Neurobiological Models of Self-Disorders in Early Schizophrenia

A. Mishara, I. Bonoldi, P. Allen, G. Rutigliano, J. Perez, P. Fusar-Poli, and P. McGuire

Self-disorders (SDs) are disturbances of the first-person perspective, long associated with schizophrenia, particularly in early phases. Although psychopathological features of SDs continue to be studied, their neurobiological underpinnings are unknown. This makes it difficult to integrate SDs into contemporary models of psychosis. The present review aims to address this issue, starting from an historical excursion revealing an interconnection between neuroscientific models and the origin of the psychopathological concept of SDs. Subsequently, the more recent neurobiological models related to SDs are discussed, particularly with respect to the onset of schizophrenia.

**Key words:** schizophrenia/psychosis/self-disorders/neuroimaging/prodromal/Heidelberg mescaline study/aberrant salience/Bayesian modeling/Gruhle/Mayer-Gross/phenomenology

**Introduction**

Self-disorders (SDs) are disturbances of “basic-self” (a “pre-reflective” sense of first-person perspective). They are characteristic features of the schizophrenic spectrum disorders, often preceding the onset of full-blown psychosis. SDs manifest as a variety of anomalous subjective experiences including depersonalization, diminished sense of existing as bodily subject, distortions of first-person perspective, reduced sense of coherence in fundamental features of self (e.g., sense of anonymity, identity confusion, etc.), and disturbed self-other/self-world boundaries. Recent evidence suggests that SDs are core phenomena of schizophrenia. In subjects experiencing first-episode psychosis, SDs are associated with schizophrenia as opposed to affective psychoses.

Greater symptom severity, neuropsychological impairment, poor treatment compliance, and clinical outcome. In people at ultra high risk for psychosis (UHR), the SDs are associated with increased risk for transition to psychosis. Despite the clinical importance of SDs, their neurobiological underpinnings are unknown, limiting their integration in contemporary models of psychosis, which have a major neurobiological component. This review aims to address this issue. We offer an historical perspective revealing that SDs and neuroscientific models of schizophrenia have been interrelated since their original conceptualization. We then examine the evidence linking SDs with contemporary models of the early phase of schizophrenia.

**Historical Definition of SDs**

The origin of the SDs term is sometimes attributed to Jaspers. Although Jaspers (1913) describes phenomena related to SDs, eg, passivity experiences, being “influenced,” inserted/withdrawn thoughts, etc., he did not systemize them under the SDs concept. The SDs concept was rather developed by Jaspers’ Heidelberg colleagues (Gruhle, Mayer-Gross, Beringer). For the first time, as far as we know, we report: (1) the role of Jaspers’ Heidelberg colleagues (the early Heidelberg School) in the origin and development of the SDs (see Hermle et al., however, in their acknowledgement of Beringer’s contribution) and (2) the fact that Gruhle (in 1915) coined the term SDs: “Although I found the experience frequently described, it was never captured by its own term...Tentatively, I call this passivity - the nonparticipation in one’s own experience - a self-disorder...” (our translation).
SDs as Aberrant Salience: The Heidelberg Psychotomimetic Mescaline Study

The Heidelberg Psychotomimetic Mescaline Study

The development of the SDs concept was associated with a pivotal study by the Heidelberg School in the 1920s, concerning the psychotomimetic effects of mescaline in modeling SDs. Beringer injected mescaline hydrochloride intramuscularly to study the phenomenological “structure” of healthy individuals’ subjective experience of the psychotomimetic SDs, thus paving the way for hypotheses about neurocognitive mechanisms. The results indicated that SDs involve disruption of embodied perceptual experience, which affects the experience of time, space, and continuity of self-experience (see table 1).

Mescaline, SDs, and Automatic Processing

These studies supported Gruhle’s observation that perceptions, movements, and hallucinations are experienced in SDs as having independence from self (ich-unabhängig). Perceiving, moving, speaking, thinking, and willing are normally supported by automatic processes. With SDs, the patient experiences these automatic processes as independent “automatisms” having foreign agency. This is consistent with the hypotheses that SDs have common neurobiological mechanisms and are disturbances of the relationship of self to its own cognitive-affective processing.

Mescaline, SDs, and Thought Disturbances

Mayer-Gross, a participant in the mescaline studies, observed that hallucinations and the “made” thoughts, etc., have a common component: it is not that thoughts are ascribed to alien agency, they are “perceived” as alien. The suggestion is that for subjects with SDs, thinking is experienced as sensory, as in thoughts that are audible to the subject (Gedankenlautwerden) and auditory verbal hallucinations (thought to be derived from inner speech). Mayer-Gross described “a making sensual (Versinnlichung)” in the “sensory representation of thoughts... Without this change, the manifestations remain inexplicable.” By observing subtle self-perceived cognitive and other disturbances in prodromal schizophrenia and in mescaline intoxication, Mayer-Gross “anticipated” Huber’s basic symptoms concept. Basic symptoms are subtle subclinical disturbances, which are experienced as arising from the self. As such, they do not overlap with SDs, which, are perceived, by definition, as happening to self, without the self’s participation. Nevertheless, mescaline models both the earlier self-perceived subclinical difficulties (Denkerschwerung) and, at higher doses, their transition to thoughts becoming sensory in the SDs leading to experiences of thought insertion/withdrawal/broadcasting, etc.

SDs as Primary Symptoms

In the 20s and early 30s, Gruhle and Mayer-Gross further developed the SDs concept by proposing that SDs were among the primary symptoms of early schizophrenia (in Jaspers’ sense). As the other primary symptoms, SDs are “nonunderstandable in terms of their historical-cultural context and the person’s biography (or motivations) because the underlying neurobiological process has interrupted the person’s development... Something new (unprecedented, nonderivable) must be present.”

It was only later that Kurt Schneider, who in 1946 became Director of the Heidelberg Psychiatric Clinic, further systemized the phenomenology of SDs in schizophrenia, which he included among its first-rank symptoms.

SDs and Aberrant Salience in Early Psychosis

Gruhle and Mayer-Gross maintained that SDs involve the interruption of the understandable context or inner connectedness of experience. This is in line with Hemsley’s hypothesis that cognitive disturbances in schizophrenia involve a change in the way stored material is integrated with sensory input. A failure to relate sensory input to a contextually appropriate frame of reference would lead to faulty or absent expectation and thus to inappropriate allocation of attention to details of the environment not normally reaching awareness (what today is called aberrant salience). As Jaspers’ primary delusions, the SDs are “nonunderstandable” for the interviewing clinician, but—unlike primary delusions—they are also nonunderstandable for the patient. Inserted thoughts concern harmless, mundane circumstances. Yet, the patient knows precisely which thoughts are inserted, eg, the thought to go to a concert is suddenly experienced as foreign.

Similarly to the aberrant salience thought to underlie prodromal delusional mood, SDs disrupt the online contribution from past experience in shaping the inner continuity of consciousness. Previous learning is experienced as irrelevant thus disrupting context. Seeming to come from “nowhere,” the made-experiences, withdrawn thoughts, or other SDs symptoms disrupt “the inner connectedness of current concerns,” (ie, goal processing). The patient himself/herself is surprised. This process suggests an analogy with what in more recent accounts is called a prediction error signal in Bayesian modeling. Due to the ongoing “interruption” by the made or influenced perceptions, movements, thoughts, etc., there is a reduction of what the patient expects or anticipates (das Vorschauende) from moment to moment; there is only the compelling sensory evidence of now (aberrant salience): “no temporal order prevails, each sensory impression is
A. Mishara et al

Table 1. Heidelberg Psychotomimetic Model of SDs by Administering Mescaline to Healthy Individuals

<table>
<thead>
<tr>
<th>Phenomenological Domains (1–9) Derived From Experimental Subjects’ Self-Reports</th>
<th>Source (Reference Number)</th>
</tr>
</thead>
</table>
| 1. Experiences resembling the later concept of aberrant salience  
The experimental subject falls into a foreign, never before experienced state…[resembling]  
psychosis.  
If I opened my eyes after closing them, the world had something surprising, new about it.  
The colors are so forceful; all the objects seem newly painted.  
A wagon rolling in the distance makes a thunderous sound…even slight sounds are experienced as painfully intense  
Each stimulus whether tactile, optic or acoustic, is equally strong, and each equally important. | 13 |
| 2. The past does not contribute to shaping present experience and goals  
Everything which I see is different, isolated, without any relationship to what has happened in the past  
I experience increasing difficulty to implement my impulses into goals and movement. | 9 |
| 3. The experience is new and compelling  
There is a loss of relationship to past and future.  
During the experiment, I happened to receive a letter of considerable importance and opportunity. I read it with complete indifference without feeling or reaction. The whole thing appeared to me to be meaningless, as if it belonged to some past time. | 13 |
| 4. Self/other boundary confusion  
Only the object in front of me exists; I forget myself and everything else around me.  
If I see a soup bowl in front of me, there exists only this soup bowl. | 13 |
| 5. Cognitions/emotions occur independently from the self’s volition  
Individual ideas, brief chains of thoughts emerged randomly, without my being the least bit conscious of where they came from. They appeared to have no connection with me. I just looked on amazed.  
What one does automatically, speaking, thinking, moving, now seems foreign  
Perception, thought, movements and will are no longer experienced as my own.  
My mouth moves without my doing anything… | 9 |
| 6. Foreign agents have odd appearance and/or power over self  
Faces appear sculpted, sharply contoured, overly expressive due to exaggerated shading like an actor in a play or caricature  
Perception, thought, movements and volition are no longer experienced as my own, and with that, the feeling that thinking is not initiated by me, but someone else, a foreign agency.  
It is as if a second person thinks and speaks, while I am unable to produce a complete thought. | 12 |
| 7. Disruption of sensorimotor integration and cognition  
Movements are experienced as abnormally slow or not seen at all: the movement of my hand is only experienced at its beginning and end positions.  
Conversely, movements are exaggerated, seen as abnormally fast, resting objects are seen as moving  
My movements appear to me as artificial, foreign, like an automaton  
At first thoughts are accomplished with remarkable ease, seeming to make the point surprisingly well followed by the increasing difficulty to think or volitionally attend at all | 9 |
| 8. Distortions of time and space  
I could not envision future or past. I lived entirely in the present, and even that in an entirely thin slice  
Time slows down, coming to a complete standstill with a sense of timelessness, or conversely, starts to speed up  
When time stands still, there is loss of sense of movement and space | 12 |
| 9. Anomalous body experiences (aberrant salience with regard to own body)  
For a moment, I randomly pulled on my earlobe and scratched at it. Then, it seemed that my entire bodily feeling was only my earlobe, which being enormous, I held in my hand. The entire rest of my body had disappeared…  
When considering my hands it is always the one I am looking at that is bigger… | 13 |

Note: SD, self-disorder.

equally valued, replacing its predecessor.”16 The experience is new and compelling and the patient’s autobiographical past seems irrelevant.14,15,19 Moreover, the prediction error disrupts the event encoding of autobiographical memories into coherent self-experience (reviewed in Mishara and Fusar-Poli19). In summary, the early Heidelberg account is compatible with later Bayesian modeling: SDs occur across different modalities (perceptual, motor, sensorimotor, volition, memories, thoughts), suggesting that a common mechanism (prediction error in current terminology) applies to each modality.
Neurobiological/Neurocognitive Models of Dysfunction Implicated in SDs

In considering the neurobiology of self-disorders we are indebted to the prior reviews and proposals by Nelson et al.23,24

Self-Monitoring Models

A failure to recognize stimuli as self-generated might result in the inability to ignore irrelevant stimuli, causing the misattribution of self-generated processes to other agents.25 Different brain areas have been implicated as playing a key role in self-monitoring.

Self-Referential Processing. Self-referential processing regards stimuli that are experienced as strongly related to one own’s person.26 There is evidence that cortical midline structures (CMS), including the medial prefrontal cortex, the anterior and posterior cingulate cortices, and the precuneus, are important for self-referential processing26 and increasingly activated during tasks requiring judgments about the self-relevance of stimuli, independently from the nature of the stimulus or sensory modality. Together with inferior parietal and lateral temporal areas, the CMS comprise part of the “default-mode network,” deactivated during engagement in nonself-referential tasks (eg, focal-cognitive/executive tasks) and are active during resting or baseline conditions.26 Since this network plays a so-called pre-reflective role during tasks requiring active reflection of self,26 it is thought to mirror the pre-reflective basic-self as a unified, stable perspective of the subject in relation to the environment, altered in SDs. Furthermore, CMS engagement across different stimuli modalities is in line with the hypothesis14–17 that SDs have common neurobiological mechanisms across the modalities of perceiving, acting, thinking, and willing. The CMS (together with insula27) are involved in assigning first-person perspective28 and contribute to form the sense of basic-self as “subject and agent of perception,” as opposed to self as “object of attribution.”29 Moreover, CMS are involved in self-representation in the past, future, and present,26 suggesting their role in maintaining the temporal continuity14–17 and uniqueness of basic-self. Dysfunctions in CMS may result in SDs, hyperreflexivity, increases in self-focus, self-other boundary confusion, and passivity phenomena.

Sensorimotor Integration (Forward-Model Deficit). Other authors propose that CMS and default-mode network are not specific to pre-reflective sense of self, but rather associated with general cognitive functions involved in inferential processing and memory retrieval.30 Specifically, they argue that the self-specific first-person perspective altered in SDs is anchored in sensorimotor integration. According to Frith and Done,25 and commentaries by Gallagher31 and Heinz,32 an effference copy of the action is matched with feedback from the action actually made, creating sense of ownership. A copy of the intention to act is sent as feed forward information to a comparator: a match with the occurrence of the action generates sense of agency. Forward-models predict and dampen the perception of planned actions, allowing discrimination of self- from non–self-generated actions, thus contributing to sense of agency. A recent functional magnetic resonance imaging study confirmed that patients with schizophrenia seem unable to predict and suppress the sensory consequences of their actions.33 Hauser and colleagues demonstrated that altered sense of agency linked to abnormal sensorimotor predictions is already present in UHR subjects.34 Other authors propose that the sensorimotor integration model could be extended to encompass other domains including mental actions (emotions and thoughts).20 In early schizophrenia, delusions of control may derive from misattribution of self-generated movements, emotions, thoughts as externally generated, resulting from a dysfunctional forward-model mechanism,25 involving disconnectivity between motor and sensory cortices. Somatosensory-related areas (postcentral gyrus, insula, temporoparietal junction) and anterior cingulate and prefrontal cortex are crucial for the sense of being an embodied subject,26 via their interactions with key motivational and limbic areas such as the ventral striatum and amygdala. In early schizophrenia, impaired prediction (in sensorimotor, emotional, perceptual, thought domains) would undermine sense of ownership and agency for both mental content and actions. For example, corollary discharge issuing from frontal areas where thoughts are generated would fail to alert auditory cortex that the thoughts are self-generated, leading to SDs involving sense of agency and ownership, the misattribution of inner speech to external sources, hearing one’s thoughts spoken aloud, and auditory hallucinations.16

Salience Models

Bayes’ theorem proposes that our expectations determine how we interpret new evidence. Conversely, subjective beliefs are updated, accounting for the new evidence. Violation of expectations (prediction errors) makes an event attention grabbing, ie, more salient. If prediction error does not fit the knowledge based on previous experience, a new inference occurs.22 The following models suggest that the dysfunction in the capacity to compare predicted and incoming stimuli and, thus, to adequately interpret the experience may be related to dysfunctional salience processing and therefore contribute to SDs.

Novelty and Motivational Salience. Subcortical dopamine release from ventral tegmental area occurs when a novel stimulus is experienced, signaling salience, orienting attention, and motivating behavior.37 The novelty signal is thought to arise in the hippocampus,38 which
may act as “comparator,” vital to the inner continuity of consciousness in terms of prediction from past experience. Aside from novelty processing, the attribution of salience to environmental stimuli is important to motivate behavior in relation to reward. For example, incentive salience is a psychological process that modifies the perception of stimuli, imbuing them with salience, making them attractive. Neurobiological alterations of salience attribution may lead to faulty, or reduced expectations, resulting in difficult integration between actual situation and one’s prediction. When mismatch occurs, innocuous environmental stimuli motivate attentional orientation due to dysregulated striatal dopamine and prediction error signaling. Attention is allocated to stimuli, which otherwise go unnoticed, leading in turn to an excessive awareness of stimuli. This disruption of the tacit structure of normal experience may account for hyperreflexivity, proposed as belonging to the SDs, introducing awkward rigidity, slowness, and sense of perplexity in the person’s interactions with world.

This model may account for the ability to automatically grasp the significance of an event, action, or

Table 2. Hypotheses Concerning Possible Neurobiological Mechanisms of SDs

<table>
<thead>
<tr>
<th>Phenomenological Domains of SDs (Early Heidelberg School)</th>
<th>Neurocognitive/Neurobiological Model</th>
<th>Brain Area/Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movements and mental processes are experienced as having independence from self&lt;sup&gt;11–13,16,17&lt;/sup&gt; (see table 1: 1.5, 1.7)</td>
<td><strong>Dysfunctional self-monitoring mechanisms:</strong></td>
<td>-Somatosensory-related cortices (postcentral gyrus, insula)&lt;sup&gt;35,45&lt;/sup&gt;</td>
</tr>
<tr>
<td>Although intensely experienced, there is increased passivity and nonparticipation in one’s own experience&lt;sup&gt;11–13,14–17&lt;/sup&gt; (see table 1: 1.4, 1.5, 1.7)</td>
<td></td>
<td>-Hippocampus, parahippocampal cortex, and cingulate cortex&lt;sup&gt;25&lt;/sup&gt;</td>
</tr>
<tr>
<td>Thoughts, feelings, actions, volition come under foreign power which paralyzes the I&lt;sup&gt;11–13,14,15&lt;/sup&gt; (see table 1: 1.6)</td>
<td></td>
<td>-Bilateral temporoparietal junction and neighboring areas</td>
</tr>
<tr>
<td>In the thought disturbances of the SDs (eg, thought insertion, thought withdrawal, thought broadcasting), thinking is experienced as sensory. The collapse between thinking and perceiving is a direct “perception.”</td>
<td></td>
<td>-CMS, insula&lt;sup&gt;26–28&lt;/sup&gt;</td>
</tr>
<tr>
<td>Possible shared mechanisms of SDs across modalities&lt;sup&gt;16,17&lt;/sup&gt; (see table 1: 1.5, 1.6)</td>
<td></td>
<td>-Dorsal ACC and lateral PFC&lt;sup&gt;39&lt;/sup&gt;</td>
</tr>
<tr>
<td>Self-other boundary confusion&lt;sup&gt;11–13&lt;/sup&gt; (see table 1: 1.1 1.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sense of being embodied subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“No temporal order prevails, each sensory impression is equally valued, replacing its predecessor”&lt;sup&gt;11–13,16,17,19&lt;/sup&gt; (see table 1: 1.1, 1.3, 1.8)</td>
<td><strong>Disrupted salience:</strong></td>
<td>-Dysregulation of the dopamine system, due to the chaotic firing of midbrain neurons and increased striatal dopamine release, leads to dysregulation of the hippocampal-VTA loop&lt;sup&gt;38,40,42&lt;/sup&gt;</td>
</tr>
<tr>
<td>Loss of “common sense” disrupts the on-line contribution from past experience (context) in shaping the inner continuity of consciousness, with attendant disturbances in time perception&lt;sup&gt;14–17&lt;/sup&gt; (see table 1: 1.2, 1.3, 1.8)</td>
<td></td>
<td>-Corticostriatal network comprising the midbrain, basal ganglia, lateral medial temporal, and prefrontal cortex&lt;sup&gt;41,42&lt;/sup&gt;</td>
</tr>
<tr>
<td>Weakening of the “grounding” effect of context:</td>
<td></td>
<td>-Hippocampus, amygdala, and medial prefrontal cortex&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>Alteration of the inner continuity of consciousness in terms of prediction from past experience</td>
<td></td>
<td>-Increased striatal dopamine release, leading to dysregulation of the hippocampal-VTA loop&lt;sup&gt;38,40,42&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Note: ACC, anterior cingulate cortex; CMS, cortical midline structures; PFC, prefrontal cortex; SD, self-disorder; VTA, ventral tegmental area.
sentence given its social context. In fact, evidence from animal and human studies suggests that the neural circuit involved in the learning and memory processes that enable context-dependent behavior includes the hippocampus, amygdala, medial prefrontal cortex, and ventral striatum.\textsuperscript{40-42} The implicit “grip” of the “rules of the game” or “common sense”\textsuperscript{22,43,44} includes a sense of the situationally appropriate. Disruption of one’s “grasp” of the field of awareness may occur with hyperreflexivity and diminished self-experience, thought to be fundamental to SDs.\textsuperscript{2}

Aberrant Salience. The aberrant salience model of psychosis mostly implicates a corticostriatial network comprising the midbrain, basal ganglia, lateral medial-temporal, and prefrontal cortex.\textsuperscript{37} Disruption of dopamine’s role in learning relevant associations and updating inferences and beliefs about the world could result in delusion formation/positive symptoms.\textsuperscript{22} Salience attribution has been studied in patients with early psychosis at both behavioral and neural levels using reward-learning tasks. The inability to distinguish between relevant and irrelevant stimuli in early schizophrenia may result in an alteration of how objects and meanings emerge from background context. The ability to recollect the past, interpret the present, and anticipate the future is disrupted. Consequences of that may be the SDs, a loss of “common sense,” and hyperreflexivity (table 2).

Conclusions and Implications for Early Schizophrenia

A model integrating phenomenological, neurocognitive, and neurobiological aspects of SDs is likely to improve our understanding of the mechanisms that underlie early schizophrenia. Conceptually, the neurochemical and neurofunctional alterations observed in UHR and first-episode subjects could be better associated with their subjective inner world and feelings. On the diagnostic/prognostic side, it may support risk stratification and individualized focused interventions in early psychosis. Among anomalous experiences in UHR, SDs have high specificity for the schizophrenic spectrum.\textsuperscript{7} Unlike attenuated positive symptoms, which have no prognostic significance, SDs (particularly passivity phenomena\textsuperscript{46}) seem to have convincing predictive power\textsuperscript{7} of conversion to schizophrenia. If results are replicated, SDs assessment could be used by ongoing international early psychosis projects (eg, PRONIA, PSYSCAN) that are developing translational diagnostic and prognostic tools by integrating psychopathology and neuroscience modalities. Finally, a comprehensive model combining neuroscience and SDs could create an experimental platform for the development and assessment of novel treatments targeting SDs in early psychosis.

Funding

P.F.P. was supported in part by a 2014 NARSAD Young Investigator Award.

Acknowledgment

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References