Inefficient Neural Self-Stabilization: A theory of spontaneous resolutions and recurrent relapses in psychosis

Lena Palaniyappan MD PhD FRCPC

Department of Psychiatry & Robarts Research Institute, University of Western Ontario & Lawson Health Research Institute, London, ON. Canada

Address correspondence to:
Lena Palaniyappan, Room 3208, Robarts Research Institute, 100 Perth Drive, London, ON N6A 5K8, Canada. Email: lpalaniy@uwo.ca
Telephone: (519) 931-5777 ext. 24398

Word counts: Abstract: 145 Text: 5828

Funding: L Palaniyappan’s work is supported by the Canadian Institute of Health Research (Foundation Grant 375104), Bucke Family Fund and the Academic Medical Organization of South Western Ontario.

Conflicts of interest: LP reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Otsuka Canada; travel support from Boehringer Ingelheim and Magstim Limited outside the submitted work. In the last 3 years, LP and/or his spouse have held shares in Shire Pharmaceuticals and Glaxo Smith Kline in their pension funds for values less than US$10,000.
ABSTRACT

A striking feature of psychosis is the heterogeneity of the phenomenon. Presentations of psychosis vary from transient symptoms with no functional consequence in the general population to a tenacious illness on the other extreme, with a wide range of variable trajectories in between. Even among patients with schizophrenia, who are diagnosed on the basis of persistent deterioration, marked variation is seen in response to treatment, frequency of relapses and degree of eventual recovery. Existing theoretical accounts of psychosis almost exclusively focus on how symptoms are initially formed, with much less emphasis on explaining their variable course. In this review, I present an account that links several existing notions on the biology of psychosis with the variant clinical trajectories. My aim is to incorporate perspectives of systems neuroscience in a staging framework to explain the individual variations in illness course that follows the onset of psychosis.
Introduction

A striking feature of psychosis is the heterogeneity of its presentation. Psychosis-like experiences occur commonly in the general population, often without notable functional consequences. In general healthcare settings, psychotic symptoms that significantly impact on one’s daily function occur in various disorders (e.g. delirium), often with full resolution. In psychiatric clinics, the course of psychosis varies from being a single, time-limited episode on one end of the spectrum to a tenacious illness on the other extreme, with a wide range of variable trajectories. Even among patients with schizophrenia, who are diagnosed on the basis of persistent deterioration, marked variation is present in response to treatment, frequency of relapses and degree of eventual recovery. Despite this, existing theoretical accounts of psychosis almost exclusively focus on how symptoms are initially formed, with much less emphasis on explaining their variable course. Thus, to date the focus of understanding the neurobiology of psychosis is largely at a "symptom formation" level, rather than at the "illness course" at an individual level. As a result, we continue to lack a physiological framework that can explain the wide range of variable outcomes that unravel to a patient experiencing an initial episode of psychosis. In this article I make an attempt to construct an account that could link several existing notions on the biology of psychosis with the variant clinical trajectories. My aim is to put forward a thesis that could be invoked during the clinical dialogue with a concerned carer wondering why his/her loved one is presenting so differently from another patient attending the same treatment program.

I first briefly review the concept of symptom resolution in psychosis and the evidence linking psychosis in general, and schizophrenia in particular, to cellular and systems level brain connectivity. Existing ideas about the role of brain development, degeneration and plasticity will be invoked to show how the concept of brain network-level homeostasis can account for the varied course of psychosis. I also argue that resolution of psychotic symptoms requires inherent homeostatic processes, which when aberrant, inhibit a fuller recovery. Finally, I highlight the aspects of psychotic illnesses that are not fully addressed by this framework, and suggest future studies that are required to test the implications of the notions proposed here. I use the term ‘neural self-stabilization’ throughout this paper for the sake of simplicity, but this refers to the homeostatic process affecting all cellular constituents (i.e. glial cells and vasculature) that enable information transfer involving the entire brain.

Onset and resolution of psychosis:

Psychotic disorders are clinically defined and diagnosed by the presence of delusions and hallucinations alongside deficits in processing speed, attention, verbal fluency, emotional
expression, logicality and coherence of thought. These symptoms have a high probability of co-occurrence in some patients and demonstrate a variable degree of resolution after the first presentation.

There are several features of the natural course of psychosis that call for explanation. Isolated psychotic experiences such as voice-hearing and delusion-like ideas occur regularly among the otherwise healthy individuals. Though these experiences are mostly transitory, the epidemiological risk factors for such experiences overlap substantially with the risk factors of psychotic disorders with conventionally poorer outcomes such as schizophrenia. Subjects with psychosis-like experiences or the more tightly defined constructs of at-risk states or schizotypal disorder are at a higher risk of developing full-blown psychosis; but the majority of those who experience such transient psychotic states do not develop a psychotic episode; the risk of conversion peaks at 2 years and drops with longer follow-up periods.

The onset of psychosis is often insidious or sub-acute in schizophrenia (53% to 70%), but can be florid and acute in many other psychotic disorders. The insidious prodromal stage in schizophrenia frequently presents with anxiety related to often numerous random coincidental associations and primitive perceptual aberrations, but at the peak of a fully evolved episode, psychosis is characterised by a limited number of stereotyped and fully formed delusions and hallucinations that tend to repeat. As psychosis evolves, patients often appear to add further elaborations to this limited set of delusions and hallucinations, rather than forming completely unrelated ideas (see Table 1 for First Person Accounts).

<table>
<thead>
<tr>
<th>Author</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chadwick, 2007</td>
<td>As my delusional system expanded and elaborated, it was as if I was not “thinking the delusion,” the delusion was “thinking me!” I was totally enslaved by the belief system. Almost anything at all happening around me seemed at least “relevant” and became, as Piaget would say, “assimilated” to it. Another way of putting things was that confirmation bias was massively amplified, everything confirmed and fitted the delusion, nothing discredited it. Indeed, the very capacity to notice and think of refutatory data and ideas was completely gone.</td>
</tr>
<tr>
<td>Chapman, 2002</td>
<td>I often misinterpreted real life occurrences such as the behaviors of others as somehow related to those conspiring against me. When people passed by (police cruisers, door-to-door salespeople), I thought they must be there to spy on me. When I half heard a conversation in the distance or the honking of a car, I would think it held special significance for me. I would randomly open a dictionary and find a word (“die,” “liar,” “evil”) and interpret how the word had special meaning for me.</td>
</tr>
</tbody>
</table>
Powell, 1998  After the first hospitalization, it was almost predictable that every 4 years my mother's behavior appeared to change in the fall. … Although my mother tried desperately to recuperate from each psychotic episode and each arrest she became increasingly reclusive and paranoid. Each episode was precipitated by an erotomanic delusion or delusions of persecution in the workplace that followed shortly after the psychiatrist decided to taper her antipsychotic medications.

**Table 1: First person accounts of delusional activity**

A substantial number of patients with first episode psychosis have only one episode \(^{20–22}\). Globally, the incidence rates of such acute and transient psychotic disorder are consistently higher than the incidence of schizophrenia \(^{23,24}\). Psychosis can resolve without treatment in some of these patients \(^{21,25}\). Longitudinal studies conducted before the advent of neuroleptics remind us that these numbers are large enough to not be dismissed \(^{26}\). When appropriate treatment is started, a large number of first-episode subjects show an early symptomatic response \(^{27,28}\); such early reduction in the severity of delusions and hallucinations predicts a favorable later outcome \(^{29–31}\). Resolution of delusions does not usually involve ‘extinction’ or ‘unlearning’ the associations underlying psychotic beliefs; instead, it involves a gradual ability to detach from the pressure of the beliefs and perceptual abnormalities \(^{15,32}\). While positive symptoms become less prominent with the course of illness, cognitive deficits and negative symptoms remain stable throughout the course of the illness \(^{33–35}\).

Relapses can occur even in adequately medicated patients (primary relapse) \(^{36,37}\); but the rates of relapse are distinctly higher in those who discontinue treatment in the early stages of illness (interventional relapse) \(^{38,39}\). When relapses occur during the course of illness, irrespective of the duration of intervening period of recovery, the same predictable set of symptoms tend to recur with each episode \(^{40,41}\). Furthermore, unlike the first episode of psychosis which is preceded by a long duration of prodromal symptoms and gradual buildup of unusual experiences, relapses often occur without a similar insidious prodrome \(^{36}\). With each relapse, treatment resistance becomes more likely, especially when the relapse occurs after antipsychotic discontinuation \(^{36,39}\). Notwithstanding this phenomenon, a small proportion of patients achieve good function, despite a high number of early relapses \(^{42–44}\).

**Progressive structural changes of the brain in psychosis:**

One of the most consistent neuroimaging observation in patients with psychosis is the reduction in the amount of grey matter volume and thickness measured using MRI \(^{45,46}\). These GM deficits are present even in the early stages of a patient’s life \(^{47,48}\), and to some
extent are shared by healthy siblings as well \(^{49-51}\). Once the early psychotic symptoms come to surface, these GM deficits appear to intensify \(^{52,53}\), especially in the first few years \(^{54}\) before slowing down \(^{55}\). Some reports indicate a continuous but low-level of ongoing GM reduction even in much later stages \(^{56,57}\). The extent of the GM deficits relates to both the severity of illness \(^{58-60}\) and the degree of exposure to agents that are associated with relapses and functional disability \(^{55,61,62}\). In addition, rather controversially, longitudinal GM reduction is more pronounced in those who have higher cumulative exposure to antipsychotic medications (with some differences between typical and atypical drugs) \(^{52,63,64}\).

Nevertheless, at several brain regions, these changes appear to be reversible. Certain types of non-drug therapies appear to reverse or slow-down the structural changes among patients \(^{65-68}\).

While not every patient with psychosis show progressive structural deficits, in those that show such GM changes, there seems to be a predilection for certain brain regions including superior temporal \(^{53,60,69-75}\), lateral frontal \(^{71,72,75,76}\), insular \(^{52,60,70,73,77}\), and anterior cingulate \(^{70,71}\) cortex on a more consistent basis, followed by thalamus \(^{75}\), precuneus \(^{76}\), inferior parietal \(^{75}\), and the hippocampal \(^{74}\) regions somewhat less consistently. Interestingly, in the large-scale organization of correlated brain activity and structural connectivity, these regions constitute the so-called ‘hubs’ \(^{78}\) or ‘rich clubs’ \(^{79}\) within the human brain. Hubs and rich-club regions typically form the most connected nodes within the overall network architecture of the human brain and consequently show higher levels of overall activity \(^{80,81}\). The connectional architecture and activity load experienced by these regions make them particularly vulnerable for structural damage in diffuse brain disorders such as schizophrenia \(^{81,82}\).

The widespread structural changes in psychotic disorders are often discussed in the context of either extended aberration in neurodevelopment and/or a limited form of neurodegeneration. While earlier theorists used the term degeneration to refer to the longitudinal changes in psychosis, it is now increasingly clear that neither a progressive neuronal loss nor a relentless clinical deterioration characteristic of true neurodegeneration occurs in psychotic disorders, leading to a preference for the term neuroprogression to denote post-onset brain changes \(^{83}\). There is now an overwhelming number of observations that are interpreted as signs suggestive of aberrant neurodevelopment in schizophrenia \(^{84-88}\). An equally strong line of argument exists for a limited form of neuroprogression \(^{89-92}\). Several attempts to bridge the two notions have been made in the recent times \(^{93-97}\). Most of the proposed compromises hinge on the notion that a healthy adult brain continues to develop and change in its structure over time \(^{98,99}\); hence a developmental aberration would continue
to affect brain structure in adult life, thus explaining the neuroprogression in schizophrenia 100.

Psychosis as a disorder of connectivity:

The notion that psychosis is related to aberrant connectivity in the brain originated in the 19th century 101–104. The initial concept of disconnection was based on the manifested disconnection in thoughts, actions and behaviour seen in patients with psychosis. Various elegant theories later resurrected, refined and pinned this idea to the ‘brain’ level, with the aid of postmortem and neuroimaging studies in the last 20 years 105–109. The three key pillars of the dysconnectivity hypothesis are (1) the reduction in neuropil, the tissue zone that normally houses a large number of neuronal synapses, observed in postmortem brains of patients with schizophrenia 107,110,111 (2) abnormal increases and decreases in the correlation of activity among various brain regions (functional connectivity) measured using Positron Emission Tomography, Magnetic Resonance Imaging, Electroencephalography and Magnetoencephalography studies 108,112. (3) abnormal increases and decreases in the indices of white matter integrity (structural connectivity) measured using Diffusion Imaging 113,114. While these observations arise from different measurement tools employed at different spatial and temporal scales and activity levels, they are notably reconcilable at a whole brain (systems) level 115,116, and speak to a reduction in the ability of information transfer within the affected brains.

While there are notable spatial variations in the patterns of resting-state functional connectivity in relation to psychosis, with improvised data processing approaches 117,118, some patterns are now emerging consistently. Functional hyperconnectivity, especially affecting the prefrontal cortex, is more pronounced during early stages of schizophrenia 119,120, relates to positive rather than the negative symptoms of the illness 119,121,122, and normalizes to some extent with antipsychotic treatment 123,119. Such functional hyperconnectivity also results from external agents that typically induce psychotic symptoms 124,125. Ketamine, an agent that produces nearly all of the core symptoms of schizophrenia in healthy humans, produces robust hyperconnectivity involving the prefrontal cortex 125–130. In particular, this resting-state hyperconnectivity involves a set of brain regions that constitute the ‘default-mode network’ 122,131,132. These regions characteristically show an elevated level of activity at rest, and appear relative deactivated when an individual is engaged in task performance 133. In contrast to the prefrontal/DMN hyperconnectivity in the early stages, a wider hypoconnected resting-state is often noted in later stages of schizophrenia 112,120,126.
Certain emerging observations provide clues as to the neural process that may underlie the hyperconnectivity between 2 brain regions seen in resting-state functional MRI. Firstly, hyperconnectivity is often seen during the initial response to neuronal injury. Secondly, training and new learning in healthy brain results in an early increase in functional connectivity in relevant brain regions, along with hypoconnectivity. Even in the absence of a learning exercise, coordinated electrical/magnetic neural stimulation (plasticity-inducing paradigms) results in functional hyperconnectivity indicating a Hebbian increase in neural communication at a synaptic level.

The concept of neural self-stabilization:

Homeostasis is a ubiquitous regulatory process that serves to maintain the function of a system at a set activity level, thus providing stability. The human brain is constantly bombarded with events and objects in the environment that evoke neuronal activity; in addition, constant spontaneous activity is also a feature of neuronal existence. Given the thousands of synapses that each neuron has with many other neurons, random coincidental firing between two (or more) neurons is highly likely within this milieu.

Hebbian rules of plasticity dictate the weakening of synaptic connectivity between the two neurons. If left unchecked, such a system can end in a hyperconnected or hypoconnected mode, neither of which is optimal for new learning or information transfer. The maintenance of both sparse functional connectivity and a steady baseline activity with low energy consumption are important for the status quo of human brain.

Several modes of homeostatic compensation operate to reduce the resulting instability; some of these involve functional rebalancing by stabilizing the firing rate of a neuron (intrinsic plasticity), tuning the inhibitory inputs (inhibitory plasticity) or down-weighting synapses (scaling). In addition, structural changes such as changes in the synaptic size, synaptic number and the dendritic spine structure (structural plasticity) either co-occur or serve as a second level of homeostatic mechanism. Reduced connectivity at synaptic level is compensated by an increase in the size and number of synapses, while reduced neuronal activity (e.g. via input deprivation) results in reduced synaptic elimination. These normal physiological processes of brain that regulate neuronal excitability or synaptic strength continuously degrade the absolute effect of synaptic coding that occurs with associative learning, though preserving the essential memory traces.
Together, the homeostatic processes serve to maintain the overall excitatory and inhibitory balance within local neuronal ensembles that constitute the global brain connectome\textsuperscript{159}. This enables a connectome-wide \textit{self-stabilization} that facilitates optimal signal processing and learning\textsuperscript{160}. Such readiness is a prerequisite for continuous adaptation to one’s environment\textsuperscript{161}. This notion is summarized in (Figure 1).

Fig 1. Efficient neural network self-stabilization or topological homeostasis facilitates optimal signal processing and learning.

Psychosis and inefficient neural self-stabilization:

In the following section, I provide an account of how a disruption in neural self-stabilization can result in the varied presentations of psychotic experiences. Firstly, I propose that several factors can disrupt the stable pattern of timing-dependent coincidence detection at a neural level.

1. Intrinsic hyperactivity: A neural tissue with anomalously high frequency of activity can result in an increased probability of associative plasticity. Depending on the location of this activity, at an experiential level, subjects may experience brief sensory or cognitive disruptions of fleeting nature. Several studies of hallucinating individuals have reported an elevated level of neural activity involving sensory cortices\textsuperscript{162,163}

2. Disturbances in the normal constraints on associative plasticity: A cardinal feature of Hebbian plasticity is the dependence on the temporal order of coincidental neuronal activation\textsuperscript{164}; under certain circumstances, this temporal window can be prolonged, increasing the probability of formation of coincidences. For example, dopamine and substances inducing an excess release of dopamine could potentiate this mechanism
At an experiential level, subjects may report heightened sensitivity and clarity for events.

3. Failure of habituation: Repeated presentation of the same stimulus elicits progressively smaller neuronal response. Disruptions in this habituation could prolong the state of evoked neuronal activity, increasing the probability of coincidences. A large body of electrophysiological studies point towards a habituation-deficit in psychosis. At an experiential level, subjects may report a lack of feeling of familiarity for events, again increasing ambiguity and uncertainty.

Figure 2: Anomalous Associations and Psychotic Experiences: A. Learning associations between two time-variable signals require tight temporal coordination (Hebbian window), shown as a narrow interval between the activation of pre- and post-synaptic neurons in the first illustration. This window can be prolonged in hyperdopaminergic states, as shown in the second illustration B. Anomalous bursts of presynaptic activity can lead to inadvertent Hebbian associations. C. Failure of habituation may lead to prolonged states of evoked activity, increasing the probability of Hebbian associations.

While these aberrations clearly have the potential to generate psychotic experiences (Figure 2), in each case, in the presence of an intact homeostatic plasticity, the synaptic coding will be weakened and eliminated. Thus, these disruptions on their own can produce periodic psychosis-like experiences and prodromal features but are insufficient to produce a psychotic episode. If agents that induce the above neural states are repetitive in frequency or massive in their dose, then a temporary overload of homeostatic mechanisms can ensue, leading to a psychotic episode, nevertheless with a high probability of full resolution with or without treatment (Figure 3). A likely example is the clinical presentation of drug-induced psychosis with full resolution. If so, then what is the necessary condition for the emergence of a schizophrenia-like illness?
Repetitive or massive doses of inducing agents

A temporary overload of homeostasis

'run-away excitation'

A psychotic episode

Homeostatic catch-up

Full resolution.

e.g. drug-induced psychosis

Figure 3. **Stabilization Lag in a Psychotic Episode**: Psychotic episodes can occur after repetitive or massive doses of inducing agents through mechanisms shown in figure 2, leading to a temporary overload of neural self-stabilization. Provided the cellular/topological self-stabilisation apparatus (homeostatic) is intact, these episodes can resolve fully.

I propose that in certain individuals, aberrant (non-structural) homeostatic plasticity leads to a lack of resolution of coincidental associations. In such cases, synaptic strengthening accumulates (often insidiously) and the neuronal ensemble harbouring such coincidental associations will eventually experience a runaway excitation, forming a strong feed-forward loop that would quickly incorporate other correlations to the initial association (delusional elaboration characteristic of a psychotic episode), resulting in a state of indiscriminate hyperconnectivity in the proximate network space of the affected ensemble. In such circumstances, a recurrent, stereotyped temporal pattern is set on motion that can persist for an abnormally long period of time ('burnt into memory' 147). This hyperconnectivity enables instantaneous recruitment of neuronal modules which are not normally involved in the processing of a given stimulus or task ('spreading activation'), thus producing apparent hyperactivity in distant sites often reported in patients with psychosis. Topologically, the connectome now appears to be ‘subtly randomized’ as sparsity in the connectivity is lost (as described by Rubinov 171). In the absence of homeostatic resetting, synaptic strengths of the affected neuronal core get fully saturated, precluding formation of new associations that are required for extinction. Thus, once fully formed in an individual with inherent defects in homeostatic plasticity, extinction of delusions becomes highly unlikely. Furthermore, the lack of neural self-stabilization also reduces the connectome’s general readiness to process further inputs, thus reducing the speed of signal processing, expressed clinically as psychomotor poverty.
With the sustained and diffuse destabilization of the connectome that affects its sparse connectivity, the need for alternate homeostatic processes involving structural plasticity get triggered \(^{154,172,173}\). Structural plasticity involves synaptic elimination and retraction of spines that aims to restore the sparse connectivity state that existed prior to the onset of psychosis. This process gradually eliminates excitatory dendritic synapses, with a consequent neuropil reduction \(^{107,174}\), and progressive grey matter loss. Due to the inherently higher activity levels and their propensity to be highly accessible to most of the aberrant neuronal ensembles, rich-club hubs of the human connectome (anterior cingulate cortex, insula, lateral prefrontal cortex, superior temporal gyrus and hippocampal regions) are most likely to be affected by this global retuning process \(^{78,81,82,175}\). With time, this leads to ‘de-escalation’ of hubs and restoration of sparse connectivity, albeit at the cost of increased segregation of functional modules and prolonged transit time within the network (Figure 5).

Aberrant functional plasticity

\[
\text{Insidious accumulation of synaptic strength} \\
\text{‘Run-away excitation’ but no resolution} \\
\text{hyper-connectivity (resting fMRI)} \\
\text{Inefficient cerebral recruitment} \\
\text{Information processing deficits}
\]

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Stabilization Shift in Psychotic Disorders: In those who have a predilection for aberrant functional plasticity, synaptic gains from associations accumulate over time, leading to run-away excitation within the neural network. The occurrence of this event may be brought forward by exposure to inducing agents as indicated in figure 3. In the absence of an intact neural self-stabilization process, this results in a hyperconnected state for resting-state brain networks with inefficient over-recruitment of task-processing regions. Subtle information processing deficits that accompanied the predilection for aberrant functional plasticity now becomes more pronounced, with the step change coinciding with the first psychotic episode. The neural self-stabilization mechanism now shifts from inefficient functional plasticity to a robust dependence on structural plasticity (i.e. spine reduction)

This structural homoeostatic compensation eventually succeeds in abolishing the recurrent spreading of the avalanche of activity originating within anomalous neuronal ensembles. This comes at a cost of somewhat lower efficiency of information transfer, which becomes
\end{figure}
apparent whenever the demands on the system increase. Patients at this stage of illness exhibit a lower level of psychotic symptoms, but continue to express negative symptoms (reduced speed of information processing, avoidance of social demands). One of the unintended consequences of hub/rich-club damage is the emergence of peripheral hubs, often located in unimodal cortices. These peripheral hubs, despite now having a higher degree of functional connectivity than other brain regions, lack the richness of hierarchical structure possessed by conventional core hubs. As a result, the flexibility of resource allocation is reduced at a global level resulting in a reduction of system-wide plasticity, producing inefficient information transfer especially when demands are placed on the networks, thus forming a fertile soil for further relapses. Psychosis-inducing triggers continue to produce further episodes at this stage, though with reduced homeostatic reserve – both intrinsic and structural – even milder doses of the inciting agents may now be sufficient to induce relapses \(^{176,177}\). Furthermore, compensation through dendritic spine reduction reaches a critical point after repeated relapses with no further room for spine reduction, leading to a state of ‘homeostatic occlusion’ (Figure 6). In those individuals who reach this stage, treating psychotic episodes becomes more difficult, taking longer time, requiring higher doses, and despite this some episodes will remain treatment-resistant.

Figure 5. Hub De-escalation in Psychotic Disorders: A. Rich-club hubs of the human connectome (anterior cingulate cortex, insula, lateral prefrontal cortex, superior temporal gyrus and hippocampal regions) have inherently high activity levels and higher topological proximity to any given brain region. B. As a result, the pathways to and from these nodes are most likely to be the sites of dendritic spine reduction occurring in response to anomalous hyperconnectivity. C. With time, this leads to ‘de-escalation’ of hubs (red-dotted circles), increased demands on remaining hubs (over-loading effect, shown with a red halo) and emergence of peripheral hubs (yellow nodes). While this helps in restoration of sparse connectivity, it comes at the cost of increased segregation of functional modules (nodes with a thunderbolt sign less well connected to other hubs) and prolonged transit time within the network.
Many other agents that may worsen the outcome of established illness can directly induce dendritic spine elimination \(^{178}\). Drugs of abuse also alter the dynamics and microstructure of both dendrites and dendritic spines \(^{179}\); chronic stress also promotes synaptic elimination \(^{180}\). In addition, activity dependence of structural synaptic plasticity means that at least some of synaptic elimination could be secondary to lack of environmental stimulation \(^{181}\). In addition, when the first psychotic episode occurs at a very early age, accompanying developmental processes synergistically hasten dendritic spine reduction resulting in an accelerated homeostatic occlusion and early emergence of treatment resistance.

Figure 6. **Saturation of Self-Stabilization**: Hub de-escalation results in a suboptimally retuned system that is prone to runaway excitation even with milder doses of psychosis inducing triggers. Recurrent relapses exhaust the structural plasticity through dendritic spine elimination, leading to a state of homeostatic occlusion. This is associated with treatment-resistant state, whereby interventions that primarily act by enhancing functional plasticity are no longer effective.

Dopamine regulates the gain of NMDA receptor mediated Hebbian associative plasticity \(^{182}\). Antipsychotics, by blocking D2 receptors, may act directly at the synaptic milieu to inhibit the associative learning process. They also facilitate the intrinsic functional homeostatic plasticity that counteracts the Hebbian potentiation, thus reducing the time taken for symptom resolution \(^{183,184}\). In particular, typical neuroleptics result in inhibition of long-term potentiation and spike-timing dependent plasticity \(^{183}\). This effect will rapidly reduce the runaway facilitation occurring at the synaptic level during a psychotic episode, thus facilitating resolution of psychosis (early responders) and restoring the efficient sparse-connectivity \(^{185}\). Nevertheless, in patients with defective intrinsic homeostasis such rapid
response may not occur; in addition, when response occurs eventually, discontinuation or dose reduction may carry a higher risk of relapse. Further, antipsychotics also increase oxidative stress that may trigger or indirectly facilitate structural synaptic plasticity\textsuperscript{186}. With long term use, certain antipsychotics (especially the atypicals) may have a propensity to reduce synaptic elimination, thus halting or reducing the neuropil reduction that may occur during the course of schizophrenia\textsuperscript{187} while the typical antipsychotics may alter synaptic proteins, thus facilitating dendritic spine regression and neuropil reduction\textsuperscript{188}. Critchlow et al.\textsuperscript{189} demonstrated a 59\% increase in primary dendritic spine density in rat hippocampal neurons upon clozapine administration, while haloperidol had an opposing effect (also see\textsuperscript{190}). This suggests that clozapine may have specific effects on dendritic spines that may help restore structural plasticity even in later stages of treatment resistance.

Neutral self-stabilization, gender and developmental age:

Brain development is characterised critical time windows where the self-stability of the brain is notably affected. Such windows are characterised by an increase in structural homeostatic activity that aids to restore the stability. For example, adolescence is associated with crucial changes in the profile of NMDA receptors which mediate functional synaptic plasticity (a subunit switch from NR2B to NR2A)\textsuperscript{191,192} and the dopamine receptor D2 that affects inhibitory plasticity\textsuperscript{193}. During this critical period, excitatory synapses are actively eliminated\textsuperscript{193–195}, thus increasing the inhibitory tone required for balanced brain activity. I posit that the emergence of psychosis is more common during late adolescence as excess demands are placed on structural plasticity at this time. Furthermore, homeostatic plasticity is the major mechanism of experience-dependent shaping of the developing brain. Any perturbation in this system, as proposed here as a mechanism for psychosis, is bound to have developmental impact, that can present itself even before the onset of psychosis (e.g. disrupted asymmetry\textsuperscript{196}, folding defects\textsuperscript{197}). In particular, it is possible that invocation of structural homeostatic plasticity is related to an increased vulnerability of oxidative stress, which plays a crucial role in synaptic elimination\textsuperscript{198,199}. This may explain why obstetric complications such as neonatal hypoxia and consequent developmental defects are more likely to be seen in patients who later develop schizophrenia\textsuperscript{200,201}.

Hebbian plasticity mechanisms appear to be modulated by sex-specific neurosteroids in animals\textsuperscript{202,203}. In particular, estradiol, allopregnanolone and related endogenous neurosteroids serve to reduce intrinsic neuronal excitability through various mechanisms\textsuperscript{204}. Provided such gender-specific mechanisms are preserved in individuals who are later develop schizophrenia, the occurrence of runaway excitation necessary for persistent
psychosis and subsequent attempts contributing to grey matter loss must be less likely, taking longer time to be established in females \(^{205}\).

Compatibility with emerging observations:

Multiple lines of evidence support the various postulates that form the building blocks of the inefficient neural self-stabilization. A large body of evidence emerging from transcranial magnetic stimulation studies support the aberrations in plasticity mechanisms in schizophrenia \(^{206-210}\). In addition, various genetic animal models that mimic various well validated aspects of the schizophrenia phenotype converge on aberrant functional plasticity \(^{211}\).

Graph-theory based accounts using data from multiple modalities of neuroimaging point towards a subtle randomisation of the topology of brain networks in schizophrenia despite the preservation of the sparse but efficient small-world structure of connectional architecture \(^{101,171,212,213}\); this effect is anticipated if the core hubs lose their prominence and peripheral hubs emerge in an attempt to restore the sparse connectivity \(^{214}\). Several other genetic loci implicated in psychotic disorders also point towards mechanisms of homeostatic synaptic plasticity \(^{215}\).

Inefficient neural self-stabilization characterised by a tilt from functional homeostatic plasticity to structural plasticity is compatible with many emerging observations that link inflammatory processes with psychosis. Genome-wide association studies now point towards a dysregulation of inflammatory mediators in schizophrenia \(^{216,217}\). Within the MHC complex, the variations in the complement cascade appears to be the most prominent genetic abnormality associated with schizophrenia \(^{217}\). Complement cascade is dubbed the ‘masterful homeostatic regulator’ of synaptic plasticity \(^{218}\). Various products of this cascade are involved in tagging the appropriate synapses to be eliminated during normal development \(^{219}\). Further, excess microglial activity which is suspected to contribute to grey matter reduction in schizophrenia \(^{220}\), can lead to homeostatic synaptic elimination \(^{221}\).

GABA and Glutamate system:

Comprehensive review of post-mortem studies suggests that the major contributor to GM volumetric loss in schizophrenia is synaptic regression involving glutamatergic excitatory synapses \(^{110,222}\). On the other hand, functional deficits related to reduced levels of GAD67 enzyme, rather than actual loss of parvalbumin containing interneurons, are now well-established in patients with schizophrenia \(^{223}\). Deficiency in the GAD67 system does not appear to be a direct result of neurotransmitter dysfunction \(^{224}\) but may indicate the failure of a key functional homeostatic process that responds to excitatory synaptic activity \(^{145,225-227}\).
this weakening of inhibitory plasticity and subsequent connectome instability may serve to tilt the balance in favour of excitatory synaptic elimination in schizophrenia. This is consistent with our proposed model of inefficient neural self-stabilization.

Glutamatergic concentration measured using MR spectroscopy appears to be elevated during early phase of illness but reduces during the course of illness, and this relates to grey matter loss in distal sites. The hallmarks of excitotoxic cell damage due to excess glutamate, if present, is yet to be demonstrated in schizophrenia, though post-mortem studies point to reduced glutamatergic dendritic spines, and reduced biological co-ordination between glutamatergic signalling and synaptic structure. The current model of inefficient self-stabilization predicts that in early stages, higher levels of glutamate signal can be found in the loci of hyperconnectivity and anomalously increased neural activity whereas lower levels of glutamate signal should accompany grey matter loss in schizophrenia, especially in later stages of illness, with excitatory synapses undergoing structural elimination. While several studies are in agreement, opposite results also exist. Cautious interpretation is required nevertheless as it is unclear how much of spectroscopic glutamate refers to synaptic activity.

Compatibility with current theories:

The current hypothesis draws from the existing accounts of associative learning and so is broadly consistent with the models based on deficiencies in reinforcement learning and prediction error. Invoking disrupted plasticity as an explanation for psychotic disorders itself is not new. Stephan et al. posit that NMDA hypofunction is the key aspect of the plasticity defect seen in schizophrenia. One of the central tenets of this postulate is the ketamine model of schizophrenia. Ketamine mimics many acute symptoms of a psychotic episode; but do not reproduce the features seen in established schizophrenia. Goto et al. highlighted the plasticity mechanisms of the prefrontal cortex and how they could be disrupted in schizophrenia. Keshavan et al. assembled a large body of evidence to argue that the positive symptoms of schizophrenia can be explained by hyperplasticity, and negative symptoms by hypoplasticity, with either one being a compensatory response to the other. More recently, Forsyth and Lewis posited that the consequences of impaired synaptic plasticity during development explains the emergence of clinical symptoms of schizophrenia. Self-stabilization processes operating at the synaptic and macro-connectomic level are not invoked in the models proposed by Goto et al., Keshavan et al. and Forsyth and Lewis.
A recent proposition highlights the role of homeostatic compensations in the putative excitation-inhibition imbalance that starts with primary glutamatergic abnormality in schizophrenia \(^{(250)}\). In contrast, the current model takes an agnostic view on the primacy of neurochemical abnormality. I propose that various disruptions in associative learning can invoke the earliest symptoms of psychosis; but the core dysfunction in established schizophrenia lies in the homeostatic aspect of plasticity that is likely to primarily involve NMDA, but can also occur as a result of disruptions in other endogenous systems such as the endocannabinoid pathways, for example. By placing the emphasis on the homeostatic aspect of plasticity, the current hypothesis attempts to explain the continuum and the course of schizophrenia rather than the acute phase of appearance of symptoms (Figure 7). In addition, I provide a framework that incorporates the emerging literature on connectomics, plasticity deficits, neuroprogression, inflammation and addresses distinct aspects of the population-level variation in the course and outcome of primary psychotic disorders.

Grades of psychosis

Figure 7. Neural Self-Stabilization and Grades of Psychosis: The degree of impairment in the homeostatic process of topological neural self-stabilization determines whether an episode of psychosis resolves fully, relapses repeatedly and fails to respond to currently prescribed interventions.

Predictions, limitations and further questions:

Physiological disruptions in neural processes related to learning based on temporal contingency are likely to be non-specific to schizophrenia and will be shared not only by other psychotic disorders but also in healthy individuals with psychosis-like experiences. But patients with schizophrenia are more likely to have specific disruptions in the homeostatic mechanisms that favour structural over functional plasticity. This vulnerability, if present from early life, could increase the risk of structural neurodevelopmental aberrations even in the absence of psychosis (e.g. in siblings). But after the onset of symptoms, more extensive structural changes would be limited to those with inefficient self-stabilization. As a result,
morphometric changes may not be directly related to the genetic liability for disease expression. But as the tissue loss is related to neural plasticity, reversal of these structural deficits is possible. Identification of individuals at risk of more severe forms of psychosis, and offering treatments that reduce or delay progressive synaptic changes at early stages may have a true disease-modifying effect. There are some promising venues in this regard.

Established post-onset grey matter loss is the sign of a 'steady-state', albeit an imperfect one, that serves to induce further relapses. Reversal of deficits may require some disruption to this suboptimal steady-state to enable introduction of alternate means neural self-stability. To make therapeutic progress in established cases, it will be important to know how to safely carry out such topological restoration. Emerging neuromodulation approaches provide some promising leads in this regard.

Given the heterogeneity of the currently accepted construct of schizophrenia, it is unlikely that any single theory could fit the multitude of observations pertaining to this disease. In particular, the notion of inefficient self-stabilization does not offer any causal explanations; it merely offers a mechanistic framework where the operations of several causal agents could converge. In line with many other prevailing hypotheses, the current framework offers a better explanation for the course of positive symptoms; there is a greater degree of uncertainty surrounding the nature of relapses and resolution of negative symptom and thought disorder in schizophrenia. Finally, the proposed theory does not provide a single index that can capture the inefficient self-stabilization that can occur in all patients. Instead, it raises the suggestion that aberrations in various indicators representing the extant pathways of intrinsic plasticity and structural plasticity can be brought together to characterize a substantial number of patients with poor outcome.
### Postulates

<table>
<thead>
<tr>
<th>Postulates</th>
<th>Approaches to test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Shift from functional to structural plasticity induces dendritic spine reduction</strong></td>
<td>Animal models of disrupted homeostatic plasticity, when subjected to repeated associative learning, would mimic the structural phenotype of schizophrenia</td>
</tr>
<tr>
<td><strong>2. Focal hyperconnectivity predates diffuse hypoconnectivity in psychotic disorders</strong></td>
<td>Longitudinal imaging from unmedicated FEP to post-psychotic, treatment responsive stages of illness</td>
</tr>
<tr>
<td><strong>3. Hub de-escalation after first episode of psychosis</strong></td>
<td>Patients with schizophrenia-like phenotype will show progressive reduction of hubness of core nodes, emergence of peripheral hubs as well as grey-matter reduction in hub regions</td>
</tr>
<tr>
<td><strong>4. Lag of neural self-stabilization in single psychotic episode</strong></td>
<td>Patients with single psychotic episodes with full recovery will lack post-psychotic morphological and connectomic changes</td>
</tr>
<tr>
<td><strong>5. Occlusion of topological homeostasis in treatment resistant subjects</strong></td>
<td>Demonstrating a lack of learning-induced changes in network connectivity in patients with treatment resistance</td>
</tr>
</tbody>
</table>

**Table 2: Experimental approaches to test neural self-stabilization theory of schizophrenia**

Aberrations in homeostatic plasticity have been suspected in many neuropsychiatric disorders\(^{147,215,247,255,256}\). Future studies aimed at discovering the aspects of this disruption that is specific to schizophrenia are required to fully understand the molecular pathways that can be targeted for treatment. If experimental models of disrupted homeostatic plasticity fail to mimic the true phenotype of schizophrenia despite the presence of conditions that disrupt associative plasticity, the theory proposed here could be refuted. To my knowledge, there is a striking lack of studies on patients who recovered after a single psychotic episode as well as longitudinal studies on connectome topology. Such studies are essential to test the premises of the neural self-stabilization theory (See Table 2).

**Conclusion:**

I speculate that psychosis, as seen in clinical practice, is not a disorder of any single neurotransmitter system or a particular brain network; instead it is a disorder of cerebral acclimatization to new learning. An important aspect of this theory is that there is an *intrinsic antipsychotic defense* mechanism that promotes self-stabilization in the human brain; this
mechanism is built upon the balancing act between neural activity and connectivity that is essential for learning statistical regularities in our immediate environment. Restoration of this intrinsic self-stabilization can reverse many features of long-term psychotic illnesses. Critical timing of interventions, in combination with approaches that provoke and redirect the pathways of neural network stability may be required to realize this goal.

Acknowledgement: I gratefully acknowledge Prof. Tim Crow (Oxford) for reviewing this hypothesis and commenting on several aspects of this manuscript. I also acknowledge Dr. Ridha Joober and the attendees of CCNP Meeting at Kingston, June 2017, for stimulating questions that led to revisions of this work.

References:


211. Balu DT, Coyle JT. Neuroplasticity signaling pathways linked to the pathophysiology of schizophrenia. *Neurosci Biobehav Rev* 2011;35:848–70.


