likelihood ratio (LLR) to assess co-occurrence connections between every possible association pairs [94]. For each possible pair of associations, we calculated the likelihood assuming statistical independence for their co-occurrence over the maximum likelihood of the observed co-occurrence:

$$\lambda(i,j) = \frac{L(j_n \cap i_n,\ i_n,\ \frac{j_n}{n})\ *\ L(j_n \cap \neg i_n,\ \neg i_n,\ \frac{j_n}{n})}{L(j_n \cap i_n,\ i_n,\ \frac{j_n \cap i_n}{i_n})\ *\ L(j_n \cap \neg i_n,\ \neg i_n,\ \frac{j_n \cap \neg i_n}{\neg i_n})},$$

where $j_n$ and $i_n$ denote the number of participants, who mentioned associations $i$ and $j$, $\neg i_n$ denotes the number respondents, who did not mentioned association $i$ and $n$ denotes the number of all participants. L(arg1,arg2,arg3) refer to the probability of a binomail distribution (L) in which *arg1* number of succes occured from *arg2* number of observations and *arg3* is the probability of *arg1*. More genenrally, the above formula measures the level of statistical dependence between $i$ and $j$ by testing whether the distribution of $j$ given that $i$ is present is the same as the distribution of $j$ given that $i$ is not present. The LLR from $\lambda$ is calculated as:

$$LLR(i,j) = \begin{cases} -ln\lambda(i,j), & \frac{j_n \cap i_n}{n} \geq \frac{j_n}{n} * \frac{i_n}{n} \\ ln\lambda(i,j), & \frac{j_n \cap i_n}{n} < \frac{j_n}{n} * \frac{i_n}{n} \end{cases}.$$

Therefore, the LLR between association $i$ and $j$ was positive (attractive) if their observed co-occurrence number was higher than the expected one and negative (repulsive) if their observed co-occurrence number was lower than the expected one. Basically LLR performs the same task as $\chi^2$-test parametric method without the requirement of normality. Multiplication of our LLR values by two can relate them to a $\chi^2$ distribution with the appropriate degrees of freedom.

We constructed CoOp networks, in which the nodes assigned by the associations and edge weights between nodes determined by LLR values. The nodes were the different associations from the total collections of associations. We ignored associations which occurred less than three times as these are not stable parts in the perception of the social object [65], or possibly related to idiosyncratic expressions, thus they do not belong to the social representation [95]. Furthermore, the removal of these nodes ensured higher robustness of the networks.

### 4.2.5 Affective Similarity

Every participant chose two emotional labels from the 20 options described above to characterize their affective relation to each of their association. The affective similarity between every pair of associations was calculated as:

$$Affective\ similarity\ (i,j) = 2 - \sum_{e=1}^{20} \left| \frac{E(e,i)}{\sum_{e=1}^{20} E(e,i)} - \frac{E(e,j)}{\sum_{e=1}^{20} E(e,j)} \right|,$$

where $i$ and $j$ are two different associations, $e$ is the emotional label (ranging from 1 to 20), and $E$ is a two dimensional matrix, in which each item $E(e,i)$ refer to the number of times $e$ emotion assigned by the respondents to association $i$. The similarity value of 2 indicates identical emotional labels while the similarity value of 0 indicates totally different emotional labels between two associations.

### 4.2.6 Module Detection

Both CoOp networks were divided into non-overlapping sets of densely linked associations (modules). A modularity maximization process [10] was applied to identify the modules of the networks. The original modularity formula is generalized [96] to deal with both the positive (attractive) and negative (repulsive) links:

$$Q = \frac{1}{v^{+} + v^{-}} \sum_{ij} [(w_{ij}^{+} + e_{ij}^{+}) - (w_{ij}^{-} + e_{ij}^{-})]\, \partial_{M_i M_j},$$

where $Q$ denotes the modularity value of a given partition of a network, $v^{+}/v^{-}$ denote the total positive/negative weights of the network, $w_{ij}^{+}/w_{ij}^{-}$ denote the positive/negative weights between node $i$ and $j$, $e_{ij}^{+}/e_{ij}^{-}$ denote the chance-expected positive/negative connections between node $i$ and $j$, $\partial_{M_i M_j}$ is an indicator function set to 1 if node $i$ and $j$ belong to the same module. The higher the modularity of a network partition, the higher the difference between the fraction of edges fall within the modules minus the expected fraction of edges fall within the same modules in a corresponding random network. The Louvain algorithm [20] was applied to identify the modular partition with the highest possible modularity, namely the highest ratio of edge weights inside the modules and lowest ratio of edge weights between modules. Therefore, the size and number of modules belong to the modular partition with the maximal modularity is parameter-independent and match with the structure of the network. In our case, it is extremely important to determine algorithmically the number of modules (i.e. number of opinion dimensions) which best describe the data. A drawback of modularity maximization that the resulting modular structure can change in each iteration [30] as the optimization process may stuck in local maxima. Therefore, a consensus partition was determined for the sake of higher reliability [97]. In the consensus partitioning process, first the consensus matrix was determined based on 5000 independent iterations of the Louvain algorithm. The edge weights between every pair of nodes in the consensus matrix determined based on the number of times two nodes fall into the same module. The consensus matrix was partitioned to non-overlapping modules 100 times by the Louvain algorithm. If the resulting 100 partitions were identical, then it was accepted as the consensus partition, otherwise the sets of 100 partitions were generated from the consensus matrix until the agreement. The average modularity score of the 5000 independent iterations and the consensus partition of the CoOp networks were determined for both samples.

### 4.2.7 Testing robustness

To demonstrate the robustness regarding co-occurrences of associations and the identified modular structure, we compared the LLR edges and modular structures of the randomly, equally divided data for Sample 1 and Sample 2 respectively. The split of the two samples were repeated 100 independent times. LLR edges and modular structures were determined for each half and compared as described below. Since associations were slightly different in the divided data, only the identical associations were compared in terms of LLR value and modular membership. The similarity of the LLR value between identical associations in the two samples was measured by Spearman's correlation. The significance of the correlation was determined by QAP [98]. A simple pairwise correlation between the LLR values of the two half would assume the independence of the edges, however a node in a network typically have similar connections, thus multiple similar edges belonging to one node can cause spurious correlation. QAP is a permutation procedure to eliminate the effect of interdependence between network edges belonging to a common node [98]. First, QAP determine the similarity between the LLR values of the two networks. This was done by Spearman's correlation in our case. Second, the edges of the CoOp network in one half of the data were randomly shuffled by permuting the rows and columns of the adjacency matrix in the same order. Third, the Spearman's correlation was calculated between the LLR values of the shuffled CoOp network and the LLR values of the CoOp network from the other half of the data. The second and third part of the QAP was repeated 5000 times and the absolute value of the simulated correlation coefficients were saved. The level of significance ($p_{QAP}$) was equal to the percentile of the simulated correlation coefficients reached the level of the correlation coefficient from the real data. Results were obtained as the average of the comparisons of the 100 independent runs.

The similarity of the modular structures was measured with the normalized mutual information (nMI) between the divided data of each samples. In order to determine whether the similarity between modular organizations indicates a nonrandom similarity, we compared the nMI calculated from the similarity of the original CoOp networks with the nMI calculated from the similarity of the null models. The simplest null-model is the Erdős-Rényi graph, in which the edges randomly rewired, however more sophisticated null model generation procedures can maintain certain parameters of the original network in the random network. Here, we generated edge, weight and strength preserving random networks [19] for both Sample 1 and Sample 2. The generation of the null model consists of two steps. First, the randomization of the network is done by connection-switching method [99] in a way that preserves positive and negative degrees of the nodes. Then, the weights are allocated and iteratively rearranged to converge to the weight distribution of the original network [19]. 5000 null models were generated and the modular structures of the null models were determined. The similarity (nMI) of the modular structures (only identical association included) was calculated for the null models. The observed nMI values were compared to the nMI values derived from the null models by independent t-test for each sample.

To demonstrate that higher number of observations offer a higher robustness of our method, we iteratively raised the threshold of the ignored associations from the default 3 to 10. The similarity of the LLR edges and modular structures were calculated for each threshold between the divided samples.

### 4.2.8 Statistical analysis

All statistical analyses were performed with MATLAB version R2014b (The MathWorks Inc, Natick, MA). The applied network measures are all available at https://sites.google.com/site/bctnet/ [100]. Differences of the POT scores were determined by an independent t-test between Sample 1 and Sample 2.

We calculated the correlation with a permutation test based on QAP [98] to test whether cognitive attraction is related to affective similarity and cognitive repulsion is related to affective dissimilarity. In the QAP procedure, we moderated the effect of near zero co-occurrence connection values. On one hand, many near zero LLR values were expected between associations never mentioned together, but these association pairs could be characterized by very heterogeneous affective similarity values. On the other hand, moderating the effect of the numerous near zero connections can generate a more balanced LLR data for the correlation analysis, in which the low and high LLR values have similar sampling. Hence, co-occurrence connection values were divided into 100 equal intervals in which the values were averaged. This way, the large number of data points representing near zero co-occurrence values were reduced into averages of a few intervals. The affective connection values were averaged for the association pairs which belonged to a given interval of the co-occurrence connection values. All correlation coefficient was calculated between these averaged values.

The final test of our method was to demonstrate that we can differentiate the modules in CoOp networks according to the attitudes toward refugees/migrants. Respondent were assigned to the CoOp modules to which the majority of their associations belonged. Respondents were compared by pairwise independent weighted t-test on their attitude scores between every pair of modules. Weighted attitude score means (WAM) and weighted attitude score variance (WAV) were calculated for each module (M) for weighted t-tests:

$$WAM_M = \frac{\sum_{i\epsilon M}(AttitudeScore_i * AssociationNumber_i)}{\sum_{i\epsilon M}(AssociationNumber_i)},$$

$$WAV_M = \frac{\sum_{i\epsilon N} AssociationNumber_i * (AttitudeScore_i - WeightedAttitudeMean_M)^2}{\sum_{i\epsilon N} AssociationNumber_i},$$

where $i$ is a respondent assigned to module $M$. Attitude scores of a respondent ($AttitudeScore_i$) were weighted equally to the number of their associations that belonged to the given module ($AssociationNumber_i$). A respondent was discarded from the attitude analysis if he or she could have been assigned to more modules with equally maximum weights. The statistical procedure was conducted on Sample 1 and Sample 2 separately.

Since our study was exploratory, we carried out statistical power estimation for a theoretically medium effect size (Cohen's d = 0.5) which we determined as the indicator of a considerable opinion difference between respondents assigned to two given modules. We concluded that 0.8 power can be achieved if sample sizes are 64 (Power was determined for Cohen's d = 0.5 with alpha = 0.05. In the calculation, normal distributions were assumed with a mean difference equal to 0.5 and with standard deviations equal to 1.).

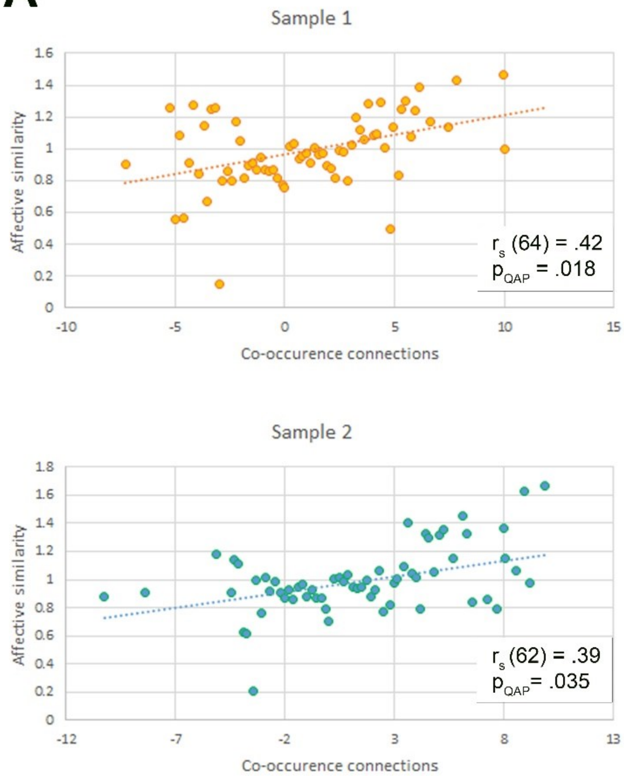
## 4.3 RESULTS

The total number of different associations was 1067 in case of Sample 1 and 1099 in case of Sample 2. After the lemmatization, the number of different associations decreased to 597 in case of Sample 1 and 533 in case of Sample 2. The number of associations mentioned at least three times - thus included in the network analysis – was 156 in case of Sample 1 and 163 in case of Sample 2. Sample 1 and Sample 2 had 114 identical associations. Thus, the analysis was performed on 1966 association tokens in Sample 1 and on 2023 association tokens in Sample 2. The POT scores had no significant overall difference between Sample 1 ($M$ = 3.33, $SD$ = 1.37) and Sample 2 ($M$ = 3.43, $SD$ = 1.36).

### 4.3.1 CoOp Connections and Affective Similarity (Hypothesis 1)

Significant correlations were found between the co-occurrence and affective similarity values (Figure 4.1.) of Sample 1 ($r_s$ $(64)$ = .42, $p_{QAP}$ = .018) and Sample 2 ($r_s$ (62)= .39, $p_{QAP}$ = .035).

**A**




**B**

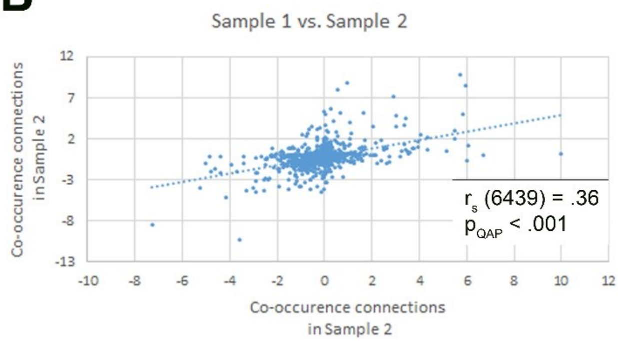


Figure 4.1. *Correlation between affective similarity values and co-occurrence connections (A) and correlation between the identical co-occurrence connections in Sample1 and Sample 2 (B).* (A) The x-axis shows the co-occurrence connection values and the y-axis shows the affective similarity values. The x-coordinates of the data points represent the averages of the co-occurrence connection values in each of the 100 equal intervals. The association pairs were determined in each interval and their affective similarity values were also averaged. The y-coordinates of the data points represent these averaged affective similarity values. (B) The similarity of the co-occurrence connections between identical associations in the two samples was measured by Spearman's correlation. The x-axis shows the co-occurrence connections in Sample 1 and the y-axis shows the co-occurrence connections in Sample 2.

DOI:10.15774/PPKE.ITK.2020.004

### 4.3.2 CoOp Modules

We labelled the modules according to the two most frequent associations (see Figure 4.2.).

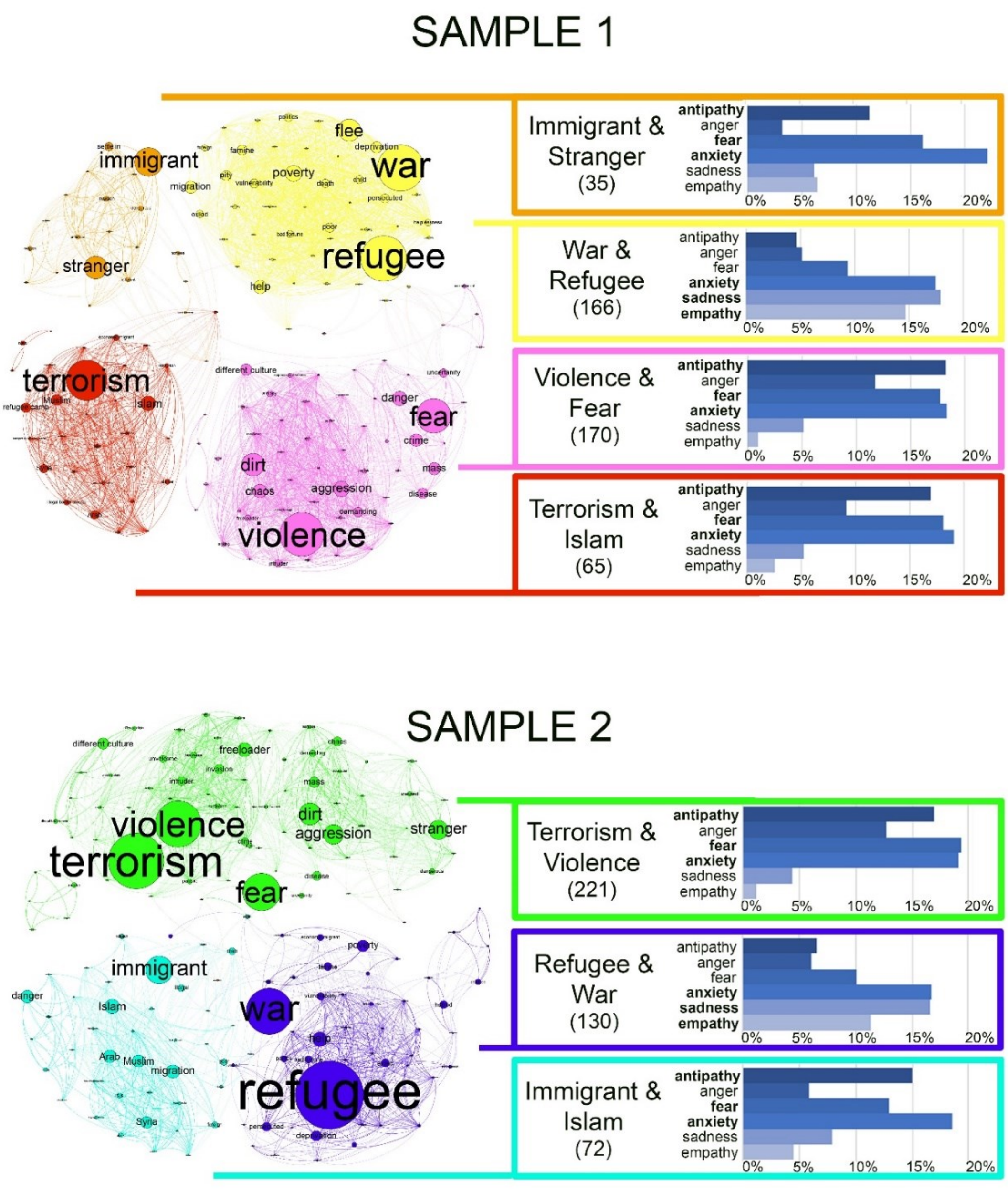



Figure 4.2. *The modules of the CoOp networks.* Each module is visualized with different colors. The size of a node and its label is proportional to the frequency of the given association. An edge means that two associations fall into a common module in the consensus partitioning procedure at least 40%. Both sample is displayed by “Yifan Hu Proportional” layout algorithm

[2] implemented in Gephi [3]. Additional information about each module is shown in a box colored identically to the corresponding module. The box contains the label of the module referring to the two most frequent associations in a given module. The number of respondents assigned to a given module is displayed below the label in parentheses. The percentages of emotional labels for every module are presented on bar charts. The percentage of the six most frequent emotions (antipathy, anger, fear, anxiety, sadness, empathy) are shown in detail. The three most frequent emotions are displayed with bold letters.

The modularity value was 0.24 for the CoOp network of Sample 1 and it was 0.23 for the CoOp network of Sample 2. The CoOp network of Sample 1 was divided into four modules while the CoOp network of Sample 2 was divided into six modules. However, in Sample 2, three modules from the identified six modules contained only a single word, each mentioned by a few respondents ("assassination", "unity", and "death"). We did not include these modules into further analysis, so the final number of modules was three in case of Sample 2.

### 4.3.3 CoOp Modules and POT scores (Hypothesis 2)

In case of Sample 1 and Sample 2, all pairwise comparisons of the modules showed significant differences in the POT score (Figure 4.3.).

DOI:10.15774/PPKE.ITK.2020.004

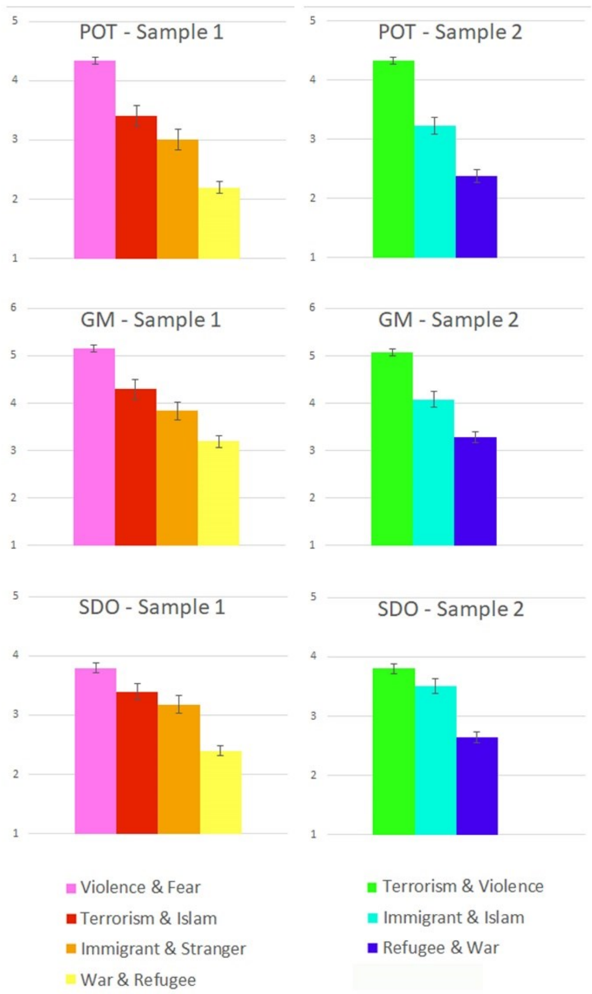



Figure 4.3. *Perceived Outgroup Threat (POT), Group Malleability (GM) and Social Dominance Orientation (SDO) scores of the modules in Sample 1 and Sample 2.* Bars represent the mean of the POT = *Perceived Outgroup Threat*, GM = Group Malleability and SDO = *Social Dominance Orientation* scores for every module. Standard error was presented on the bars. All pairwise comparisons of the modules showed significant differences in POT scores. See detailed POT analysis results below.

In case of Sample 1, respondents assigned to *War & Refugee* module ($M$ = 2.25, $SD$ =1.20) showed significantly lower POT score than respondents assigned to the *Immigrant & Stranger* ($t(199)$ = -3.23, $p$ < .001, d = 0.57), *Terrorism & Islam* ($t(229)$ = -6.65, $p$ < .001, d = 1.01) and *Violence & Fear* ($t(334)$ = -18.49, $p$ < .001, d = 2.03) modules. The *Immigrant & Stranger* module ($M$ = 2.93, $SD$ =1.12) had significantly lower POT score than the *Terrorism &*

*Islam* (*t(98)* = -2.27, *p* = .013, d = 0.45) and *Violence & Fear* modules (*t(203)* = -6.91, *p* < .001, d = 1.62). The *Terrorism & Islam* module (*M* = 3.50, *SD* =1.31) had significantly lower POT score than the *Violence & Fear* module (*t(233)* = -4.64, *p* < .001, d = 0.84). The *Violence & Fear* module showed the highest POT score (*M* = 4.30, *SD* =0.78). In case of Sample 1, the statistical comparisons involving the *Immigrant & Stranger* and the *Terrorism & Islam* did not have sufficient power.

In case of Sample 2, respondents assigned to the *Refugee & War* module (*M* = 2.10, *SD* = 1.21) had significantly lower POT score than respondents assigned to the *Immigrant & Islam* (*t(200)* = -6.92, *p* < .001, d = 0.99) and *Terrorism & Violence* (*t(349)* = -18.71, *p* < .001, d = 2.27) modules. The *Immigrant & Islam* module (*M* = 3.25, *SD* =1.13) had significantly lower POT score than the *Terrorism & Violence* modules (*t(291)* = -7.39, *p* < .001, d = 1.17) The *Terrorism & Violence* module showed the highest POT score (*M* = 4.31, *SD* =0.83). In case of Sample 2, all comparisons can be considered with a power 0.8.

Similarly, to the POT scores, the GM and SDO scores were compared across the modules. In the most of the cases—similarly to POT—these measure could differentiate the modules. Here, we only give a short overview about the few exceptions in which we did not get significant difference or sufficient power. In case of Sample 1, the comparisons of every module gave a significant difference for the GM analysis, but the *Immigrant & Stranger* and the *Terrorism & Islam* comparison did not have sufficient power. In case of Sample 2, all comparisons were significant with a sufficient power. In case of Sample 1, the comparison of the modules along the SDO scores failed to detect a significant difference between the *Immigrant & Stranger* and the *Terrorism & Islam* modules and the comparison of the *Terrorism & Islam* and *Violence & Fear* modules did not have sufficient power. In case of Sample 2, the comparison of the modules along the SDO scores were all gave significant differences, although the comparison of the *Immigrant & Islam* and *Terrorism & Violence* modules did not reach the sufficient power. In sum, POT, GM, and SDO showed very similar patterns in the most of the cases.

#### 4.3.4 Robustness (Hypothesis 3)

To test the robustness of our method, we derived edge level and modular level comparisons between the randomly, equally divided data of Sample 1 and Sample 2 respectively. The LLR level comparison performed by the correlation of the LLR values between the identical association pairs. We have found a significant correlation between the LLR values of the separated data for 100 independent runs in Sample 1 (mean $r_s$ (2209) = .26, mean $p_{QAP}$ < .001) and Sample 2 (mean $r_s$ (2924) = .28., mean $p_{QAP}$ < .001). The modular level similarity was determined by the nMI value of the modular membership of the identical associations between the separated data of Sample 1 and Sample 2. The similarity between the modular structures of the divided data were significantly higher than the similarity of the corresponding null models (for Sample 1: $M_{real}$=0.3, $SD_{real}$=0.056, $M_{null}$=0.21, $SD_{null}$=0.042, t(198)=12.52, p<.001; for Sample 2: $M_{real}$=0.26, $SD_{real}$=0.058, $M_{null}$=0.2, $SD_{null}$=0.04, t(198)=8.9, p<.001).

The changes of the LLR level and modular level similarity between the divided data were determined by ignoring associations occurring less than a given threshold. The threshold was iteratively raised from the default 3 to 10. Strong and significant correlation was detected between the threshold and the LLR (for Sample 1 $r_s$ (6) = .91, p = .002; for Sample 2 $r_s$ (6) = .97, p < .001) and modular (for Sample 1 $r_s$ (6) = .98, p < .001; for Sample 2 $r_s$ (6) = .96, p < .001) level similarity. Ignoring sparse associations from the analysis could raise the edge and modular level similarity between the divided data in Sample 1 and Sample 2 respectively. Details about the edge and modular level similarity for every threshold is presented in Figure 4.3. Here, we only present the LLR level similarity (for Sample 1 mean $r_s$ (188) = .35, mean $p_{QAP}$ < .001; for Sample 2 mean $r_s$ (208) = .36, mean $p_{QAP}$ < .001) and modular level similarity (for Sample 1 mean nMI = .41; for Sample 2 mean nMI = .34) for the analysis ignoring associations occurred less than 10 times (Figure 4.4.).

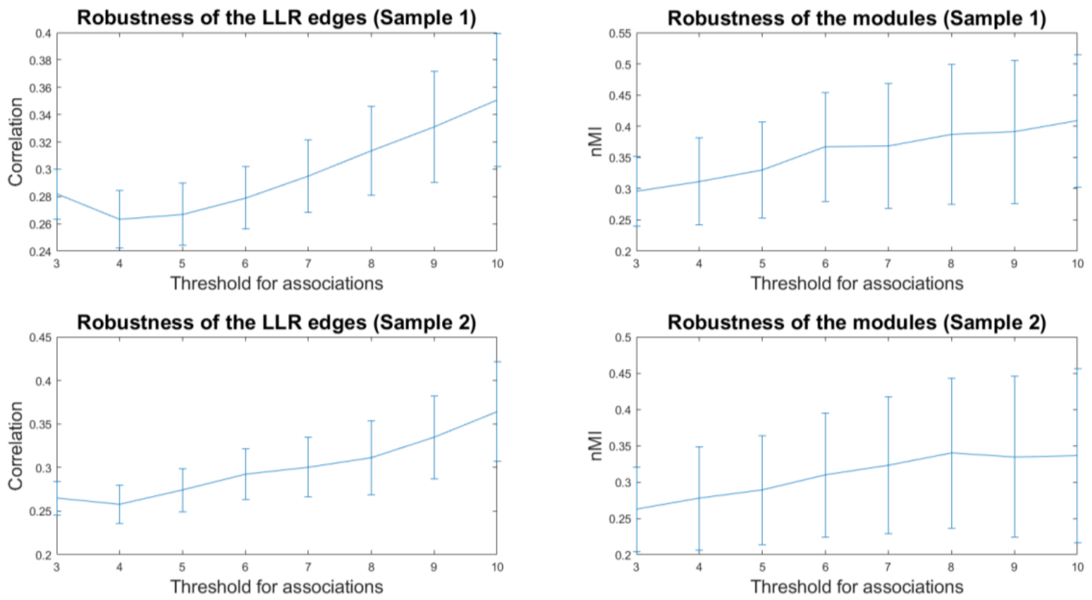



*Figure 4.4. Correlation between the robustness and the exclusion of rare associations from the analysis.* The x-axis shows the minimal number of occurrence of an association. Below that occurrence number, an association was excluded from the analysis. The y-axis shows the LLR (correlation) and modular (nMI) level similarity between the randomly divided samples. Error bars represents the standard deviation of the 100 independent runs. Exclusion of rare associations was resulted in higher LLR similarity and higher modular similarity in Sample 1 and Sample 2.

## 4.4 Discussion

In this study, we aimed to introduce and validate a method which identifies groups of associations reflecting distinct attitudes and emotions toward demonstrative cue: migrants. In line with Hypothesis 1, the co-occurrence of the associations (CoOp networks) reflected the emotional similarity between the associations. In line with Hypothesis 2, the distinct cohesive structures of associations (CoOp modules) reflected different results on the POT, GM and SDO measures. For example, between modules reflecting on *violence* (Violence & Fear, Terrorism & Violence) and *refugee* (War & Refugee, Refugee & War) always demonstrated significant differences in the three measures (POT, GM, and SDO). In line with Hypothesis 3, the modular structures of CoOp networks showed considerable robustness in the two independent samples. In sum, the present results demonstrated that analyzing the modular organization of CoOp networks can be an inductive tool for identifying the most important dimensions of public opinions about relevant social issues.

CoOp networks can be seen as a subtype of large-scale semantic networks [48], [101], [102]. Semantic networks are built from multiple cues and organized by constant lexical relations. Our study demonstrated that co-occurrences of multiple free word associations can also follow affective similarity patterns regarding a social issue. This is in line with cognitive studies on roles that emotions play in mental process e.g., message acceptance/rejection and information recall [77], [103]. Our results also highlight that module detection in CoOp networks yields a psychologically meaningful mapping of context behind attitudes. The modular membership of the associations creates a context for the interpretation of each individual association. Furthermore, the jointly interpreted associations can link the attitudes to the relevant context. More generally, consistent patterns in individual association sequences can reveal the most prominent frames of opinions regarding a social issue.

The polarization of opinions was consistent in the two samples with a pole indicated by terms such as “Refugee”, “War” or “Help” and a pole indicated by terms such as “Violence”, “Fear” or “Terrorism”. Furthermore, modules reflecting these poles comprised the majority of all the respondents in both samples. The *Violence & Fear* (Sample 1) and *Terrorism & Violence* (Sample 2) modules had the highest POT scores. These modules indicate explicit hostility [104] such as labeling refugees/migrants as morally inferior [105] (e.g., “dirt”, “lazy”, “demanding”, “freeloader” associations) or emphasizing perceived threats (e.g., “terrorism”, “crime”, “invasion” associations) [80], [81], [83]. The *War & Refugee* (Sample 1) and *Refugee & War* (Sample 2) modules reflect *(i)* humanitarian concerns and show the lowest POT scores compared to other modules. The scores and the contents of these modules indicate that considering refuges/migrantsas refugees who are forced to leave their homes (e.g., “war”, "famine", "death", "flee" associations) is linked to social solidarity (e.g., "help", "pity" associations) [106]–[108].

Compared to Bansak et al. (2016), we could identify modules referring to *(i)* humanitarian concerns (the *War & Refugee* (Sample 1) and *Refugee & War* (Sample 2) modules) and *(ii)* the anti-Muslim sentiment (*Terrorism & Islam* (Sample 1) and *Immigrant & Islam* (Sample 2) modules). Interestingly, we could not differentiate any modules reflecting economic reasoning, which indicates that no subgroup of the respondents considered refuges/migrants solely as an economic threat. This is in line with previous research which showed that only economic considerations are not sufficient explanations to the perception of refugees/migrants [109], [110]. Humanitarian concerns are unequivocally present in Hungarians' perception of refugees/migrants consistently with Bansak et al. (2016).

The LLR values between the identical associations in Sample 1 and Sample 2 showed significant correlation and the modular structures of the CoOp networks referring to a relative robustness compared to the null model in both samples. Our method also showed a higher robustness in case of the frequent associations compared to the rare associations. From an information theoretical point of view, these results suggest that frequent associations resulted in a more stable pattern of co-occurrences. Following this logic one can reach the desired robustness by increasing the sample size. From the social psychological point of view, frequent associations more likely to belong to the core structure referring to a higher reliability over time than rare peripheral associations [31], [65]. It is possible that complex influential factors such as media can more likely affect the peripheral elements of the representation. This is in line with Abric's [65] description of progressive transformation in social representations. In sum, reducing the effect of influential factors and the sparsity of the data by excluding rare associations increased the robustness of the results, which suggests the reliability of the applied methodological framework.

The measure on word co-occurrence and the appropriate clustering method were selected based on the following considerations. First, frequency of associations - similarly to word occurrence in a corpus – had a power law function [112], thus an adequate similarity measure should deal with associations occurring sparsely. The LLR was successfully used in previous text processing designs to measure typical word co-occurrences in large corpus of sentences [94], [113]. In our case, a five-associations-long response sequence was considered as a sentence and the typical pattern of co-occurrence across the sequences was measured by the LLR. In our case, a five-associations-long response sequence was considered as a sentence and the typical pattern of co-occurrence across the sequences was measured by the LLR. The first advantage of LLR that it does not depend on normality as well as it allows the comparison of the co-occurrence of both rare and common associations [94]. Second, the LLR can handle the attraction and repulsion of association pairs based on the expected number of co-occurrence in case of independence for two associations. In contrast, a simple co-occurrence count can only distinguish between weak and strong connections. For example, simple co-occurrence count gave a relatively high value (i.e. strong connection) between the *Violence* and *Refugee* associations (6/13 in Sample 1/Sample 2) compared to other co-occurrence values in our data. However, based on the frequency of the two associations (93/99 for *Violence* and 97/146 for Refugee in Sample1/Sample2) expected co-occurrence should have resulted in a higher co-occurrence count (17/28 in Sample 1/Sample 2). The expected co-occurrence was related to observed co-occurrence count in the LLR formula and resulted in a high negative value (i.e. strong repulsive connections) (-7.27/-8.4 in Sample 1/ Sample 2). Third, LLR can be related to the cumulative distribution of $\chi^2$-test with one degree of freedom, hence one can calculate the significance of the co-occurrences. The modularity clustering procedure can give a partitioning which match with the structure of the network without selecting parameters. Most importantly, the size and number of modules are not predefined (like K-means clustering) or assigned by the researcher based on a dendrogram (like Ward's method). The parameter free and unconstrained characteristics of the modularity formula ensures the data-driven clustering of the associations.

The major limitation is that connections of the CoOp networks were often created from relatively few observations. As a consequence of sparsity, it is important to be careful with interpretations based on a single connection and rely more on the modules which were proved to be a meaningful indicators of different attitudes. Furthermore, the modular investigation of the CoOp network is as an exploratory analysis. Therefore, a minimum number of respondents cannot be guaranteed in each module. As an example, three modules were identified containing only one association in case of Sample 2 ('assassination', 'unity', and 'death'). As a consequence, we cannot provide a lower bound (holding for all comparisons) for statistical power. However, small modules can be filtered according to future study designs to achieve a desired statistical power for a given effect size.

We would like to provide a few recommendations for further similar studies to choose an appropriate sample, cue and additional questionnaires for the associations. Large and diverse sample is recommended to increase the robustness of the method (increased threshold for ignoring associations increase the robustness) and to capture the heterogeneity of opinions in the target group. Selection of the appropriate cue for the study is crucial. Most importantly, the respondents should have an elaborated opinion about the provided cue. For example, there should be an active group-level discourse about the topic in the target group. In our case, during data collections migration was a prominent topic in the political public and media discourses for the Hungarian population. Indefinite cues should be avoided; different respondents can easily provide different meanings for a cue, hence the segregation of the CoOp modules can easily reflect to semantic differences. For instance, the cue *play* can refer to *sport*, *music*, or *games* [114]. An appropriate cue should be a single word. Even for compound words certain respondents may associate to the first word as others to the second word. Further studies can also guide associations by manipulating the instructions. For example, simply asking "climate change" as a cue may be result in a CoOp module structure in which technical terms, beliefs and associations for "climate" are segregated. If one is interested whether different beliefs for climate change may change, the instruction could be restricted to opinions. For the preprocessing of the associations, automatized lemmatization methods are available in case of English responses; e.g., the Porter's algorithm [115]. For sake of higher reliability, we recommend further studies to apply additional questionnaires to test the relevance of the CoOp modules. Although we demonstrated that only the co-occurrence analysis of associations can yield meaningful results, we only tested and validated for a single cue. Based on our results, not only an explicit questionnaire about the cue (POT), but questionnaires measuring more abstract constructs (GM and SDO) can differentiate between CoOp modules. This suggests that a broad spectrum of dependent questionnaires is appropriate for testing the modules. Emotional similarity between associations provided a validation metric for LLR values. However, further studies can use the emotional similarity between associations to construct networks and modules. Applying the label of the associations for a similarity measure can help to link directly associations to certain emotional constructs and also gives a less sparse data than co-occurrence measures. Emotional labeling of the association may be applied to sentiment analysis of thematic corpora. Similarly to the Sentiwordnet, in which emotions spread through the logical connections of the Wordnet's synsets [116], we may expand our findings by adding emotional labels to specific seed words and based on their co-occurrence spreading the emotions in the corpus. As the human coders could judge words' sentiment related to specific topic and the co-occurrence of the words corresponds the topic, this approach may generate a context dependent sentiment analysis approach as an opposite to the general solution of the Sentiwordnet. It is also important to emphasize that emotional labeling of the associations can be changed to other appropriate labels (e.g.: valence, etc.). However, we recommend applying a diverse set of potentially relevant labels to maintain the unrestricted nature of the association task.

Future studies could investigate network topological parameters to determine how in individual associations are distributed across modules. These parameters can link the identified modules to individual response patterns. Studying the relation between individual response patterns and the higher level structure can relate the group-level opinion dynamics to cognitive processes such as biased assimilation [117] or socio-psychological differences such as SDO or

DOI:10.15774/PPKE.ITK.2020.004

GM in our case. In future studies, the influence of a social object on association relations can be assessed by comparing these relations to a "resting state" baseline of the mental organization among lexical concepts such as large-scale semantic networks [48], [101], [102]. Furthermore, constructing questionnaires from data-driven constructs (CoOp modules) can help to converge theoretical and observed dimensions regarding a social object. For example, as opposed to previous studies which found emphasis on economic concerns if respondents' attention was explicitly directed on them, economic concerns did not appear as a governing factor in free individual opinions about refugees/migrants. Cross-cultural studies can also apply CoOp network analysis to study how corresponding social objects vary in different cultures and refine questionnaires according to specific cultures [118].

In sum, traditional questionnaires without inductive focus can hardly reflect the dynamic contents constituting a social object although these can form a link between social constructs and actual actions [65]. The inductive nature of CoOp modules can contribute to the classification of changing contents that constitute a social object and it can provide a data-driven representation of characteristic social frames at a particular time and space.

# 5 SUMMARY

## 5.1 THESIS GROUP I. APPLICATION OF THE MODULARITY VALUE ON LOCAL SCALE IN FUNCTIONAL BRAIN NETWORKS

Graph theoretical analyses have demonstrated that brain functional networks have a modular structure. In general, integration within a module allows efficient local processing, while segregation between modules may be necessary to avoid internal or external noise. Describing the changes of modularity in different clinical states can help to understand the alterations of the global processing mechanisms. In the current study we applied the modularity in a local scale. I tested the proposed methods on resting-state fMRI data of 20 young and 20 elderly subjects.

**Thesis Ia. I measured the extent of the modularity decrease caused by changing the community membership of brain regions. The occipital regions showed the highest local modularity difference between the young and elderly group, but the local modularity value of the regions in the default mode network was slightly affected by aging.** I found increased local modularity of the occipital regions in the young, which is in line with the previously reported increased segregation of the visual cortex in the young. I determined that regions of the dorsal attention network were characterized with an increased local modularity in the elderly which suggest that these regions preserve the modular structure in advanced age.

**Thesis Ib**. **I evaluated the group-level differences in the community membership of the brain regions between the two age groups by applying approximation node shifts. I showed that the changes of the community assignment of the young 'fronto-temporal' module have a significant impact on the modular organization of the brain networks in advanced age. I** identified 4 modules in the young and 3 in the elderly group. I found the merging of the DMN to the 'fronto-temporal' module in the elderly group.

## 5.2 THESIS GROUP II. MODULAR INVESTIGATION OF FREE WORD ASSOCIATION NETWORKS

Free word association is a widely used technique in market research and psychology to encourage respondents to express openly their underlying motivations, beliefs, attitudes or feelings regarding a specific topic. This technique enhances the unconstrained expressions of respondents and overcomes limitations of predefined questionnaires, but it produces scattered outcome. I developed and validated a network solution for data-driven grouping of freely expressed individual associations. I demonstrated that analyzing the modular organization of association networks can be a tool for identifying the most important dimensions of public opinions about a relevant social issue without using pre-defined constructs.

**Thesis IIa. I constructed semantic networks from multiple free word associations for the cue “Hero/Everyday Hero”. I experimentally evaluated global parameters of the networks to describe relevant socio-psychological constructs.** The scale-free properties of the association networks correspond to the classical core/periphery social representation structure. In the statistically modular network, I determined the modules and their modular hubs, which may represent socio-cognitive patterns in the social representations.

**Thesis IIb. I demonstrated in case of the “migrant” cue that the co-occurrence based relations of free word associations reflected emotional similarity and the modules of the association network were validated on well-established measures.**

**Thesis IIc. I tested and showed the LLR and modular level robustness of the CoOp networks. Furthermore, I demonstrated that exclusion of rare associations increases the robustness of the modular structure, which suggesting the validity of the applied framework.** First, I showed the correlation between the LLR values between the randomly divided data for both samples. Second, I showed the modules of the CoOp networks are more similar to each other—based on normalized mutual information—than a large number of randomized networks (null-models). Third, I showed that the exclusion of rare associations increases the robustness of the LLR values and the modular structure.

# 6 THE AUTHOR’S PUBLICATIONS

## 6.1 PUBLICATIONS RELATED TO THE PRESENT THESIS

### 6.1.1 Papers:

File, B., Klimaj, Z., Somogyvári, Z., Kozák, L. R., Gyebnár, G., Tóth, B., ... & Molnár, M. (2016). Age-related changes of the representative modular structure in the brain. In Pattern Recognition in Neuroimaging (PRNI), 2016 International Workshop on (pp. 1-4). IEEE.

File, B., Keczer, Z., Vancsó, A., Bőthe, B., Tóth-Király, I., Hunyadi, M., ... & Orosz, G. (2019). Emergence of polarized opinions from free association networks. Behavior research methods, vol. 51, no. 1, p. 280-294.

Keczer, Z., File, B., Orosz, G., & Zimbardo, P. G. (2016). Social Representations of Hero and Everyday Hero: A Network Study from Representative Samples. PloS one, vol. 11, no. 8, p. e0159354.

### 6.1.2 Oral presentations:

Bálint File, Dániel Gerő, Marco Bueter, Zsolt Keczer, Gábor Orosz, Zoltán Somogyvári, Júlia Góth, Noreen Hinrichs, Matteo Müller, István Ulbert (2019): Véleménykinyerés szabad szó asszociációs hálózatok moduláris vizsgálatával, Doktoranduszok Országos Szövetsége, Debrecen

File, B., Keczer, Z., Vancsó, A., Bőthe, B., Tóth-Király, I., Hunyadi, M., ... & Orosz, G. (2018). Polarized Opinions from Free Association Networks, Singapore Conference on Applied Psychology 2018, Singapore

### 6.1.3 Poster presentations:

Bálint File, Zsolt Keczer, Gábor Orosz, Beáta Bőthe, István Tóth-Király, Anna Vancsó, Márton Hunyadi, Adrienn Ujhelyi, István Ulbert, Júlia Góth (2017): Attitudes toward migrants: free word association networks bridging social and cognitive representations. 18th General Meeting of the European Association of Social Psychology, Granada, Spain

## 6.2 PUBLICATIONS NOT RELATED TO THE PRESENT THESIS

### 6.2.1 Papers:

Gero, D., File, B., Justiz, J., Steinert, R. E., Frick, L., Spector, A. C., & Bueter, M. (2019). Drinking microstructure in humans: A proof of concept study of a novel drinkometer in healthy adults. Appetite, 133, 47-60.

File, D., File, B., Bodnár, F., Sulykos, I., Kecskés-Kovács, K., & Czigler, I. (2017). Visual mismatch negativity (vMMN) for low-and high-level deviances: A control study. Attention, Perception, & Psychophysics, 79(7), 2153-2170.

Kardos, Z., Tóth, B., Boha, R., File, B., & Molnár, M. (2014). Age-related changes of frontal-midline theta is predictive of efficient memory maintenance. Neuroscience, 273, 152-162.

Tóth, B., Kardos, Z., File, B., Boha, R., Stam, C. J., & Molnár, M. (2014). Frontal midline theta connectivity is related to efficiency of WM maintenance and is affected by aging. Neurobiology of learning and memory, 114, 58-69.

Brigitta Tóth, Bálint File, Roland Boha, Zsófia Kardos, Zoltán Hidasi, Zsófia Anna Gaál, Éva Csibri, Pál Salacz, Cornelis Jan Stam, Márk Molnár: EEG network connectivity changes in mild cognitive impairment — Preliminary results, International Journal of Psychophysiology, Available online 6 February 2014, ISSN 0167-8760, http://dx.doi.org/10.1016/j.ijpsycho.2014.02.001.

Boha R., Tóth B., Gaál Zs. A., Kardos Zs., File B., Molnár M. Spectral, phase-synchronization, and graph theoretical EEG changes related to mental arithmetics. IDEGGYÓGYÁSZATI SZEMLE/CLINICAL NEUROSCIENCE 66(5–6): 175–183 (2013)

Molnár M., Tóth B., Boha R., Gaál Zs. A., Kardos Zs., File B., Stam J. C. Aging effects on ERP correlates of emotional word discrimination CLINICAL NEUROPHYSIOLOGY 24: 1986–1994 (2013)

### 6.2.2 Oral presentations:

Bálint File, Domonkos File, Flóra Bodnár, István Sulykos, Krisztina Kecskés-Kovács, Zsófia Kardos, Brigitta Tóth, Roland Boha, Márk Molnár, Ágota Tóth, Dániel Fabó, István Ulbert, István Czigler: Resting state brain networks predict the characteristics of visual evoked potentials, (2016), NEURONUS 2016 IBRO & IRUN Neuroscience Forum, Krakow, Poland.

### 6.2.3 Poster presentations:

Bálint File, Emília Tóth, Virág Bokodi, Zsolt Keczer, László Entz, István Ulbert, Dániel Fabó (2017): “Functional connectivity networks of cortico-cortical evoked potentials”, Fens Regional Meeting, Pécs, Hungary

Bálint File, Zsolt Keczer, Gábor Orosz, Beáta Bőthe, István Tóth-Király, Anna Vancsó, Márton Hunyadi, Adrienn Ujhelyi, István Ulbert, Júlia Góth (2017): Attitudes toward migrants: free word association networks bridging social and cognitive representations. 18th General Meeting of the European Association of Social Psychology, Granada, Spain

Bálint File, Emília Tóth, Virág Bokodi, Zsolt Keczer, László Entz, István Ulbert, Dániel Fabó (2017): “Functional connectivity analysis of cortico-cortical evoked potentials”, The 16th European Congress of Clinical Neurophisiology, Budapest, Hungary

B. File, D. File, F. Bodnár, I. Sulykos, K. Kecskés-Kovács, Z. Kardos, B. Tóth, R Boha, M. Márk, Á. Tóth, D. Fabó, I. Ulbert, I. Czigler (2016): Resting state functional subnetworks predict visual event related potentials, Fifth Biennal Conference on Resting State and Brain Connectivity, Vienna, Austria

Bálint File, Emília Tóth, László Entz, Tibor Nánási, Loránd Erőss, István Ulbert, Dániel Fabó (2016): Discovering the functional network behind cortico-cortical evoked potentials, IBRO workshop, poster presentation, Budapest, Hungary

Bálint File, Tibor Nánási, Brigitta Tóth, Márk Molnár, Cornelis J. Stam, Hillebrand Arjan, István Ulbert, Laszlo Entz, Lorand Erőss, Daniel Fabó: Functional connectivity analysis of the human epileptic intracranial EEG networks, 2015, 45th annual meeting of the Society for Neuroscience, Chicago, Illinois, USA

Bálint File, Brigitta Tóth, Roland Boha, C.J. Stam, Arjan Hillebrand, Márk Molnár (2015) Age-related alterations in the connectivity of resting state functional modular structure. 9th World Congess of International Brain Research Organization, Rio de Janeiro, Brazil. jul. 7-11.

Bálint File, Tibor Nánási, Brigitta Tóth, Márk Molnár, Cornelis J. Stam, Hillebrand Arjan, István Ulbert, László Entz, Loránd Erőss, Daniel Fabó (2015). Functional network analysis of human intracranial EEG recordings XV. Conference of the Hungarian Neuroscience Society, 2015, Budapest

Bálint File, Brigitta Tóth, Roland Boha, Zsófia Kardos, Márk Molnár (2013). Source level based investigation of age-related changes in resting state functional connectivity, The 53rd AnnualMeeting of the Society for Psychophysiological Research (Florence, Italy, October). Psychophysiology, Vol. 50, Supplement, p. S25.

Bálint File, Brigitta Tóth, Roland Boha, Zsófia Kardos, Márk Molnár (2013): Age-related changes in resting state functional network characteristics, V. Dubrovnik Conference on Cognitive Science (DuCog), Sleep, neural oscillations and cognition, 2013, Dubrovnik. Learning & Perception, Volume 5, Supplement 1., p. 40.

File B., Tóth B., Roland Boha, Molnár. M. (2013) Age-related alterations in resting state functional modular structure of neural networks. XIV. Conference of the Hungarian Neuroscience Society, 2013, Budapest

File B., Tóth B., Boha R., Hidasi Z., Gaál Zs.A., Stam C.J., Molnár M. (2012). Longitudinal study of resting state functional network characteristics in mild cognitive impairment, 52nd Annual Meeting of Society of Psychophysiology Research, September 19-23, 2012, New Orleans, Louisiana, USA, Psychophysiology, Vol. 49, Supplement , p. S30

# 7 ACKNOWLEDGEMENTS

First, I would like to say thank to Professor György Karmos for his wonderful lectures about electrophysiology and his recommendation, which made me possible to join to the research group of Professor Márk Molnár. Without his references I may have not started studying and applying signal processing and graph theory during my further works.

I'd like to express my gratitude to Márk Molnár (PhD, Dsc) and Brigitta Tóth (PhD) for supervising me as an undergraduate student. They taught me the stable foundation of EEG processing and graph theory. They also encouraged me to keep searching for my methodological problems individually.

I'm really grateful for Pofessor C.J. Stam and Arjan Hillebrand (PhD) from the VU Universtiy, Amsterdam for receiving me during my Erasmus internship. I have learned several usefull techniques under their supervision and also met with an ideal structure of lab meetings.

As my PhD supervisor, I'm especially grateful for Professor István Ulbert. He gave me the opportunity to join to different projects, work with numerous other researchers from various fields. However, with his brilliant intelligence he was always ready to solve my research problems, when I was in a need. This kind of freedom was necessary for discovering my scientific interests and learn the basics of successful teamwork.

I am much obliged to Zoltán Somogyvári (PhD). During my entire PhD studies he gave me valuable methodological advices both in the field of signal processing and network science.

During the association network analysis, I had the privilege to collaborate Zsolt Keczer. This fruitful cooperation was necessary for the existence of present dissertation. Also I'm grateful for Gábor Orosz (PhD) for the supervision of the social psychological parts of the study. These outstanding social psychologists encouraged and largely helped my work to create and validate the network analysis of associations.

Although, some of my research works are not involved in this thesis, the contribution of these projects had a significant impact of my scientific progress. Hence, I wish to sincerely thank the numerous advices from the coworkers of these projects.

I would like to say thank you to the European Social Fund for the sponsorship (EFOP-3.6.3-VEKOP-16-2017-00002) during the preparation of the present work.

Last but not least, I would like to thank to my family. Without their unconditional love, encouragement and support I would not have succeeded.

# 8 BIBLIOGRAPHY

[1] H. W. Kuhn, "The Hungarian method for the assignment problem," *Nav. Res. Logist. Q.*, vol. 2, no. 1-2, pp. 83–97, 1955.

[2] Y. Hu, "Efficient, high-quality force-directed graph drawing," *Math. J.*, vol. 10, no. 1, pp. 37–71, 2005.

[3] M. Bastian, S. Heymann, M. Jacomy, and others, "Gephi: an open source software for exploring and manipulating networks.," *ICWSM*, vol. 8, pp. 361–362, 2009.

[4] C. J. Stam and E. C. W. van Straaten, "The organization of physiological brain networks.," *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1067–87, Jun. 2012.

[5] H.-J. Park and K. J. Friston, "Structural and functional brain networks: from connections to cognition.," *Science*, vol. 342, no. November, p. 1238411, 2013.

[6] O. Sporns and R. F. Betzel, "Modular Brain Networks," *Annu. Rev. Psychol.*, vol. 67, no. 1, pp. 613–640, 2016.

[7] D. Meunier, S. Achard, A. Morcom, and E. Bullmore, "Age-related changes in modular organization of human brain functional networks," *Neuroimage*, vol. 44, no. 3, pp. 715–723, 2009.

[8] W. de Haan, W. M. van der Flier, T. Koene, L. L. Smits, P. Scheltens, and C. J. Stam, "Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer's disease.," *Neuroimage*, vol. 59, no. 4, pp. 3085–93, Feb. 2012.

[9] A. F. Alexander-Bloch *et al.*, "Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia.," *Front. Syst. Neurosci.*, vol. 4, no. October, p. 147, 2010.

[10] M. Newman and M. Girvan, "Finding and evaluating community structure in

networks," *Phys. Rev. E*, pp. 1–16, 2004.

[11] R. Guimerà and L. a Nunes Amaral, "Functional cartography of complex metabolic networks.," *Nature*, vol. 433, no. 7028, pp. 895–900, Mar. 2005.

[12] A. Alexander-Bloch, R. Lambiotte, B. Roberts, J. Giedd, N. Gogtay, and E. Bullmore, "The discovery of population differences in network community structure: new methods and applications to brain functional networks in schizophrenia," *Neuroimage*, vol. 59, no. 4, pp. 3889–3900, 2012.

[13] M. N. Moussa, M. R. Steen, P. J. Laurienti, and S. Hayasaka, "Consistency of network modules in resting-state FMRI connectome data.," *PLoS One*, vol. 7, no. 8, p. e44428, Jan. 2012.

[14] L. Geerligs, R. J. Renken, E. Saliasi, N. M. Maurits, and M. M. Lorist, "A brain-wide study of age-related changes in functional connectivity," *Cereb. Cortex*, vol. 25, no. 7, pp. 1987–1999, 2015.

[15] D. Meunier, S. Achard, A. Morcom, and E. Bullmore, "Age-related changes in modular organization of human brain functional networks.," *Neuroimage*, vol. 44, no. 3, pp. 715–23, Feb. 2009.

[16] M. Y. Chan, D. C. Park, N. K. Savalia, S. E. Petersen, and G. S. Wig, "Decreased segregation of brain systems across the healthy adult lifespan," *Proc. Natl. Acad. Sci.*, vol. 111, no. 46, pp. E4997–E5006, 2014.

[17] N. Tzourio-Mazoyer *et al.*, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage*, vol. 15, no. 1, pp. 273–289, 2002.

[18] M. L. Schölvinck, A. Maier, Q. Y. Frank, J. H. Duyn, and D. A. Leopold, "Neural basis of global resting-state fMRI activity," *Proc. Natl. Acad. Sci.*, vol. 107, no. 22, pp. 10238–10243, 2010.

[19] M. Rubinov and O. Sporns, "Weight-conserving characterization of complex functional brain networks," *Neuroimage*, vol. 56, no. 4, pp. 2068–2079, 2011.

[20] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech. Theory Exp.*, vol. 2008, no. 10, p. P10008, 2008.

[21] M. Meilă, "Comparing clusterings—an information based distance," *J. Multivar. Anal.*, vol. 98, no. 5, pp. 873–895, 2007.

[22] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[23] R. Guimera and L. Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, no. February, pp. 895–900, 2005.

[24] D. Tomasi and N. D. Volkow, "Aging and functional brain networks," *Mol. Psychiatry*, vol. 17, no. 5, pp. 549–558, 2012.

[25] C. Grady, S. Sarraf, C. Saverino, and K. Campbell, "Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks," *Neurobiol. Aging*, vol. 41, pp. 159–172, 2016.

[26] X. J. Chai, A. N. Castañón, D. Öngür, and S. Whitfield-Gabrieli, "Anticorrelations in resting state networks without global signal regression," *Neuroimage*, vol. 59, no. 2, pp. 1420–1428, 2012.

[27] J. Song *et al.*, “Age-related reorganizational changes in modularity and functional connectivity of human brain networks,” *Brain Connect.*, vol. 4, no. 9, pp. 662–676, 2014.

[28] K. Onoda and S. Yamaguchi, “Small-worldness and modularity of the resting-state functional brain network decrease with aging,” *Neurosci. Lett.*, vol. 556, pp. 104–108, 2013.

[29] B. C. M. Van Wijk, C. J. Stam, and A. Daffertshofer, “Comparing brain networks of different size and connectivity density using graph theory,” *PLoS One*, vol. 5, no. 10, p. e13701, 2010.

[30] B. H. Good, Y. A. De Montjoye, and A. Clauset, “Performance of modularity maximization in practical contexts,” *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 81, no. 4, pp. 1–20, 2010.

[31] E. L. Kinsella, T. D. Ritchie, and E. R. Igou, “Zeroing in on heroes: A prototype analysis of hero features.,” *J. Pers. Soc. Psychol.*, vol. 108, no. 1, p. 114, 2015.

[32] G. R. Goethals and S. T. Allison, “Making heroes: The construction of courage, competence, and virtue,” in *Advances in experimental social psychology*, vol. 46, Elsevier, 2012, pp. 183–235.

[33] L. E. Rankin and A. H. Eagly, “Is his heroism hailed and hers hidden? Women, men, and the social construction of heroism,” *Psychol. Women Q.*, vol. 32, no. 4, pp. 414–422, 2008.

[34] M. P. Sullivan and A. Venter, “Defining heroes through deductive and inductive investigations,” *J. Soc. Psychol.*, vol. 150, no. 5, pp. 471–484, 2010.

[35] Z. E. Franco, K. Blau, and P. G. Zimbardo, “Heroism: A conceptual analysis and differentiation between heroic action and altruism,” *Rev. Gen. Psychol.*, vol. 15, no. 2, pp. 99–113, 2011.

[36] M. P. Sullivan and A. Venter, “The hero within: Inclusion of heroes into the self,” *Self Identity*, vol. 4, no. 2, pp. 101–111, 2005.

[37] F. Farley, “The Real Heroes of" The Dark Knight"[Internet],” *Psychol. Today*, 2012.

[38] S. Moscovici, *La psychanalyse, son image et son public*. Presses universitaires de France, 2015.

[39] S. Moscovici, “The phenomenon of social representations.,” *Soc. Represent.*, pp. 3–69, 1984.

[40] E. Kirchler, B. Maciejovsky, and F. Schneider, “Everyday representations of tax avoidance, tax evasion, and tax flight: Do legal differences matter?,” *J. Econ. Psychol.*, vol. 24, no. 4, pp. 535–553, 2003.

[41] J.-C. Abric, “Méthodologie de recueil des représentations socials,” *Paris PUF*, 1994.

[42] J.-C. Abric, *Méthodes d’étude des représentations sociales*. Erès, 2003.

[43] J.-C. Abric, “L’approche structurale des représentations sociales: développements récents,” *Psychol. société*, vol. 4, no. 2, pp. 81–104, 2001.

[44] P. Érdi *et al.*, “Prediction of emerging technologies based on analysis of the US patent citation network,” *Scientometrics*, vol. 95, no. 1, pp. 225–242, 2013.

[45] C. Böde, I. A. Kovács, M. S. Szalay, R. Palotai, T. Korcsmáros, and P. Csermely, “Network analysis of protein dynamics,” *Febs Lett.*, vol. 581, no. 15, pp. 2776–2782,

2007.

[46] C. J. van Stam and E. C. W. Van Straaten, “The organization of physiological brain networks,” *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1067–1087, 2012.

[47] R. S. Burt, M. Kilduff, and S. Tasselli, “Social network analysis: Foundations and frontiers on advantage,” *Annu. Rev. Psychol.*, vol. 64, pp. 527–547, 2013.

[48] M. Steyvers and J. B. Tenenbaum, “The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth,” *Cogn. Sci.*, vol. 29, no. 1, pp. 41–78, 2005.

[49] G. G. Knyazev, N. V. Volf, and L. V. Belousova, “Age-related differences in EEG connectivity and network topology,” *Neurobiol. Aging*, pp. 1–11, 2015.

[50] L. Li, D. Alderson, J. C. Doyle, and W. Willinger, “Towards a theory of scale-free graphs: Definition, properties, and implications,” *Internet Math.*, vol. 2, no. 4, pp. 431–523, 2005.

[51] M. C. González, H. J. Herrmann, J. Kertész, and T. Vicsek, “Community structure and ethnic preferences in school friendship networks,” *Phys. A Stat. Mech. its Appl.*, vol. 379, no. 1, pp. 307–316, 2007.

[52] P. F. Jonsson and P. A. Bates, “Global topological features of cancer proteins in the human interactome,” *Bioinformatics*, vol. 22, no. 18, pp. 2291–2297, 2006.

[53] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, “Uncovering the overlapping community structure of complex networks in nature and society,” *Nature*, vol. 435, no. 7043, p. 814, 2005.

[54] P. Vergès and C. Guimelli, “Approche du noyau central: propriétés quantitatives et structurales,” *Struct. Transform. des représentations Soc.*, vol. 277, pp. 233–253, 1994.

[55] M. L. Rouquette and C. Flament, “Anatomie des idées ordinaires,” *Comment {é}tudier les Represent. Soc. Paris Arman. Colin*, 2003.

[56] J.-C. Abric, “Les représentations sociales: aspects théoriques,” *Prat. Soc. représentations*, vol. 2, pp. 11–37, 1994.

[57] G. Karypis, “Evaluation of item-based top-n recommendation algorithms,” in *Proceedings of the tenth international conference on Information and knowledge management*, 2001, pp. 247–254.

[58] D. Billsus and M. J. Pazzani, “Learning Collaborative Information Filters.,” in *Icml*, 1998, vol. 98, pp. 46–54.

[59] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science (80-. ).*, vol. 286, no. 5439, pp. 509–512, 1999.

[60] A. Clauset, C. R. Shalizi, and M. E. J. Newman, “Power-law distributions in empirical data,” *SIAM Rev.*, vol. 51, no. 4, pp. 661–703, 2009.

[61] M. E. J. Newman, “Fast algorithm for detecting community structure in networks,” *Phys. Rev. E*, vol. 69, no. 6, p. 66133, 2004.

[62] Y. Sun, B. Danila, K. Josić, and K. E. Bassler, “Improved community structure detection using a modified fine-tuning strategy,” *EPL (Europhysics Lett.*, vol. 86, no. 2, p. 28004, 2009.

[63] R. Nagarajan, A. T. Kalinka, and W. R. Hogan, “Evidence of community structure in biomedical research grant collaborations,” *J. Biomed. Inform.*, vol. 46, no. 1, pp. 40–

46, 2013.

[64] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian, “ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software,” *PLoS One*, vol. 9, no. 6, p. e98679, 2014.

[65] J. C. Abric, “Central system, peripheral system: their functions and roles in the dynamics of social representations,” *Pap. Soc. Represent.*, vol. 2, no. 2, pp. 75–78, 1993.

[66] J. Wachelke, “Representations and social knowledge: an integrative effort through a normative structural perspective,” *New Ideas Psychol.*, vol. 30, no. 2, pp. 259–269, 2012.

[67] S. Moscovici, “The Phenomenon of Social Representations. W: RM Farr, S. Moscovici (red.), Social Representations (s. 3--69).” Cambridge: Cambridge University Press, 1984.

[68] W. Wagner *et al.*, “Theory and method of social representations,” *Asian J. Soc. Psychol.*, vol. 2, no. 1, pp. 95–125, 1999.

[69] K. Bansak, J. Hainmueller, and D. Hangartner, “How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers,” *Science (80-. ).*, vol. 354, no. 6309, pp. 217–222, Oct. 2016.

[70] D. Lazer, R. Kennedy, G. King, and A. Vespignani, “The parable of Google Flu: traps in big data analysis,” *Science (80-. ).*, vol. 343, no. 6176, pp. 1203–1205, 2014.

[71] H. Joffe and J. W. B. Elsey, “Free association in psychology and the grid elaboration method.,” *Rev. Gen. Psychol.*, vol. 18, no. 3, p. 173, 2014.

[72] B. P. Bradley, K. Mogg, and R. Williams, “Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety,” *Behav. Res. Ther.*, vol. 33, no. 7, pp. 755–770, 1995.

[73] P. M. Niedenthal, J. B. Halberstadt, and Å. H. Innes-Ker, “Emotional response categorization.,” *Psychol. Rev.*, vol. 106, no. 2, p. 337, 1999.

[74] J. B. Halberstadt, P. M. Niedenthal, and J. Kushner, “Resolution of lexical ambiguity by emotional state,” *Psychol. Sci.*, vol. 6, no. 5, pp. 278–282, 1995.

[75] L. B. Szalay and J. E. Brent, “The analysis of cultural meanings through free verbal associations,” *J. Soc. Psychol.*, vol. 72, no. 2, pp. 161–187, 1967.

[76] W. Wagner, J. Valencia, and F. Elejabarrieta, “Relevance, discourse and the 'hot' stable core social representations: A structural analysis of word associations,” *Br. J. Soc. Psychol.*, vol. 35, no. 3, pp. 331–351, 1996.

[77] R. L. Nabi, “A cognitive-functional model for the effects of discrete negative emotions on information processing, attitude change, and recall,” *Commun. theory*, vol. 9, no. 3, pp. 292–320, 1999.

[78] L. B. Szalay, R. M. Kelly, and W. T. Moon, “Ideology: Its meaning and measurement,” *Comp. Polit. Stud.*, vol. 5, no. 2, pp. 151–173, 1972.

[79] L. B. Szalay, C. Windle, and D. A. Lysne, “Attitude measurement by free verbal associations,” *J. Soc. Psychol.*, vol. 82, no. 1, pp. 43–55, 1970.

[80] S. M. Holmes and H. Castañeda, “Representing the ‘European refugee crisis’ in Germany and beyond: Deservingness and difference, life and death,” *Am. Ethnol.*, vol. 43, no. 1, pp. 12–24, 2016.

[81] A. Kallius, D. Monterescu, and P. K. Rajaram, “Immobilizing mobility: Border ethnography, illiberal democracy, and the politics of the ‘refugee crisis’ in Hungary,” *Am. Ethnol.*, vol. 43, no. 1, pp. 25–37, 2016.

[82] N. Kteily, E. Bruneau, A. Waytz, and S. Cotterill, “The ascent of man: Theoretical and empirical evidence for blatant dehumanization.,” *J. Pers. Soc. Psychol.*, vol. 109, no. 5, p. 901, 2015.

[83] W. G. Stephan, C. W. Stephan, and S. Oskamp, “An integrated threat theory of prejudice,” *Reducing Prejud. Discrim.*, pp. 23–45, 2000.

[84] R. Schweitzer, S. Perkoulidis, S. Krome, C. Ludlow, and M. Ryan, “Attitudes towards refugees: The dark side of prejudice in Australia,” *Aust. J. Psychol.*, vol. 57, no. 3, pp. 170–179, 2005.

[85] E. Halperin, A. G. Russell, K. H. Trzesniewski, J. J. Gross, and C. S. Dweck, “Promoting the Middle East peace process by changing beliefs about group malleability,” *Science (80-. ).*, vol. 333, no. 6050, pp. 1767–1769, 2011.

[86] F. Pratto, J. Sidanius, L. M. Stallworth, and B. F. Malle, “Social dominance orientation: A personality variable predicting social and political attitudes.,” *J. Pers. Soc. Psychol.*, vol. 67, no. 4, p. 741, 1994.

[87] S. De Deyne and G. Storms, “Word associations: Norms for 1,424 Dutch words in a continuous task,” *Behav. Res. Methods*, vol. 40, no. 1, pp. 198–205, 2008.

[88] G. Orosz and C. Roland-Lévy, “Social representation of competition and fraud,” *Citizsh. Teach. Learn.*, vol. 8, no. 2, pp. 157–177, 2013.

[89] P. Ekman, “Are there basic emotions?,” 1992.

[90] A. Ortony and T. J. Turner, “What’s basic about basic emotions?,” *Psychol. Rev.*, vol. 97, no. 3, p. 315, 1990.

[91] D. L. Robinson, “Brain function, emotional experience and personality,” *Neth. J. Psychol.*, vol. 64, no. 4, pp. 152–168, 2008.

[92] C. E. Izard, *Human emotions*. Springer Science & Business Media, 2013.

[93] D. E. Beaton, C. Bombardier, F. Guillemin, and M. B. Ferraz, “Guidelines for the process of cross-cultural adaptation of self-report measures,” *Spine (Phila. Pa. 1976).*, vol. 25, no. 24, pp. 3186–3191, 2000.

[94] T. Dunning, “Accurate Methods for the Statistics of Surprise and Coincidence,” *Comput. Linguist.*, vol. 19, no. 1, pp. 61–74, 1993.

[95] M. Sarrica, “War and peace as social representations: Cues of structural stability.,” *Peace Confl. J. Peace Psychol.*, vol. 13, no. 3, p. 251, 2007.

[96] S. Gómez, P. Jensen, and A. Arenas, “Analysis of community structure in networks of correlated data,” *Phys. Rev. E*, vol. 80, no. 1, p. 16114, 2009.

[97] A. Lancichinetti and S. Fortunato, “Consensus clustering in complex networks,” *Sci. Rep.*, vol. 2, 2012.

[98] W. Simpson, “QAP: The quadratic assignment procedure,” in *North American Stata Users’ Group Meeting, March*, 2001, pp. 12–13.

[99] N. C. Wormald, “Models of random regular graphs,” *London Math. Soc. Lect. Note Ser.*, pp. 239–298, 1999.

[100] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: uses and interpretations,” *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[101] D. L. Nelson, C. L. McEvoy, and T. A. Schreiber, “The University of South Florida free association, rhyme, and word fragment norms,” *Behav. Res. Methods, Instruments, Comput.*, vol. 36, no. 3, pp. 402–407, 2004.

[102] S. De Deyne and G. Storms, “Word associations: Network and semantic properties,” *Behav. Res. Methods*, vol. 40, no. 1, pp. 213–231, 2008.

[103] R. L. Nabi, “Exploring the framing effects of emotion: do discrete emotions differentially influence information accessibility, information seeking, and policy preference?,” *Communic. Res.*, vol. 30, no. 2, pp. 224–247, 2003.

[104] J. F. Dovidio, K. Kawakami, and S. L. Gaertner, “Implicit and explicit prejudice and interracial interaction.,” *J. Pers. Soc. Psychol.*, vol. 82, no. 1, p. 62, 2002.

[105] N. Haslam and S. Loughnan, “Dehumanization and infrahumanization,” *Annu. Rev. Psychol.*, vol. 65, pp. 399–423, 2014.

[106] L. D. Appelbaum, “Who deserves help? Students’ opinions about the deservingness of different groups living in Germany to receive aid,” *Soc. Justice Res.*, vol. 15, no. 3, pp. 201–225, 2002.

[107] M. Verkuyten, “Emotional reactions to and support for immigrant policies: Attributed responsibilities to categories of asylum seekers,” *Soc. Justice Res.*, vol. 17, no. 3, pp. 293–314, 2004.

[108] A. M. Nickerson and W. R. Louis, “Nationality Versus Humanity? Personality, Identity, and Norms in Relation to Attitudes Toward Asylum Seekers1,” *J. Appl. Soc. Psychol.*, vol. 38, no. 3, pp. 796–817, 2008.

[109] J. Citrin, D. P. Green, C. Muste, and C. Wong, “Public opinion toward immigration reform: The role of economic motivations,” *J. Polit.*, vol. 59, no. 3, pp. 858–881, 1997.

[110] P. M. Sniderman, L. Hagendoorn, and M. Prior, “Predisposing factors and situational triggers: Exclusionary reactions to immigrant minorities,” *Am. Polit. Sci. Rev.*, vol. 98, no. 01, pp. 35–49, 2004.

[111] B. News, “Nice attack: At least 84 killed by lorry at Bastille Day celebrations,” Jul-2016.

[112] G. K. Zipf, “The psycho-biology of language.,” 1935.

[113] S. Bordag, “A comparison of co-occurrence and similarity measures as simulations of context,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4919 LNCS, pp. 52–63, 2008.

[114] A. Lancichinetti, F. Radicchi, J. J. Ramasco, and S. Fortunato, “Finding statistically significant communities in networks,” *PLoS One*, vol. 6, no. 4, p. e18961, 2011.

[115] M. F. Porter, “An algorithm for suffix stripping,” *Program*, vol. 14, no. 3, pp. 130–137, 1980.

[116] A. Esuli and F. Sebastiani, “Sentiwordnet: A publicly available lexical resource for opinion mining.,” in *LREC*, 2006, vol. 6, pp. 417–422.

[117] C. G. Lord, L. Ross, and M. R. Lepper, “Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence.,” *J. Pers. Soc. Psychol.*, vol. 37, no. 11, p. 2098, 1979.

[118] J. Hainmueller and D. J. Hopkins, “Public attitudes toward immigration,” *Annu. Rev. Polit. Sci.*, vol. 17, pp. 225–249, 2014.