Simply Clustering. Making New Sense In The Five Facets Mindfulness Questionnaire

Elena Druică (elena.druica@faa.unibuc.ro)
University of Bucharest, Romania

Rodica Ianole - Calin (rodica.ianole@faa.unibuc.ro)
University of Bucharest, Romania

ABSTRACT

A common approach in examining data collected based on different scales is to look at their structure by means of factor analysis. This article provides a way to look not only into the overall mindfulness score as an individual characteristic, but also at how the mindfulness dimensions cluster together to provide potentially consistent individual profiles. The novelty of our contribution is two fold: we reached our goal with the help of cluster analysis, and not by the means of previous methodological approaches. Also, we applied the most popular tools used to measure mindfulness, the face facets mindfulness questionnaire, on a sample of Romanian participants which makes this research the first study on mindfulness conducted on a Romanian sample. We found that, despite the existence of some stable groups that share similar characteristics, the degree of homogeneity across individuals is pretty high. In addition, the levels of mindfulness corresponding to our participants seems to be unrelated with background variables like age, gender, and working place.

Keywords*: Mixed methods, Health care, Survey research, Regression analysis*

JEL Classification*: C19, C38*

1. INTRODUCTION

The realm of methodological tools of investigation for social sciences has largely extended in the last decades, both in terms of research designs and coverage. Beyond its extraordinary potential, the complexity and diversity of available data exposes to new challenges when choosing the most appropriate methods of analysis. With respect to scale measurement and development, the confirmatory factor analysis framework has become extremely popular in serving the objective of testing the construct validity of item sets (Gerbing & Anderson, 1988; Reise, Waller and Comrey, 2000). In the area of application that motivates our study, mindfulness research, especially in regard to national measurement, reliability and factor structure seem to be to most targeted objectives. On the one hand, this is a positive fact due to the need of applying objective filters for the numerous questionnaires aiming to capture the essence of mindfulness (Baer, 2011). On the other hand, the risk of falling under the principles of a very standardized path can materialize in a certain level of ignorance for other properties of the data.

The current paper was originally conceived in order to build up a strong argument for putting mindfulness on the national public research agenda in Romania. This would be a natural development since the practice of mindfulness has been gradually accepted in the Western psychology (Shapiro, 2009) and medicine (Sauer et al., 2011). There is generous stream of studies assessing the clinical benefits of mindfulness (Keng, Smoski and Robins, 2011; Davis and Hayes, 2011; Baer, 2003) or in attention enhancement, present-moment awareness, and its stress reduction effects (Purser & Milillo 2015), thus creating strong premises for considering it an important tool in the contemporary mental healthcare portfolio. Moreover, the positive impact of mindfulness can be noticed also in organizational settings, in relation with decision-making (Fiol and O'Connor, 2003) and wise actions (Weick and Putnam 2006), leadership development (Goldman-Schuyler et al., 2017) and subjective wellbeing (Brown et al., 2009).

Objectives like measuring and capitalizing on the benefits of an increased mindfulness level for different populations would be much easier promoted if a national study would assess the cross-cultural validity of one of the existing questionnaires. Since our pilot study is just a preliminary step for conducting the testing on a representative sample, we took the opportunity to resort to a more unconventional statistical approach: an exploratory cluster analysis. The operational advantage is that we avoid the limitations of the minimum sample size recommended for factor analysis (Mundfrom, Shaw and Ke, 2005). Free from that constraint, the study makes two unique contributions. First, we examine the structure of the mindfulness score captured with the Five Facets Mindfulness Questionnaire (FFMQ), applied for a first time to a Romanian sample. By applying the hierarchical clustering algorithm and two linkage methods, Average and Ward, we determined a relatively high level of stability, thus homogeneity between the four clusters obtained, hinting to the fact that subjects posses similar characteristics corresponding to the five dimensions of mindfulness. Second, we consider the detected clusters and three background variables - gender, age and workplace – as potential predictors for the FFMQ score. The statistical significance of the groups and the lack of significance of all the other covariates further strengthen the existing differences on mindfulness scores between groups, and the rather irrelevance of the background variables in explaining the level of mindfulness in our sample.

The rest of the article is structured as follows: section two consists of a brief review of the most recent approaches in examining the structure of FFMQ questionnaire and the emphasis put on factor analysis. In the empirical part we provide justification for the number of clusters we choose, we decide on the clustering algorithm and the appropriated linkage methods, and provide detailed explanations on clusters stability. Finally, we look into some potential determinants of the FFMQ score. We conclude with both methodological and content implications for the study of mindfulness, along with a careful revision of the limitations of our study.

2. MEASURES AND COMPUTATIONAL APPROACHES FOR MINDFULNESS

2.1. The Five Facets of Mindfulness Questionnaire (FFMQ)

The set of investigative tools for mindfulness ranges from one-dimensional examples, like the Mindfulness and Attention Awareness Scale (MAAS) or the Freiburg Mindfulness Inventory (FMI), to most popular multi-dimensional ones: with two factors - Toronto Mindfulness Scale (TMS), with four factors - the Kentucky Inventory of Mindfulness (KIMS) or with five factors - the Five Facets Mindfulness Questionnaire (FFMQ). As it is natural, parsimony proves more efficient in some well-defined cases (patients dealing with depression, anxiety etc), or for some groups (e.g. regular versus non-regular meditators), but at the level of the general population there is still much to discuss regarding the best and accurate coverage of mindfulness. The findings of reputed scholars in the field like Ruth Baer and her colleagues (2006) commend the higher level of sharpness captured by a multifaceted construct of mindfulness, thus offering an important choice criteria.

Bergomi, Tschacher and Kupper (2013) offer a good overview on advantages, disadvantages and particularities in implementation of these different scales. In line with their opinion that the Five Facet Mindfulness Questionnaire – FFMQ (Baer et al., 2008) is the best choice for conducting evaluations at the level of the general population, the current paper takes the tool as a reference point for exploring mindfulness on a Romanian sample.

Moreover, taking into account the objective of national comparability, the widespread use of FFMQ, strengthens the choice of the scale. Several studies have examined the psychometric properties of the scale and grassroots implications for different countries: Spain (Cebolla et al., 2012), Sweden

(Lilia et al, 2011), Italy (Giovannini et al., 2014), France (Heeren et al., 2011), Netherlands (Bohlmeijer et al., 2011), Hong Kong (Hou et al., 2014), Japan (Sugiura et al., 2012), Brazil (Barros et al., 2014) etc. However, we found little or almost no evidence for calibrating this particular tool in Romania and Eastern Europe in general. Thus, the paper wants to address and minimize this gap.

The 39-item scale of FFMQ was translated into Romanian and implemented through an on-line questionnaire, sent to a small sample of participants located in Bucharest, both students and working adults.

2.2. The analytics around FFMQ

While confirmatory factor analysis probably holds the highest position for favorite approach in dealing with the FFMQ scale, there are many technical nuances to account for. The original study conducted by Ruth Baer and her colleagues (2006) was focused on the procedure of item parceling, with other papers following not only in a replication mindset (Williams et al., 2014; Deng et al., 2011; Aguado, 2015) but also in implementing slight changes. The immediate alternative was to employ individual items as indicators (Neuser, 2010), under the assumption that this manner shows a more salient proof for model fit and model specification. Tran et al. (2013) makes a strong argument about the problems associated to item parcels procedures but nonetheless there is no consensual view on the matter. From a general theoretical perspective concerning factor analysis, more recent interventions, like Zhang and Preacher (2015), suggest increase attention for the cases in which we may obtain similar rotated factor loading but very different standard errors, depending on the method chosen for factor rotation.

Often, structural equation modeling (SEM) is also used to investigate the factorial structure of a questionnaire, with different preferences for the maximum likelihood minimization path (Bollen and Lennox, 1991) or for bootstrapping goodness of fit measures (Bollen and Stine, 1992). However, despite the consistent alternative, in the case of SEM analysis applied to mindfulness, the focus is more on explaining the links between different other variables, and not on the internal configuration of the concept: a structural model of procrastination, mindfulness and health relationships (Sirois and Tosti, 2012) with mindfulness as a mediator variable or a structural model of mindfulness, sleep and wellbeing (Howell, Digdon and Buro, 2010). One crucial aspect here is also the fact that those SEM studies usually employ one-dimensional scales for mindfulness (MAAS or FMI), and thus they somewhat avoid the analytical layers of the five dimensions proposed by FFMQ.

3. **DATA AND EMPIRICAL ANALYSIS**

3.1. Sample characteristics and several implications

The sample consists of 111 subjects, among which 47.2% are females and 52.8% are males. 43.6% of the participants in this study work in public institutions, while the rest of 56.4% work in private sector. Nearly half of the participants (46.4%) work in the IT domain, 25.5% have jobs that involve managerial decisions, while the rest of 28.1% belong to a separate category of "support jobs", like for example being a translator, or a specialist in communication. The descriptive statistics for the five numerical dimensions of the FFMQ, as well as for the total score, are presented in Table 1 below.

Descriptive statistics

Table 1

Dimension	Min	Mean	Median	Max	Standard Deviation
Observing	12	24.81	25	39	5.26
Describing	17	30.22	31	40	5.74
Awareness	16	30.84	32	40	5.53
Non-judging	12	27.03	27	39	5.97
Non-reactivity	11	21.11	22	32	4.13
FFMQ	95	134	135.5	175	15.26

Source: Authors' work

In each case, both the mean and the standard deviation are very similar to the results obtained for the Spanish (Cebolla et al., 2012) and Italian samples (Giovannini et al., 2014), though slightly higher compared to the Chinese study (Hou et al., 2014) in the awareness and non-judging dimension. It is noteworthy that these exact two elements are the most salient in respect to dealing with different forms of psychological distress (Hayes & Feldman, 2004).

However, there is a little bit more that can be derived from these simple characteristics of the sample. We choose to conduct a preliminary analysis of the five dimensions and final score, and investigate whether our variables are normally distributed. This fact will be essential, as we will explain later. To conduct the normality test we used the "moments" package in R, and the function jarque.test() for the Jarque – Berra Normality test (Jarque&Bera 1980), available in the "moments" package in R, test created by Frederick Novomestky (Komsta& Novomestky 2015). The null hypothesis is that the data comes from a normal distribution, while the alternative hypothesis states that there are significant differences between the distribution of the data and the normal distribution, in the sense that skewness and kurtosis are significantly different than their default values in case of normally distributed data, namely 0 and respectively 3. Table 2 presents the results and shows that only the awareness dimension is not –normally distributed.

The Jarque – Berra Normality Test

Table 2

Dimension	JB - statistic	Result
Observing	1.37 *(0.5041)*	Fail to reject the null
Describing	4.4405 *(0.1086)*	Fail to reject the null
Awareness	7.3017 *(0.026)*	Reject the null
Non-judging	1.0924 *(0.579)*	Fail to reject the null
Non-reactivity	0.469 (0.791)	Fail to reject the null
FFMQ	0.342 (0.843)	Fail to reject the null

(p – values in parentheses)

Source: Authors' work

One of the implications of this preliminary finding is that, according to the properties of a normal distribution, approximately 99.9% of a population for which our sample might be representative would score to each dimension except for awareness, values situated three standard deviations about the mean. Table 3 summarizes the lower and the upper values of these possible scores.

99.9% of the population would score

Table 3

Dimension	Lower limit	Mean	Upper limit
Observing	9.02	24.81	40.60
Describing	12.99	30.22	47.45
Non-judging	9.11	27.03	44.94
Non-reactivity	8.69	21.11	33.52
FFMQ	88.19	134	179.81

Source: Authors' work

The “awareness” dimension calls for other two separate tests that would clarify how the distribution of this variable differs from a normal distribution. In this sense we conduct two tests: the D’Agostino test for skewness (D’Agostino 1970 now available in the “moments” package in R, developed by Lukasz Komsta, 2015), and the Anscombe-Glynn test of kurtosis (Anscombe&Glynn 1983, also available in the “moments” package cited before).

The D’Agostino test is built upon the fact that the skewness of a normal distribution is 0. The real data rarely obey to this rule, in the sense that the skewness might be in fact different than zero, but not significantly different to support the idea that data are not normally distributed. This test is aimed to detect significant non – zero skewness, under the null that the data is normally distributed. In a similar manner, the Anscombe-Glynn test addresses another characteristic of a normal distribution: it is built upon the fact that a normal distribution has a kurtosis of 3, and captures the significant differences between the kurtosis of a certain data and this benchmark value. The null hypothesis of this test is that data is normally distributed, while the alternative states that it is not. Each of the two tests can be conducted as one – tailed, or two tailed tests.

The skewness and kurtosis tests for Awareness (two tailed tests)

Table 4

Dimension	D’Agostino skewness test	Anscombe-Glynn test for kurtosis
Awareness	-0.62654 *(0.008)*	2.849 *(0.95)*

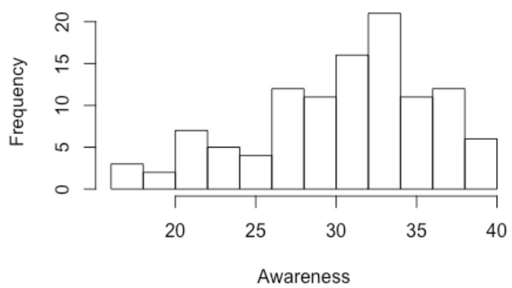
(p – values in parentheses)

Source: Authors' work

In our case, the second column in Table 4 shows that we reject the null hypothesis of a skewness equal to 0, but the kurtosis is near enough to 3 to consider that the distribution has a similar kurtosis as a normal distribution. Moreover, the one - tailed D’Agostino test reveals that the distribution of awareness is in fact negatively skewed, suggesting that some small extreme values occur in this variable for which Figure 1 presents the histogram.

Some small extreme values occur for the awareness dimension

Figure 1



Source: Authors' work

At the end of this descriptive subsection we can summarize that, of the five components of the mindfulness overall score, the only dimension that does not follow a normal distribution is awareness. The bias is induced by some extreme small values that would potentially suggest either some type of environmental or contextual noise at the moment of filling the questionnaire, or a low level of commitment when answering to this part of the survey for example.

3.2. An exploratory approach through cluster analysis: between the method and the empirical results

The FFMQ score has been developed and used before to rank subjects based on their level of mindfulness (Baer, 2011). The aim of this section is to explore whether within the structure of the FFMQ score there are significant patterns that group subjects into relevant categories. Let's take for example the overall score of 148: it can be achieved from numerous combinations between its five components, like observing = 25, describing = 36, awareness = 25, non – judging = 28, and non – reactivity = 34. Our question is: among all possibilities, are some of them more likely to provide this score of 148, than others? And if the answer is yes, do these combinations create stable patterns that may predict the overall score in a consistent way? Our intention in this section is therefore to explore whether we can gain in understanding if we look into the structure of the overall score, and not only to its total value.

We have chosen to further conduct this exploration through a cluster analysis applied to a data set comprising of the five components that lead to the final FFMQ score: observing, describing, awareness, non – judging, and non - reactivity. Cluster analysis is an unsupervised learning tool that aims at finding groups within data. The observations within groups are similar with each other, while the observations in different groups are not. This type of investigation is very common in statistical studies and has been used in various fields, though in what concerns the mindfulness studies another tool has been used widely, namely the factor analysis. To our knowledge, this is the first contribution in the area of mindfulness that uses cluster analysis to make a deeper sense of the five factors that have been derived through the preliminary investigation of the FFM questionnaire.

In the first stage we compute the Hopkins statistic, used as a measure for clustering tendency. This is a sensitive issue in the area of any type of unsupervised learning, where the analysis will tend to generate clusters even if in reality they are meaningless. To calculate the Hopkins statistic, we use the "clustertend" package in R, and the function hopkins(), and get in this case a value of 0.37. Having in mind that the critical value of the Hopking statistic, the one that indicates that the data is uniformly distributed and there are no meaningful clusters within is 0.5, we admit that there is a certain tendency to clustering in our case, though not that high (Hopkins&Skellam 1954, Lawson&Jurs 1990).

There are currently an impressive number of clustering algorithms, but choosing among them the best one to conduct a particular analysis is not an easy task (Flynt and Dean, 2016). Moreover, the problem of determining the right number of clusters, like for example k – mean or k – medoid requires, is also a difficult choice. The performance of the final clusters is evaluated based on several characteristics (compactness, well – separation, connectedness and stability), but also the relevance of the results in terms of practical interpretation is important. To identify the right clustering method and the right number of clusters, we used in the first instance the "clValid" package. The clValid() function in this package provides validation measures for several clustering algorithms and a specified number of clusters. Among the measures available with this function, internal and stability measures are of interest for our purpose. The internal measures are based on connectivity, Silhouette Width, and Dunn Index, while the stability measures include the average proportion of non-overlap (APN), the average distance (AD), the average distance between means (ADM), and the figure of merit (FOM). A detailed description of each measure and index is available in the clValid package description (Brock et al. 2008) as well as in Datta, S. and Datta (2003, 2006). Table 5 provides the recommendations based on each measure.

As it is very common in such cases, there is no clear best algorithm or number of clusters, each of the methods providing different recommendation.

To overcome this issue, we used the "RankAggreg" package that performs aggregation of ordered lists based on the ranks, via the Cross-Entropy Monte Carlo algorithm or the Genetic Algorithm (Pihur et al. 2007).

The recommended clustering method and number of clusters (clValid function)

Table 5

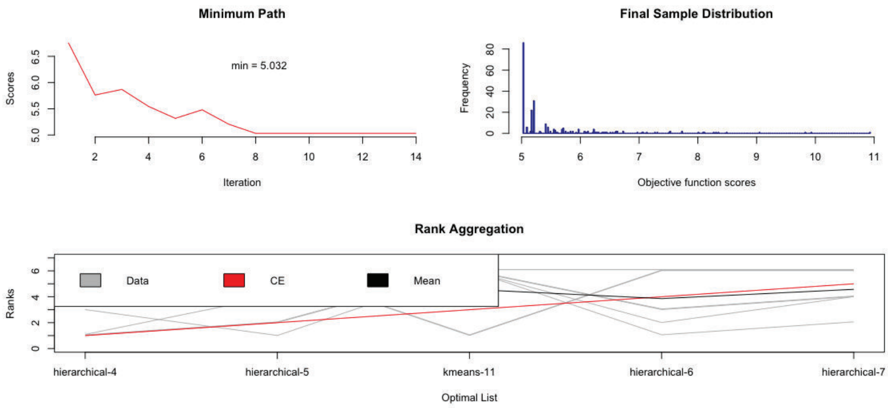
Index	Recommendation
APN	hierarchical - 4
AD	kmeans - 11
ADM	hierarchical - 5
FOM	kmeans - 11
Connectivity	hierarchical - 4
Dunn	hierarchical - 6
Silhouette	hierarchical - 4

Source: Authors' work

The RankAggreg function comes out with a recommendation for hierarchical clustering with 4 clusters. Figure 2 graphically presents the final output of this preliminary investigation.

The RankAggreg function recommends hierarchical clustering with 4 clusters

Figure 2



Source: Authors' work

The hierarchical agglomerative clustering is perhaps one of the most common clustering algorithms. Initially, each data point is assigned to its own singleton cluster, and then an iterative process begins. The similarity between clusters is evaluated based on a variety of pre-defined distances, the most similar clusters being merged at each step into one singe cluster. Hierarchical clustering provides a dendogram, a graphical presentation of the agglomerative process, which can be cut at various levels to produce the desired number of clusters. Figure 3 will illustrate this fact.

The problem of clustering algorithm being solved, the selection of the best linkage method comes next. This is a sensible issue, as many other issues in the area of unsupervised techniques. Usually, the clustering method is selected based on the logic of the result, of its power of interpretation, or on their appropriateness for the data to be clustered. The most used methods, among the wide variety available for hierarchical clustering, are: average linkage method, complete linkage method, and Ward's method. In our paper, we will compare the results provided by these three methods.

The Average Linkage method (Sokal&Michener 1958) is very intuitive and defines the distance between two clusters as the average distance between the data points in the first cluster, and the data points in the second cluster. As a result, at each stage of the analysis we will merge the clusters that have the smallest average linkage distance. The Average Linkage distance between two clusters G and H is defined as follows:

$$d_{AL} = \frac{1}{N_G N_H} \sum_{x \in G} \sum_{y \in H} d_{xy}$$

Here, N_G is the size of the cluster G, and N_H is the size of the cluster H. By d_{xy} we denoted the Euclidean distance between two generic points, x from G and respectively y from H. More exactly:

$d_{xy} = \sum_{i=1}^{n} \|x_i - y_i\|$, n being the total number of variables to be clustered.

In our particular case, this method will tend to merge groups of subjects that are, on average, very close to each other in terms of the scores achieved to each of the five components of the FFMQ score. Since taking the average in a data set has the potential to mitigate the differences between extreme values, we can expect that the clusters created based on this linkage method will be permissive and include a wider variety of individuals than other methods would include.

The Complete Linkage method creates groups for which the similarity is defined as the maximum distance between any single data point in the first cluster and any single data point in the second cluster. The Complete Linkage distance between two clusters G and H is defined as follows:

$$d_{CL} = max\{d_{xy}, x \in G, y \in H\}$$

The Euclidean distance between two generic points x and y was defined above. Therefore, the method pursues to combine clusters that have the smallest complete linkage distance between them and seek for final groups that differ on the basis of the maximum distance. Based on this definition, we expect that for our data the Complete Linkage method will create clusters where the extreme values will play a significant role. More precisely, we expect that this distance will avoid merging groups close to each other in terms of average distance, because of potential outliers that are far apart.

Unlike these approaches, the Ward method (Ward 1963) is distance – free and treats the clustering process as an analysis of variance. It looks for the clusters that explain the most of the variance in the data set, and aims at maximizing the value of R – square, defined as the ratio between the variance explained by the clusters over the total variance of the initial data set.

There is a high probability that the clusterization provided by these three methods will lead to different structure of the groups. As recommended in the literature, the logic behind choosing the appropriate linkage method should be the interpretation power as well as the stability of the clusters. While there are arguments stating that choosing a linkage measure that leads to interpretable clusters if a form of psychological validation, without too much scientific value, the stability of the clusters that also have a practical meaning can be a valuable selection tool.

The stability of the clusters created via these three methods is evaluated using the Jaccard coefficient. This coefficient, originally introduced by Sneath (1957), is defined as the ratio between the size of the intersection and the size of the union of two clusters G and H as follows:

$$J(G, H) = \frac{|G \cap H|}{|G \cup H|}$$

The clusterboot() function in the “fpc” R package provides a mean for assessing the stability of the clusters based on a bootstrap resampling scheme, and the value of the Jaccard coefficient of similarity is the main driver in reaching a conclusion. The idea behind this procedure is the following: we conduct an initial clustering and record the clusters. Then, the procedure is repeated, this time on a different data set resulted from resampling the observations in the initial data. A comparison between the original and the new clusters will provide information regarding the most similar cluster we found after resampling, the similarity being evaluated based on the Jaccard coefficient of similarity defined above. The algorithm compares the value with 0.5: whenever the distance between the old and new cluster is less than 0.5, the cluster is considered unstable and is dissolved. We count the number of time a cluster is dissolved and decide based on the rule described in the next paragraph whether the cluster includes relevant information, or only noise.

The critical values of this coefficient are 0.5, 0.6 and 0.75: highly unstable clusters correspond values of the Jaccard coefficient less than, or at most 0.5. If the coefficient ranges between 0.6 and 0.75, we admit that there are some patterns in the data, but it is unclear which data points belong to which cluster. Whenever the Jaccard coefficient is higher than 0.75, as admit that we found stable clusters. Values of 0.85 or higher point toward highly stable results. (Hennig 2007, 2008 and Zumel 2015). In our study we run a bootstrap with 1000 iterations, and the results are recorded in Table 4.

The Complete Linkage method provides unstable groups, each of the coefficients being below the trustworthy level of 0.6. Therefore, we will not consider the Complete Linkage method as an alternative to explore our data. Unlike this case, both the Ward method and the Average Linkage lead to the conclusion that there are some patterns in the data, but it is only the Average Linkage method that provides indeed a stable group, namely the first group with a Jaccard coefficient of 0.84. Moreover, the Average Linkage method results in three groups that may be considered as relevant patterns, while the Ward method identifies only two clusters with a Jaccard coefficient higher than 0.6.

The Jaccard coefficient for each linkage method, and cluster

Table 6

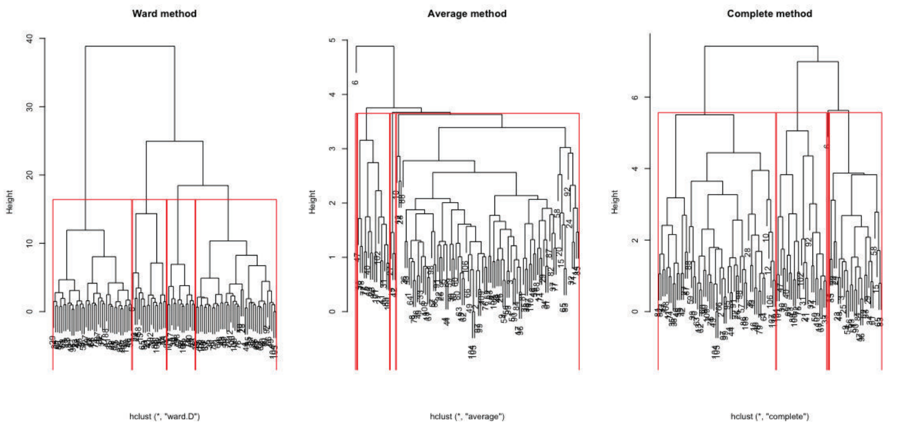
Method	Group 1	Group 2	Group 3	Group 4
Ward	0.6636	0.7220	0.5182	0.5021
Average	0.8432	0.6100	0.3620	0.6027
Complete	0.5339	0.4475	0.4688	0.3504

Source: Authors' work

Figure 3 graphically presents the dendograms that correspond to the three linkage methods, using hierarchical clustering and 4 clusters. The Ward method provides the most balanced groups in terms of size of the clusters, while the other two methods lead to groups with only few individuals. In what follows we will maintain the Average Linkage method and the Ward's method to further conduct the cluster analysis.

Hierarchical clustering with 4 clusters: Ward, Average and Complete Linkage methods

Figure 3



Source: Authors' work

Table 7 shows that the most consistent group created by the Average Linkage method – the one that proved to have the highest stability, according to the results presented in Table 6 – is the first group, comprising of 81.82% of the participants, namely 91 of the subjects. The least stable group, with a Jaccard coefficient of 0.362, is the third group, while the rest of the groups points toward some existing patterns, but not very stable. These remained groups include very few subjects, as follows: group 2 includes only one person; group 4 includes 16 persons, while the third group is left with 3.

Clusters characteristics – The Average Linkage Method

Table 7

Groups	Observing	Describing	Awareness	Non-judging	Non-reactivity	Proportion
1	24.63	31.14	32.42	27.89	21.13	0.8182
4	25.25	23.13	21.50	20.38	19.94	0.1455
3	31.33	38.33	38.00	34.67	30.00	0.0273
2	14.00	36.00	16.00	33.00	11.00	0.009

Source: Authors' work

By simply looking at the results recorded in Table 5, it is difficult to set description of the groups, based on the differences among them. The group that detaches itself from the rest is group 3, characterized by higher values than the rest for each dimension. Group two is also different in what concern the lowest values scored at observing, awareness and non-reactivity.

Compared with the remained groups, four and one, these subjects seem to be on average better at describing and have the rare quality of non – judging at nearly the similar level with the champion group, three. Group one and four seem to be more balanced in terms of scores to each dimension. While pretty similar at observing and non – reactivity (with slightly lower values in the fourth group), for the other three dimensions the first group scores up to 11 points more than group 4.

The first conclusion that derives from our result is that the sample is characterized by a high homogeneity, as long as up to 82% of the participants belong to the same groups, and this homogeneity results in pretty high scores at each dimension. We may interpret the results concerning this group as a type of "majority" profile. Although much small than group 1, group 4 includes an important percent of the subjects, and balanced scores in the sense that there are no extreme values as in group 2, for instance.

The Ward's method produces the groups recorded in Table 8. As mentioned earlier, the size of the groups are, in this case, more balanced that in the Average Linkage case. The first two groups, which are somehow consistent, share around 36% of the subjects each, while the rest of nearly 30% are split between the two unstable groups.

Clusters characteristics – The Ward's Method

Table 8

Groups	Observing	Describing	Awareness	Non-judging	Non-reactivity	Proportion
1	27.88	30.50	31.25	24.95	20.00	0.3636
2	23.10	33.92	33.87	30.64	24.33	0.3545
3	18.41	24.47	30.24	30.35	17.65	0.1545
4	28.57	26.07	21.93	18.86	19.50	0.1274

Source: Authors' work

As much as with the previous linkage method, the Ward's method produces clusters that are difficult to describe by only looking at the average group values foe each dimension. The first two groups seem to consist of subjects that scored pretty high value to all five dimensions. Unlike these cases, the third group achieved on average lower scores at observing and non – reactivity, whiles the fourth group scores less at non – judging and non – reactivity. However, these last groups are unstable, therefore we cannot rely on a specific profile they may point toward.

Our next step in the investigation is whether the clusters we detected are significant in explaining the final FFMQ score, and to what extend other variables like gender, age and public versus private employment may have a significant impact on the overall level of mindfulness.

3.3. Looking into some potential FFMQ determinants

In this subsection we conduct a regression analysis to explore whether the clusters we detected explain the FFMQ score, and how this score relates with other variables. While a similar analysis can be conducted based on the numerical values of each individual dimension, such an approach would fail to account for possible consistent patterns associated with similar scores of mindfulness, a phenomenon usually called cross – sectional dependence.

The model we fit based on our data is the following:

$$FFMQ = \beta_0 + \beta_1 Groups + \beta_2 Age + \beta_3 Gender + \beta_4 Working\ place + \varepsilon$$

where ε is the error term. The fact that the FFQM scores follow a normal distribution, as found in normality testing (section 3.2.) becomes a very comfortable result in this context. Table 7 provides the output of this model for each of the two clustering methods, along with the information regarding the statistical significance of the predictors.

A preliminary evaluation of the model that corresponds to the Average Linkage clustering method shows that the regressors are able to explain up to 60% of the variation in FFMQ overall score. The model is statistically significant at an overall level, although three of the background variables, gender, age and working place do not prove to be of help in explaining the variability in the overall level of mindfulness. More exactly, and according to our results, males seem to be less mindful than women, and also those subjects who work in private institutions scores less, on average, than those who work in public workplaces, but Table 7 presents the p – values that point toward differences that are not statistically significant. Age seems to be a variable that, at least in this framework, is not able to bring any contribution to the model.

However, of the four group identified in our preliminary clustering, the first – and most stable – group has been set as reference group. Compared with this cluster, the second one has an average a lower mindfulness score, a result that also holds for the fourth cluster. Unlike these two, the third group points toward a significantly higher FFMQ score compares with the reference.

The third column in Table 9 shows the results for the Ward's linkage method. The overall explanatory power of the model is slightly lower than in the first case, but the background variables prove to be as non- significant as before. The only notable difference is that working place and age became marginally significant, namely significant at 10% level.

The regression model for each clustering linkage method

Table 9

Model: FFMQ score	Estimated coefficients (Average Linkage method)	Estimated coefficients (Ward's distance method)
Intercept	131.04*** *(< 2e-16)*	126.141*** *(< 2e-16)*
Group:		
Group 1:	Reference	Reference
Group 2:	-26.575***(0.0097)*	12.153*** *(1.86e-06)*
Group 3:	34.823****(3.44e-08)*	-12.852*** *(7.32e-05)*
Group 4:	-26.642*** *(< 2e-16)*	-19.998*** *(1.96e-08)*
Gender		
Female:	Reference	Reference
Male:	-0.4892 *(0.812)*	-3.051 *(0.168)*
Working place		
Public:	Reference	Reference
Private:	-0.129 *(0.9502)*	3.758 *(0.089)*
Age	0.213 *(0.109)*	0.250 (0.0812)
R – Squared	60.07%	55.15%
Adjusted R – Squared	57.75%	52.54%
F - statistic	25.83 *(< 2.2e-16)*	21.11 *(5.111e-16)*

P – values in parentheses
Source: Authors' work

The second cluster provides on average significantly higher FFMQ scores than the reference group, while groups 3 and 4 are associated with lower average scores than the reference. However, since clusters 3 and 4 are unstable, we cannot claim that we detected a particular profile that leads to lower mindfulness scores than other profiles.

4. CONCLUSIONS, DISCUSSIONS AND FURTHER RESEARCH

The Five Facets Mindfulness Questionnaire is documented as one of the valid instruments aimed at measuring the level of mindfulness. It is built upon a questionnaire of 39 questions, as a valuable index that combines five different dimensions: observing, describing, awareness, non-judging and non-reactivity. The result of this combination is an overall score that helps ranking subjects according to their level of mindfulness.

The aim of our research was to look into the structure of this score and detect potential stable patterns that may explain the variability in mindfulness level. Based on a sample of 111 subjects, we conducted first a cluster analysis based on two different linkage methods and found four groups, suggesting four types of persons built upon the scores recorded to the five dimensions mentioned above. Of these four groups created by the Average Linkage method, one was highly stable, and included the majority of the participants, suggesting that among the subjects there is a significant degree of homogeneity. Two of the other groups pointed toward existing patterns, but without any warranty regarding the type of subjects belonging to these groups. One last group was highly unstable. The results obtained in this exploratory stage suggest that, at least in this particular sample and using the Average Linkage method, the majority of the participants share similar characteristics in what concerns the levels they scored to the five dimensions.

The Ward's method identified two groups that can be associated with consistent patterns, but the level of stability was weak: although we can admit that these patterns may exist, it is unclear which persons belong to which group. Since these groups account for nearly 70% of the participants, the idea of a certain homogeneity among subjects comes up again: the results suggest that the participants may switch the groups randomly, without creating any significant changes in the clustering results.

In the last stage we fitted a regression model aimed at explaining the variations in the mindfulness level, measured by the FFMQ score, based on the groups we detected, and other three background variables: gender, age and workplace. We found that although the background variables did not bring any contribution, all four groups proved to be statistically significant in each linkage case. This result suggest that the five dimensions of the FFMQ overall score cluster the participants in our study in groups that are systematically different in their mindfulness level, even for this particular sample some of the groups are highly unstable. This result invites to further investigations regarding possible relations between the patterns we found, and possible personality profiles the mindfulness level may be associated with.

One of the limits of our study consists of the fact that the Average Linkage method led to a group that includes the majority of the subjects. This result prevents us from fitting and comparing any models with interactions, models that could have clarified potential interplays between the background variables and the clusters. If, for example, a model with interaction between groups and working place is fitted for the Ward's method clustering, age becomes statistically significant, as well as the working place. That means that it is not only age, or the working place that impact the mindfulness overall score, but the interaction between the patterns we detected within the mindfulness dimensions and the background variables. Having in mind the size of the clusters in the Average Linkage case, such a model with interaction cannot be built for comparison.

Another important limit is the convenience sampling that has been used as the sampling method, while a replication of this investigation having more background variables might be a good idea for further research.

Despite the evident limitations discussed above, our work contributes to the existing literature in that it provides a methodological framework that can be applied to different, and preferably richer data, both in what concerns the preliminary, exploratory analysis, but also in the second stage of assessing the significance of the clusters. The value added to the literature consists of a deeper perspective over the components of the mindfulness overall score, components treated here not only as numbers that should be added to calculate a total value, but also as potential determinants of some specific, and stable patterns that are more likely to be associated with certain values of FFMQ scores than others.

Last, but not least, to our knowledge this is the first study on mindfulness applied to Romanian subjects. The fact that we found similarities between the values recorded by Romanian participants, and subjects belonging to other countries, is a first and important step in a deeper analysis that may concern the level of mindfulness in Romania, in the sense that it invites to further interventions to increase the mindfulness level based on similar models that have been implemented worldwide.

References

1. **Aguado, J., Luciano, J. V., Cebolla, A., Serrano-Blanco, A., Soler, J., & García-Campayo, J.**, 2015. *Bifactor analysis and construct validity of the five facet mindfulness questionnaire (FFMQ) in non-clinical Spanish samples*. Frontiers in psychology, 6, 404.
2. **Anscombe, F.J., Glynn, W.J.**, 1983, *Distribution of kurtosis statistic for normal statistics*. Biometrika, 70, 1, 227-234
3. **Baer, R. A.**, 2011, *Measuring mindfulness*. Contemporary Buddhism, 12(01), 241-261.
4. **Barros, V. V. D., Kozasa, E. H., Souza, I. C. W. D., & Ronzani, T. M.**, 2014, *Validity evidence of the Brazilian version of the Five Facet Mindfulness Questionnaire (FFMQ)*. Psicologia: Teoria e Pesquisa, 30(3), 317-327.
5. **Bergomi, C., Tschacher, W., & Kupper, Z.**, 2013, *The assessment of mindfulness with self-report measures: Existing scales and open issues*. Mindfulness, 4(3), 191-202.
6. **Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R.**, 2011, *Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form*. Assessment, 18(3), 308-320.
7. **Bollen, K., & Lennox, R.**, 1991, *Conventional wisdom on measurement: A structural equation perspective*. Psychological bulletin, 110(2), 305.
8. **Bollen, K. A., & Stine, R. A.**, 1992, *Bootstrapping goodness-of-fit measures in structural equation models*. Sociological Methods & Research, 21(2), 205-229.
9. **Brock, G., Pihur, V., Datta, S. and Datta, S.**, 2008, *clValid: An R Package for Cluster Validation* Journal of Statistical Software 25(4) http://www.jstatsoft.org/v25/i04
10. **Brown, K. W., Kasser, T., Ryan, R. M., Linley, P. A., & Orzech, K.**, 2009, *When what one has is enough: Mindfulness, financial desire discrepancy, and subjective well-being*. Journal of Research in Personality, 43(5), 727-736.

11. **Cebolla, A., García-Palacios, A., Soler, J., Guillen, V., Baños, R., & Botella, C.**, 2012, Psychometric properties of the Spanish validation of the Five Facets of Mindfulness Questionnaire (FFMQ). The European Journal of Psychiatry, 26(2), 118-126.
12. **D'Agostino, R.B.**, 1970, *Transformation to Normality of the Null Distribution of G1*. Biometrika, 57, 3, 679-681.
13. **Datta, S.** and **Datta, S.**, 2003, *Comparisons and validation of statistical clustering techniques for microarray gene expression data*. Bioinformatics 19(4): 459-466
14. **Datta, S.** and **Datta, S.**, 2006, *Methods for evaluating clustering algorithms for gene expression data using a reference set of functional classes*. BMC Bioinformatics 7:397 http://www.biomedcentral.com/1471-2105/7/397
15. **Davis, D. M.** & **Hayes, J. A.**, 2011, What are the benefits of mindfulness? A practice review of psychotherapy-related research. *Psychotherapy*, *48*(2), 198.
16. **Deng, Y. Q., Liu, X. H., Rodriguez, M. A., & Xia, C. Y.**, 2011, *The five facet mindfulness questionnaire: psychometric properties of the Chinese version*. Mindfulness, 2(2), 123-128.
17. **Gerbing, D. W., & Anderson, J. C.**, 1988, *An updated paradigm for scale development incorporating unidimensionality and its assessment*. Journal of marketing research, 186-192.
19. **Giovannini, C., Giromini, L., Bonalume, L., Tagini, A., Lang, M., & Amadei, G.**, 2014, *The Italian five facet mindfulness questionnaire: a contribution to its validity and reliability*. Journal of Psychopathology and Behavioral Assessment, 36(3), 415-423
19. **Fiol, C. M., & O'Connor, E. J.**, 2003, *Waking up! Mindfulness in the face of bandwagons*. Academy of management review, 28(1), 54-70.
20. **Flynt, A., & Dean, N.**, 2016, **A survey of popular r packages for cluster analysis**. Journal of Educational and Behavioral Statistics, 41(2), 205-225.
21. **Goldman-Schuyler, K., Skjei, S., Sanzgiri, J., & Koskela, V.**, 2017, *"Moments of Waking Up" A Doorway to Mindfulness and Presence*. Journal of Management Inquiry, 26(1), 86-100.
22. **Hayes AM, Feldman G**., 2004, *Clarifying the construct of mindfulness in the context of emotion regulation and the process of change in therapy*. Clinical Psychology: Science and Practice. 2004;11:255–262.
23. **Hennig, C.**, 2007, *Cluster-wise assessment of cluster stability*. Computational Statistics and Data Analysis, 52, 258-271.
24. **Hennig, C.**, 2008, *Dissolution point and isolation robustness: robustness criteria for general cluster analysis methods*. Journal of Multivariate Analysis 99, 1154-1176.
25. **Heeren, A., Douilliez, C., Peschard, V., Debrauwere, L., & Philippot, P.**, 2011, *Cross-cultural validity of the Five Facets Mindfulness Questionnaire: Adaptation and validation in a French-speaking sample*. Revue européenne de psychologie appliquée/European Review of Applied Psychology, 61(3), 147-151
26. **Hopkins B, Skellam JG.**, 1954, *A new method for determining the type of distribution of plant individuals*. Ann Bot Lond 1954; 18: 213–227.
27. **Hou, J., Wong, S. Y. S., Lo, H. H. M., Mak, W. W. S., & Ma, H. S. W**., 2014, *Validation of a Chinese version of the Five Facet Mindfulness Questionnaire in Hong Kong and development of a short form*. Assessment, 21(3), 363-371.
28. **Howell, A. J., Digdon, N. L., & Buro, K**., 2010, *Mindfulness predicts sleep-related self-regulation and well-being*. Personality and Individual Differences, 48(4), 419-424.
29. **Keng, S. L., Smoski, M. J., & Robins, C. J.**, 2011, *Effects of mindfulness on psychological health: A review of empirical studies.* Clinical psychology review, 31(6), 1041-1056

30. **Komsta L., Novomestky F.**, 2015, *Package "moments", available on R – cran*, at http://www.r-project.org, http://www.komsta.net/
31. **Lawson, R.G.** and **Jurs, P.C.,** 1990, *New index for clustering tendency and its application to chemical problems*. Journal of Chemical Information and Computer Sciences, 1990, 30(1):36-41
32. **Mundfrom, D. J., Shaw, D. G., & Ke, T. L.**, 2005, *Minimum sample size recommendations for conducting factor analyses*. International Journal of Testing, 5(2), 159-168.
33. **Neuser, Ninfa J.**, 2010, *Examining the Factors of Mindfulness: A Confirmatory Factor Analysis of the Five Facet Mindfulness Questionnaire* (Doctoral dissertation, Pacific University). Retrieved from:http://commons.pacificu.edu/spp/128
34. **Pihur, V., Datta, S., and Datta, S**., 2007, *Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach*, Bioinformatics, 23(13):1607-1615
35. **Purser R., Milillo J.**, 2015, *Mindfulness Revisited*, Journal of Management Inquiry Vol 24, Issue 1, pp. 3 - 24, doi 10.1177/1056492614532315
36. **Reise, S. P., Waller, N. G., & Comrey, A. L**., 2000, *Factor analysis and scale revision*. Psychological assessment, 12(3), 287.
37. **Sauer, S., Lynch, S., Walach, H., & Kohls, N**., 2011, *Dialectics of mindfulness: implications for western medicine*. Philosophy, Ethics, and Humanities in Medicine, 6(1), 10.
38. **Shapiro, S. L.**, 2009, *The integration of mindfulness and psychology*. Journal of clinical psychology, 65(6), 555-560.
39. **Sirois, F. M., & Tosti, N.**, 2012, *Lost in the moment? An investigation of procrastination, mindfulness, and well-being*. Journal of Rational-Emotive & Cognitive-Behavior Therapy, 30(4), 237-248.
40. **Sneath, P. H. A.**, 1957, *Some thoughts on bacterial classification*, Journal of General Microbiology 17, 184-200.
41. **Sokal R. R., Michener C. D.**, 1958, *A statistical method for evaluating systematic relationships [J]*, Univ. Kans. Sci. Bull, 28, 1409–1438.
42. **Stephens, M.A**., 1986, *Tests based on EDF statistics*. In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.
43. **Sugiura, Y., Sato, A., Ito, Y., & Murakami, H**., 2012, *Development and validation of the Japanese version of the Five Facet Mindfulness Questionnaire*. Mindfulness, 3(2), 85-94.
44. **Thode Jr., H.C**., 2002, *Testing for Normality*. Marcel Dekker, New York.
45. **Tran, U. S., Glück, T. M.**, and **Nader, I. W**., 2013, *Investigating the five facet mindfulness questionnaire (FFMQ): construction of a short form and evidence of a two-factor higher order structure of mindfulness*. J. Clin. Psychol. 69, 951–965. doi: 10.1002/jclp.21996
46. **Ward, J. H., Jr.**, 1963, *Hierarchical Grouping to Optimize an Objective Function*, Journal of the American Statistical Association, 58, 236–244.
47. **Weick K., Putnam T.**, 2015, *Organizing for Mindfulness. Eastern Wisdom and Western Knowledge*, Journal of Management Inquiry, vol 15, issue 3, pp 275 - 287
48. **Williams, M. J., Dalgleish, T., Karl, A., & Kuyken, W**., 2014, *Examining the factor structures of the five facet mindfulness questionnaire and the self-compassion scale*. Psychological assessment, 26(2), 407.
49. **Zhang, G., & Preacher, K. J.**, 2015, *Factor Rotation and Standard Errors in Exploratory Factor Analysis*. Journal of Educational and Behavioral Statistics, 1076998615606098.