Mentalization and theory of mind in borderline personality organization: exploring the differences between affective and cognitive aspects of social cognition in emotional pathology

Dominika Górska, Monika Marszał

Institute of Psychology, Adam Mickiewicz University, Poznań, Poland
Head: prof. dr hab. J. Brzeziński

Summary

Aim: This article addresses the problem of explaining emotional pathology (levels of personality organization) using the concepts of theory of mind (ToM) and mentalization. Although these terms are used interchangeably to describe the “ability to interpret the behavior of others in terms of mental states,” they do not have identical status in emotional disorders. ToM refers to a “cold” knowledge, whereas mentalization requires the activation of relational and emotional representations, as well as processing of emotional experience (whether reflection or defense). The aim of the study was to compare the cognitive (ToM) and affective (mentalization) aspects of “understanding the behavior of others in terms of mental states” in the clinical group – consisting of patients with borderline personality organization (N = 30); and the control group (N = 30).

Method: The Borderline Personality Inventory was used as a diagnostic questionnaire for the organization of personality, the Strange Stories Test was employed to measure ToM, and the Mental States Task instrument measured mentalization.

Results: With respect to mentalization, different patterns of results were obtained: the activation of overwhelming mental states and primitive defenses in the clinical group; as well as the inhibition of the recognition of mental content by defenses, such as denial and suppression of emotions, in the control group. No differences were observed in ToM between groups.

Conclusions: In explaining the personality organization levels, only the affective, and not the cognitive, aspects of “understanding the behavior of others in terms of mental states” are significant. People with borderline personality organization, as well as healthy individuals, use “cold” knowledge about internal states. However, the activation of relational and emotional representations triggers different mental states in both groups.

Key words: theory of mind, mentalization, borderline personality organization
Introduction

In descriptions of the ability to interpret the behavior of others in terms of mental contents (such as beliefs, desires, feelings, needs, thoughts, and intentions), the concepts of theory of mind (ToM) and mentalization are frequently used interchangeably [1]. Although some aspects of these notions are similar, they do not possess identical status and importance in explaining emotional disorders. This is evidenced not only by theoretical conceptualizations involving the various factors and mechanisms behind development (both normal and impaired), but also by the results of empirical research, particularly concerning emotional and personality disorders. This paper is an attempt to answer the questions about the differences between theory of mind (ToM) and mentalization, and to indicate their importance for explaining emotional pathology—here represented by the levels of personality organization (LPO) [2].

Theory of mind is defined as the ability to attribute mental states to others in order to understand and predict social behavior [3]; it is usually explored through the false-belief paradigm. ToM involves two processes: the detection (and decoding) of mental states, and reasoning about mental states. This latter includes more advanced skills, such as predicting behavior on the basis of false beliefs [4]. ToM has been criticized as too narrow a term, as discounting the relational and emotional aspects of understanding the behavior of others [5]. Therefore, attempts have been made to separate the different aspects or functions of ToM, especially in order to distinguish its cognitive and emotional components [6]. Mentalization, on the other hand, is a specific form of social cognition based on the performance of imaginative mental activity that allows an understanding of the behavior of other people in terms of their intentions, needs, desires, or goals [7]. Mentalization is described as a multidimensional construct incorporating three dimensions: modes (implicit and explicit), two subjects (self and others), and two aspects (cognitive and affective) [8]. Besides understanding the mental states of others, it subsumes the regulation and transformation of one’s own emotions, through obtaining understandings of others’ intentions, feelings, and beliefs. Mentalization thus comes to play an important role in the mechanisms of emotional processing and regulation, which are especially important for people with emotional disorders.

In addition to the differing theoretical conceptualizations of ToM and mentalization, research results also confirm that the distinction between these two terms is significant from theoretical and clinical points of view. The deficits of ToM have been observed in autism spectrum disorders [9], as well as in disorders characterized by severe deficits and impoverishment in multiple functions, such as schizophrenia [10]. However, studies of emotional psychopathology (i.e., of anxiety disorders) have shown that there is no difference between the clinical and the control group in the basic tests relating to ToM [11]. Also, no association has been shown between ToM and the patterns of defense mechanisms [12] and the quality of the attachment relationships [13]. In addition, a mother’s level of verbal skill happens to be a predictor of a high level of ToM in her children, but the mother’s frequency of reference to mental states and
secure attachment happens not to be [14]. These studies suggest that ToM and emotional functioning are relatively independent of each other.

At the same time, research on mentalization has revealed significant deficits in this function in emotional disorders, such as deficits in mentalization in borderline disorder [15], the relationship between mentalizing and the maturity of defense mechanisms [16], and the level of personality organization [2, 17]. It has also been shown that the ability to mentalize is closely linked to attachment security, and that its development is dependent on the quality of the emotional relationship between the mother and the child [18]. The basic feature in the development of the ability to mentalize is the mirroring of the child’s internal states by his or her mother [19]. Threatening experiences can lead to defensive attempts to block the inference of the mental states of the caregiver. The results also indicate that the quality of the child’s bond of attachment to the mother can be predicted from the mother’s ability to mentalize, [20] and that child’s subsequent ability to mentalize can be inferred from the child’s level of attachment [21].

Although the definition of “explaining the behavior of others in terms of their thoughts, feelings and intentions” matches both the concept of theory of mind and the concept of mentalization, in light of the current research, both terms can be seen to differ in key areas:

- **emotional arousal:** inferences about the mental states of others in ToM do not assume personal emotional involvement in the story of a person whose mind is being recognized (there is no activation of the attachment system)—a feature that is necessary in mentalization [22];

- **relational aspects:** in ToM, there is an inference about mental states, in which the figure is another person not emotionally related to the subject. This is opposed to mentalization, which refers to a mind of a person remaining in an emotional relationship (whether real or imaginary) with the subject [8, 7]. This is associated with a dyadic (relational) structure of representations, which is activated in the case of mentalization postulated in both psychodynamic theories (“self in an emotional relationship with the object”) [2] and in attachment theory (internal working models) [21].

- **emotional experience processing:** in mentalization, emotional activation is accompanied by the activation of defense mechanisms or reflection. The interaction of these two is disclosed in the mentalization ability [23]. Neither of these factors is taken into account in the model of ToM;

- **regulatory functions:** in addition to understanding one’s own mental content, the ability to mentalize assumes the transformation and regulation of one’s own emotions because of the understanding gained of others’ intentions, feelings, and beliefs. Nothing comparable happens in ToM [24];

- **developmental roots:** it has been suggested that difficulties in mentalization may serve as defenses against intense emotional experiences, including early trauma. ToM deficits are associated with dissociation in the development of certain cognitive functions [9].
For the purpose of this study, the following definitions have been employed. It is assumed that ToM requires understanding of the influence of belief on an individual’s behavior and the ability to take the perspective of the person (both cognitively and emotionally), but—at the moment of reasoning—without the involvement of a relationship or the activation of an emotional state. In turn, mentalization, as it is defined here, is a much more advanced ability, being more than the ability to understand the emotions of other people or feel like another person feels (an emotional ToM) [6]. It also goes well beyond the understanding of the causes of emotions occurring in the other person. Mentalization involves emotional arousal in a relationship with someone whose mind is being recognized, and thus activation of the internal emotional and relational representation [2, 25, 26] as an effect of the stimulus coming from the relationship. It thus assumes processing of one’s own emotional experience (including defense and reflection), together with the building of hypotheses about one’s own mental contents.

The most discussed disorder in the context of deficits in mentalization is borderline pathology. Depending on which aspects of the ability to reason about mental states is under analysis, the results reveal different tendencies: no deficits or even better performance on cognitive aspects [27, 28], or else significantly reduced functioning in emotional and relational aspects [22], accompanied by severe emotional dysregulation [15, 29]. This study combines the analysis of both these components, based on the conceptualization of borderline pathology proposed by Kernberg, which includes aspects of personality structure. The whole concept of levels of personality organization (LPO) [2] distinguishes three levels—borderline (BPO), neurotic, and integrated—with regard to the following criteria: maturity of defense mechanisms, reality-testing ability, the level of identity diffusion, and the maturity of object relations. Therefore, the present study was designed to compare the theory of mind and mentalization in groups with different levels of personality organization (BPO and above BPO). The method was used to examine the cognitive aspects (ToM) on one hand and the emotional aspects (the ability to mentalize in emotional arousal) on the other. It was predicted that differences in mentalization would occur between the groups, even in the absence of these differences in ToM. Moreover, we expected the relative independence of ToM and mentalization.

**Materials and method**

**Participants**

The clinical group (N = 30) consists of patients from the Hospital for Mental Diseases in Międzyrzecz and Medison Private Health Care in Koszalin, who have been diagnosed as having BPO on the basis of the results of the screening test. The control group consists of participants from the general population (N = 30), not participating in psychiatric treatment or psychotherapy. These participants are qualified as having personality levels higher than borderline based on the results of the screening test.
To eliminate the influence of age, gender, and education on the differences between the groups, a procedure of matching pairs was used. In each group, there were 17 women and 13 men, aged $M = 30.77$, $SD = 9.21$, $min = 18$, $max = 50$. The highest levels of education obtained were as follows: 20% had completed primary and vocational school, 23.3% had completed high school, 26.6% had begun but not completed third-level education, and 20% had completed third-level education. Informed consent was obtained prior to involvement in the study.

The Borderline Personality Inventory (BPI [30]) is a true–false self-report instrument based on Kernberg’s structural theory of personality organization, and which is recommended for large-scale screening of BPO. The most discriminatory items are included in the shorter Cut-20 scale, which provides a cut-off score: Cut-20 results $\geq 10$ are recommended as the cut-off point for the proper classification of borderline patients. In our research, we used Cut-20 $\geq 10$ for assignment to the clinical group and Cut-20 $\leq 5$ for assignment to the control group. The BPI has good internal consistency, test–retest reliability, and satisfactory rates of sensitivity (0.85 to 0.89) and specificity (0.78 to 0.90) [30].

Strange Stories [31] is a method of measuring advanced ToM, consisting of a set of stories that refers to a character’s behavior in everyday situations. The participant’s task is to explain the character’s behavior. Successful performance requires attribution of mental states, sometimes including higher-order mental states, such as one character’s belief about the beliefs of another character. Eight social stories taken from the original 24 were used (stories referring to physical behavior were excluded). The accuracy of each response was rated on a 0–2 scale by two competent judges [32], and agreement between judges was excellent (ICC=0.9).

The Mental States Task (MST [23]), evaluates individual differences relative to two processes: representation/elaboration and openness/modulation to one’s subjective experience. First, participants were primed with the 3BM card of the Thematic Apperception Test [33], in order to evoke emotional arousal and regulation strategies. They were then asked to write down a story that came to mind in response to the image. Next, participants responded to 24 items assessing their mental states during the previous task. The MST measures the following six mental states, which reflect the interactions between the activation of mental representations and their modulation: Concrete thinking: an important defect in the elaboration of the subjective experience, low awareness of one’s mental contents; Low Defensive Level: the activation of representational contents makes the subject emotionally overwhelmed, and the mental contents are defended against through immature defenses (e.g., splitting and acting out); Intermediate Defensive Level: the recognition and elaboration of the representational contents are impeded, and the person’s subjective experience is obliterated or its meaning is downplayed by defenses of denial, minimization, or emotional suppression; Objective-Rational: the subjective experience is treated with an objectifying attitude and distance; High Defensive Level: elaboration on and openness to the true subjective experience is present, but is defended against by more mature defenses and adaptive emotional regulation strategies; Reflective Thinking: the capacity to...
recognize and elaborate the full subjective experience, associated with some use of mature defenses and emotional regulation strategies. The score for each subscale reflects the scores for each mental state, and the total MST score is obtained by using an equation, which includes weights to reproduce the reflective continuum. The MST has good reliability coefficients (0.79–0.58 for the English version and 0.82–0.62 for the French version) [23].

Results

The correlations between variables are presented in Table 1.

Table 1. Intercorrelations between variables

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>BPI</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cut 20</td>
<td>0.96”</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>TU</td>
<td>-0.11</td>
<td>-0.09</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>CT</td>
<td>-0.15</td>
<td>-0.22”</td>
<td>-0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>LDL</td>
<td>0.54”</td>
<td>0.55”</td>
<td>-0.001</td>
<td>-0.38”</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>IDL</td>
<td>-0.18</td>
<td>-0.26”</td>
<td>-0.01</td>
<td>0.29”</td>
<td>-0.38”</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>O-R</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.13</td>
<td>0.15</td>
<td>-0.26”</td>
<td>-0.22”</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>HDL</td>
<td>-0.10</td>
<td>-0.12</td>
<td>-0.03</td>
<td>0.14</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Refl</td>
<td>-0.08</td>
<td>-0.06</td>
<td>-0.16</td>
<td>-0.19</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.29”</td>
<td>0.24”</td>
</tr>
<tr>
<td>10</td>
<td>MST</td>
<td>-0.22”</td>
<td>-0.17</td>
<td>-0.08</td>
<td>-0.39”</td>
<td>-0.19</td>
<td>-0.28”</td>
<td>0.28”</td>
<td>0.57”</td>
</tr>
</tbody>
</table>

CT-Concrete Thinking, LDL-Low Defensive Level, IDL-Intermediate Defensive Level, O-R-Objective-Rational, HDL-High Defensive Level; Refl.-Reflective Thinking

The results of the comparisons of the mentalization levels (Table 2) indicate that there are differences between the clinical group and the control group with regard to the two categories of mental states Low Defensive Level and Intermediate Defensive Level. In case of Low Defensive Level, BPO participants obtained significantly higher scores than the control group, whereas the opposite pattern of results was observed for the other scale, where the participants above BPO obtained scores higher than the clinical group. In other categories of mental states, there was no difference between groups. Comparison of the results relating to the ToM revealed no differences between the groups, either in relation to any single story or to general result.

Table 2. Comparison of groups in terms of mentalization and ToM

<table>
<thead>
<tr>
<th></th>
<th>CT</th>
<th>LDL</th>
<th>IDL</th>
<th>O-R</th>
<th>HDL</th>
<th>Refl</th>
<th>MST</th>
<th>ToM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above-BPO group M (SD)</td>
<td>14.2 (5.50)</td>
<td>10.57 (4.28)</td>
<td>10.03 (3.75)</td>
<td>16.97 (4.25)</td>
<td>15.9 (5.70)</td>
<td>16.47 (5.48)</td>
<td>1.96 (.15)</td>
<td>12.33 (1.84)</td>
</tr>
</tbody>
</table>

*table continued on the next page*
Results discussion

The present research project is the first to link in a single study an exploration of the cognitive and affective aspects of “understanding the behavior of others in terms of internal states,” while pointing to their significance for the level of personality organization. Firstly, it was confirmed here that there is a relationship between the level of personality organization and mentalization in the absence of a relationship between ToM and LPO. Secondly, the hypothesis of independence between ToM and mentalization was confirmed.

The test procedure relating to mentalization assumed the activation of internal relation-emotional representations, usually accompanied by a tendency to defend against painful and overwhelming affect [23]. An interesting pattern of results was obtained when comparing the two treatment groups: the clinical group showed the prevalence of the Low Defensive Level mental state (indicating the activation of excessively overwhelming mental representations with very primitive defenses, i.e., splitting and acting out), while the control group revealed more Intermediate Defensive Level (indicating impeded recognition of activated mental content with attempts to obliterate its personal meaning, with defenses such as denial, minimization, and suppression of emotions being used for this purpose) [23]. These patterns represent empirical confirmation of the assumptions on the differentiation of personality organization: 1) primitive and immature representations, intolerance of negative affect, greater impulsivity and defenses based on splitting and externalization of affect in BPO; 2) attempts to weaken the affect by mechanisms of emotional suppression in those above BPO [2, 33, 34].

At the same time, no differences were revealed between the two groups in the tasks related to theory of mind in any of the types of stories. Individuals in both the control and treatment groups explained the reason for the behavior of the character in similar ways in the case of stories examining both the basic ability to understand the behavior of others in terms of mental states, as well as more advanced abilities like “third-order theory of mind.” This suggests a lack of relationship between theory of mind and the level of personality organization, between understanding the intentions of others and intrapsychic structures of varying emotional pathology.

Our results extend the understanding of psychopathology of LPO, and especially the issues of disclosure of symptoms, differences in activated mental states, defenses,
and emotion regulation strategies. The research results reveal that people with lower and higher PO equally understand others as having the internal mental states and use this knowledge to interpret their behavior. However, a major limitation to the use of this knowledge in a comparable way is the activation of the relational–emotional internal representations [2,25,26]. This activation involves the expression of different mental states. The recognition and elaboration of the representational contents are disturbed or impeded, and the mental contents are defended against through a number of defensive patterns [23]. While both populations are similar in the same way they use “cold” knowledge about the mental states of others, when it comes to emotions in relationships and “hot” knowledge, the groups vary in their range of mental states, defenses, and emotional regulation strategies.

These results also shed light on the issue of the interaction between the cognitive and emotional aspects of knowledge of the mental states of others, and the question of the origin of their normal and impaired development. Arguments are given in the literature for both the interdependence and independence of these competences [35]. Moreover, distinct factors seem to play a key role in the origin of ToM and of mentalization. The factor responsible for the achievement of ToM is considered to be the development of language, including the social use of language, family characteristics, and executive functioning [1]. On the other hand, emotional and relational context—i.e., the quality of attachment [5, 7]—is considered most significant in shaping the role of mentalization. The results of our study speak for the independence of theory of mind and mentalization, and suggest trajectories of development of these abilities that differ, at least in part.

The results of this study should be viewed as preliminary and considered in the light of the limitations of the research. One limitation is the small sample size and the large intragroup differences in terms of education and age. The impact of the latter variables has been minimized through the use of the matching pairs procedure. Careful generalization of the results should thus be in line with the replication of this study in larger and more homogeneous populations.

Conclusions

1. There is not only clinical, but also empirical, justification for the importance of differentiating between the concepts of mentalization and theory of mind, understood as relating to the emotional–regulatory versus cognitive–perceptual aspects of “understanding the behavior of others in terms of internal states.”
2. Theory of mind does not help to differentiate between levels of personality organization; “cold” knowledge of the internal states of others plays an equivalent role in individuals with high and low levels of emotional and structural pathology.
3. Mentalization differentiates between levels of personality organization; “hot” knowledge of the internal states of others and activation of internal representations in a relational context leads to different mental states and is associated with different levels of pathology.
4. The results suggest a significant contribution of mentalization, though not of ToM, to the pathogenesis of levels of personality organization.
5. These data argue for the independence of the two abilities and their (at least partially) different trajectories of development.

References


