Fuzzy system approach to symptoms in schizophrenia

CORNELIU MOSOIU¹, ALIN GILBERT SUMEDREA¹, VICTORIA BURTEA², PETRU IFTENI²
¹Department of Psychology, Lucian Blaga University, Sibiu, 40 Victoriei Blvd, 550024
²Department of Medicine, Transilvania University Brasov, 25 Eroilor Blvd, 500030
ROMANIA
mosoiu_corneliu@yahoo.com

Abstract
Schizophrenia has long been characterized by significant variability in symptoms, clinical profiles, and course of illness. To reduce the heterogeneity of this complex disorder, we try to identify homogeneous syndrome that will facilitate the identification of links between symptoms. This article aim is to investigate symptoms in schizophrenia with the aid of fuzzy system and evaluation scale (Positive and Negative Syndrome Scale - PANSS).

Key-Words: schizophrenia, fuzzy system, evaluation scale

1. Introduction
Schizophrenia has been characterized by significant variability in symptoms, course of illness, and clinical profiles. To reduce the heterogeneity of this complex disorder, it is important to identify subtypes (clusters of symptoms) that will improve the identification of links between symptoms. To explain the pattern within a set of observed variables, we used factors that attempts to identify underlying variables. This method is often used to generate hypotheses regarding causal mechanisms, to screen variables for subsequent analysis or in data reduction to identify factors that explain most of the variance that is observed in a much larger number of variables. Also objects belong to the same cluster if this one defines a concept common to all that objects. In other words, objects are grouped according to their fit to descriptive concepts, not according to simple similarity measures. This article aim is to investigate the subtypes of schizophrenia with the aid of fuzzy systems that was applied to the Positive and Negative Syndrome Scale (PANSS) that quantify the severity of symptomatology. Quantification has been regarded as an important topic in our theoretical approach. The fuzzy quantifiers provide an example of such formal concept. The use of fuzzy quantifiers is one way to attach semantic labels to fuzzy sets and can also be seen as flexible tools making the existing fuzzy models more readable. Although the application of quantifier-based models can be found in the literature [1], such research in psychopathology and clinical psychiatry is relatively limited. However, several definitions of fuzzy quantifiers have been proposed including so-called absolute quantifiers and relative quantifiers. One of the main issues in the use of fuzzy quantification on fuzzy models is that most of the existing techniques do not deal adequately with practical interests such as monotonicity and duality of the quantifiers. This has significantly restricted the take-up of fuzzy quantification techniques [2]. This paper proposes a method which uses fuzzy quantifiers to cluster symptoms, based on existing factor analysis-based modeling of schizophrenia symptomatology.

2. Material and methods
Participants
Fifty-six individuals with DSM-IV-TR [3] diagnoses of schizophrenia were recruited after provided informed consent to participate in the study. Participants received either typical or atypical antipsychotics, depending on the protocol being conducted when the patient was studied.

Inclusion and Exclusion Criteria
Included were participants who met the following criteria: age 18 or older and a DSM-IV-TR diagnosis of schizophrenia disorder. Excluded were individuals who had other DSM-IV-TR Axis I or II disorders.

Diagnosis
All participants were evaluated by a psychiatrist. Diagnosis was assessed using a semi-structured diagnostic interview and review of medical records that is used to gather demographic, diagnostic, and course of illness information for DSM-IV Axis I and II disorders. Diagnoses were achieved by consensus among interviews, past medical records, and symptom ratings.

Methods
The aim of the study was to investigate evidence of a more complex structure of schizophrenia than the relatively simple negative-positive symptoms concept. For this purpose, fuzzy system approach based on factors model of schizophrenia was applied to a sample of patients who had been scored on the Positive and Negative Syndrome Scale (PANSS). The PANSS is a 30-item, 7-point (1=Absent–7=Extreme) rating scale. Total score on PANSS range from 30 (no illness) to 210 points [4].
A potentially confusing feature of the PANSS is that even those without any mental ill will score 30. This means that 30 must be subtracted from the patient’s score in order to gain meaningful information so we consider only scores above 30 points. PANSS was chosen because of its known validity and because it includes negative, positive, and general psychopathology items (such as depression and anxiety). Those items have been found to be internally consistent and highly reliable. In our study scale items were rated according to the definitions and criteria provided by the manual and structured interview (SCI-PANSS).

There are many reports on the factor (syndrome) structure of PANSS items, with much controversy over whether data best fit a three-, four-, five- or even six-factor solution [5]. This might reflect the ambiguous definitions of some symptom items [6]. Data gathered from this assessment procedure are applied to the PANSS ratings. Each of the 30 items is accompanied by a specific definition as well as detailed anchoring criteria for all seven rating points.

These seven points represent increasing levels of psychopathology, as follows:
- 1-absent
- 2-minimal
- 3-mild
- 4-moderate
- 5-moderate severe
- 6-severe
- 7-extreme

In assigning ratings, one first considers whether an item is at all present, as judging by its definition.
- A rating of 1 denotes that item is absent. If it is present we must determine its severity by reference to the particular criteria from the anchoring points.

For ambiguous situations of ratings, the highest applicable rating point is always assigned, even if the patient meets criteria for lower points as well. In judging the level of severity, the rater must utilise all information available in deciding which anchoring point best characterises the patient’s functioning and rate accordingly, whether or not all elements of the description are observed.

The rating points of 2 to 7 correspond to incremental levels of symptom severity:
- A rating of 2 (minimal) denotes questionable or subtle or suspected pathology, or the extreme end of the normal range.
- A rating of 3 (mild) denotes a symptom whose presence is clearly established but not pronounced and interferes little in functioning.
- A rating of 4 (moderate) denotes a symptom which, though representing a serious problem, either occurs only occasionally or intrudes on daily life only to a moderate extent.

- A rating of 5 (moderate severe) denotes marked manifestations that distinctly impact on one’s functioning but are not all-consuming and usually can be contained at will.
- A rating of 6 (severe) represents gross pathology that is present very frequently, proves highly disruptive to one’s life, and often calls for direct supervision.
- A rating of 7 (extreme) refers to the most serious level of psychopathology and also the manifestations drastically interfere in most or all major life functions, typically necessitating close supervision and assistance in many areas.

Fuzzy logic known as a basic concept for embedding structured human knowledge into workable algorithms and was initially introduced by Zadeh in order to improve tractability, robustness and low-cost solutions for real world problems. The theory of fuzzy sets is a theory of graded concepts and membership grades. Defined fuzzy sets or classes for each variable allows intermediate grades of membership in them which means each set could have elements that belong partially to it; the degree of belonging is called membership functions ranging from 0 to 1. If X is the Universe of discourse and its elements be denoted as x, in contrast with crisp set, the fuzzy set A of the X has a characteristics function associated to it:
\[ \mu : X \rightarrow [0,1] \]
- \( \mu (X) = 1 \) if x is totally in A
- \( \mu (X) = 0 \) if x is not in A
- \( 0 < \mu < 1 \) if x is partly in A

For expert evaluation we used qualifiers of symptoms. Those are used to grade the degree to which a particular attribute is given. There are five fundamental judgment dimensions:
- Intensity \([I]\), e.g., not, a little, rather, very, extremely;
- Frequency \([F]\), e.g., never, sometimes, often, always;
- Probability \([P]\), e.g., unlikely, hardly, possibly, for sure.
- Quality \([Q]\), e.g., bad, acceptable, satisfactory, excellent;
- Agreement with statements \([S]\), e.g., don’t accept, agree, true for me.

All collected words/expressions were allocated to these 5 categories, and further ones were created by combining single modifiers into combined ones, e.g., very often \((I+’F’\)), not likely \((I+’P’\)), rather good \((I+’Q’\)), often true for me \((F+’S’\)). To determine these relationships between symptoms and diagnoses, occurrence and confirmation, they have been defined as fuzzy sets. Consequently the psychiatrist had to specify these relationships by only giving answers like always, almost always, very often, often, unspecific, seldom, very seldom, almost never, and never. Also the psychiatrist choose fuzzy sets which have been defined by their membership functions. In this case, the membership functions for values of occurrence and...
confirmability could be defined as relative frequencies. Critics of this fuzzy set approach occur because disagreement exist amongst raters, especially regarding the measurement of fuzziness and the properties of the resulting membership functions. We presented a model for determining the measurement function based on consensus among human expert raters of symptomatology. For our study membership is defined as a function of the distance \( d(x) \) between a given object \( x \) and \( x \) standard (ideal, “golden standard”):

hence \( d(x) = 0 \Rightarrow \mu(x) = 1 \) and \( d(x) = 1 \Rightarrow \mu(x) = 0 \).

It has the additional advantage that it can easily be linearised by transformation, and easy to check.

Therefore, a fuzzy membership function \( \mu(x) \) indicate the degree of belonging of some element \( x \) to the universe of discourse \( X \) and it maps each element of \( X \) to a membership grade between 0 and 1. The grade of membership (\( \mu \)) for all levels and symptoms thus describes a fuzzy set. Grade of membership is normally a real number between 0 and 1. The higher the number, the higher the membership, but finally membership depends on one’s own view. There is no formal basis for how to determine the grade of membership.

The membership grade is a precise, subjective measure that depends on the context and also is different from a statistical probability distribution.

It is not easy to find common features among these approaches, although we consider that:

1. membership functions are continuous,
2. membership functions map an interval \([a, b]\) to \([0, 1]\), \( \mu[a, b] = [0, 1] \).
3. membership functions are (a) either monotonically increasing, or (b) monotonically decreasing, or (c) could be divided into a monotonically increasing or decreasing part.

The determination of any membership function is possible with four parameters:

- HEIGHT or magnitude (usually normalized to 1)
- WIDTH (of the base of function),
- SHOULDERING (height at maximum if an outer function. Shouldered functions evaluate as 1.0 past their center)
- CENTER points (center of the member function shape)
- OVERLAP (typically about 50% of width but can be less).

Based on expert opinion on fuzzy qualifiers, from different membership functions we choose triangular membership function (TMF) which is widely in use as following:

For evaluation of schizophrenia symptomatology based on PANSS scale description we used triangular membership functions (TMF) with 3 memberships defined as low, medium, and high. The memberships were defined as low, if TMF was between 0 and 0.5 with center based at 0.25; medium if TMF was between 0.25 and 0.75 with center based at 0.5; and high if full TMF was between 0.5 and 1 with center based at 0.5.

Expert knowledge was used and can be expressed using linguistic variables, which are described by fuzzy inference system. Linguistic rules describing the system consist of two parts; an antecedent block (between the IF and THEN) and a consequent block (following THEN). Depending on the system, it may not be necessary to evaluate every possible input combination, since some may rarely or never occur. The expert knowledge for these variables can be formulated as a rules like, for example, IF feature A low AND feature B medium AND feature C medium AND feature D medium AND feature E high THEN Cluster = cluster 1. The rules can be combined in a table that is rule based.

3. Results

The results of the approach based on principal component analysis after the components were rotated using varimax method suggest that schizophrenia has 5 distinct and interpretable factors.

| Component Matrix* |
|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                   | A               | B               | C               | D               | E               |
| p1                | .415            | .463            | .240            | .332            | .576            |
| p3                | .315            | .287            | .777            | .005            | .109            |
| p5                | .705            | .208            | .456            | .209            | .143            |
| g9                | -.366           | .252            | .363            | .361            | .673            |
| g1                | .387            | .563            | .405            | .435            | .030            |
| p2                | .184            | .251            | .919            | .001            | .011            |
| p4                | .102            | .854            | .180            | .101            | .076            |
The major finding of the present study is that the structure of schizophrenia symptomatology consists of 5 distinct syndromes based on fuzzy rules used. Future studies should replicate this structure.

The implication for the rating scale is that items which load high on more than one factor must be replaced by two or more items, each of which load on a single factor, which results in lengthier scales. This study supports that schizophrenia is made up of 5 distinct syndrome as we show based on our fuzzy approach. As our goal was to assess the structure of schizophrenia, the analysis included only participants who met overall criteria for it [8]. By making this type of evaluation, usually done by an experienced operator, fewer rules can be evaluated, thus simplifying the processing logic and perhaps even improving the fuzzy logic system that we used. Instead of assuming that everything must be defined crisply that is the binary view, the fuzzy view holds that most things fall somewhere in between, that is in varying degree of membership, and should be analyzed as such. So any syndrome specified is not absolute and can be expressed within a probability between 0 and 1, that is, between certainly present and certainly absent. Fuzzy sets incorporates all of this information as well as the predictors suggested by human heuristics into formal rule bases for data analysis. The inability of statistics to analyze the degree of causal relationships from natural language quantifiers is the essence of using the fuzzy inference rules.

**4. Conclusions**

The major finding of the present study is that the structure of schizophrenia symptomatology consists of 5 distinct syndromes based on fuzzy rules used. Future studies should replicate this structure.

The implication for the rating scale is that items which load high on more than one factor must be replaced by two or more items, each of which load on a single factor, which results in lengthier scales. This study supports that schizophrenia is made up of 5 distinct syndrome as we show based on our fuzzy approach. As our goal was to assess the structure of schizophrenia, the analysis included only participants who met overall criteria for it [8]. By making this type of evaluation, usually done by an experienced operator, fewer rules can be evaluated, thus simplifying the processing logic and perhaps even improving the fuzzy logic system that we used. Instead of assuming that everything must be defined crisply that is the binary view, the fuzzy view holds that most things fall somewhere in between, that is in varying degree of membership, and should be analyzed as such. So any syndrome specified is not absolute and can be expressed within a probability between 0 and 1, that is, between certainly present and certainly absent. Fuzzy sets incorporates all of this information as well as the predictors suggested by human heuristics into formal rule bases for data analysis. The inability of statistics to analyze the degree of causal relationships from natural language quantifiers is the essence of using the fuzzy inference rules.

**Limitations**

One limitation of the present study is the relatively modest sample size (56 subjects). Also a reliable principal component analysis for fuzzy approach requires a higher participant/variable ratio than we achieved in this study. Thus, given the sample size and possibility that the resulting factor structure may not be reliable, these results need to be replicated with a larger sample.

### Table 1. Component matrix for PANSS items

<table>
<thead>
<tr>
<th>Item</th>
<th>p6</th>
<th>p7</th>
<th>n1</th>
<th>n2</th>
<th>n3</th>
<th>n4</th>
<th>n5</th>
<th>n6</th>
<th>n7</th>
<th>g2</th>
<th>g3</th>
<th>g4</th>
<th>g5</th>
<th>g6</th>
<th>g7</th>
<th>g8</th>
</tr>
</thead>
<tbody>
<tr>
<td>p6</td>
<td>.668</td>
<td>.553</td>
<td>.118</td>
<td>.101</td>
<td>.415</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p7</td>
<td>.268</td>
<td>.732</td>
<td>.177</td>
<td>.007</td>
<td>.486</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n1</td>
<td>.481</td>
<td>.290</td>
<td>.120</td>
<td>.585</td>
<td>.510</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n2</td>
<td>.873</td>
<td>.037</td>
<td>.123</td>
<td>.283</td>
<td>.196</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n3</td>
<td>.944</td>
<td>.125</td>
<td>.197</td>
<td>.007</td>
<td>.116</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n4</td>
<td>.780</td>
<td>.089</td>
<td>.231</td>
<td>.527</td>
<td>.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n5</td>
<td>.186</td>
<td>.091</td>
<td>.787</td>
<td>.038</td>
<td>.337</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n6</td>
<td>.787</td>
<td>.010</td>
<td>.402</td>
<td>.086</td>
<td>.162</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n7</td>
<td>.700</td>
<td>.587</td>
<td>.035</td>
<td>.045</td>
<td>.209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g2</td>
<td>.376</td>
<td>.792</td>
<td>.183</td>
<td>.347</td>
<td>.106</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g3</td>
<td>.087</td>
<td>.090</td>
<td>.264</td>
<td>.884</td>
<td>.086</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g4</td>
<td>.280</td>
<td>.851</td>
<td>.087</td>
<td>.041</td>
<td>.034</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g5</td>
<td>.219</td>
<td>.210</td>
<td>.190</td>
<td>.098</td>
<td>.816</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g6</td>
<td>.421</td>
<td>.181</td>
<td>.061</td>
<td>.754</td>
<td>.181</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g7</td>
<td>.108</td>
<td>.074</td>
<td>.125</td>
<td>.198</td>
<td>.643</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g8</td>
<td>.363</td>
<td>.005</td>
<td>.056</td>
<td>.881</td>
<td>.138</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g9</td>
<td>.033</td>
<td>.143</td>
<td>.063</td>
<td>.890</td>
<td>.376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g10</td>
<td>.466</td>
<td>.149</td>
<td>.654</td>
<td>.309</td>
<td>.265</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g11</td>
<td>.547</td>
<td>.036</td>
<td>.549</td>
<td>.008</td>
<td>.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g12</td>
<td>.417</td>
<td>.844</td>
<td>.045</td>
<td>.266</td>
<td>.297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g13</td>
<td>.086</td>
<td>.271</td>
<td>.654</td>
<td>.321</td>
<td>.395</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g14</td>
<td>.400</td>
<td>.039</td>
<td>.391</td>
<td>.382</td>
<td>.668</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g15</td>
<td>.409</td>
<td>.299</td>
<td>.116</td>
<td>.296</td>
<td>.155</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Component matrix show loadings on factors for all items of PANSS and we used the values of factors for each item as input for triangular membership function and therefore for fuzzy inference rules as stated above.

The first cluster contains patients having high scores on negative items (emotional withdrawal, poor rapport, passive/apathetic social withdrawal, lack of spontaneity and flow of conversation, stereotyped thinking) and low scores on positive and define the *negative syndrome*.

The second cluster contains patients having high scores on positive (excitement, suspiciousness, hostility), and general psychopathological items (tension, anxiety, impulse control), had low scores on all the negative (except stereotyped thinking), of the PANSS and define *excitement syndrome* [7].

The third cluster contains patients having high scores on some positive items (conceptual disorganization), and general psychopathological items (lack of judgement and insight, disturbance of volition, preoccupation), had low scores on negative items and define a *cognitive syndrome*.

The fourth cluster contains patients that were higher on general psychopathological items (depression, somatic concerns, guilt feelings, poor attention) and had low scores on the positive and negative items and define *mood syndrome*.

The fifth cluster contains patients that have high scores on positive (delusion), and general psychopathological items (unusual thought content, mannerism and posturing, motor retardation, active social avoidance) and define the *positive syndrome*.
Data gathered from this small sample of schizophrenic patients show a five factors model instead of positive/negative model of disease. In order to improve validity and reliability of clustering method further studies based on larger samples need to be performed and maybe to apply slightly different membership function and inference rules.

It may also be argued that the current analysis was done on one clinician's heuristics, and perhaps the results may be better if another more experienced clinician was to select input memberships. We believe that knowledge of experienced clinicians with this type of psychopathology are likely to select mostly similar input memberships function as ours.

References: