Taxonomizing the Neural Correlates of Delusion

Taxonomizing the neural correlates of delusions

Rebecca Hammond

21 April 2019

Mentored by David S. Silbersweig, MD, PhD, Department of Psychiatry, Brigham & Womens Hospital

Taxonomizing the neural correlates of delusions

Question: What are the neural correlates of delusions?

Student role: Perform a review of the literature and write a perspectives piece

Abstract

Title: Taxonomizing the neural correlates of delusions

Purpose: Delusions represent a core symptom of schizophrenia and are a major cause of morbidity associated with psychiatric illness, yet the neural correlates of delusions remain relatively under-characterized. Part of the challenge is that delusions are difficult to define, and it is not known whether different types of delusions share underlying etiologies, or where to draw the lines of taxonomy in the first place. The purpose of this paper is to review current evidence for the neurobiological mechanism of delusions, with special focus on persecutory delusions and delusions of reference, which represent two of the more commonly accepted subtypes. A critical discussion of the classification of delusions, the conceptualization of their mechanism, and the impact of these two features on the study of delusions will also be performed.

Methods: A review of studies using fMRI to examine the neural correlates of persecutory delusions and delusions of reference was performed. This review was expanded to include papers outlining foundational concepts upon which the assumptions of the fMRI studies were based. Several studies were selected for detailed analysis for the purposes of both providing evidence for a mechanism of delusion and as examples of the challenges of studying delusions.

Results: The aberrant salience hypothesis is the prevailing theory for the neural basis for delusions, and the interpretation of fMRI data are informed by this framework and also support it. Persecutory delusions and delusions of reference both appear to share abnormalities in the so-called salience network, and each exhibit abnormalities in specific other networks of the brain that make sense given our understanding of their phenomenology. Though these mechanisms create a sensical picture of each delusion type, much of the data are not perfectly accounted for by prevailing theory.

Conclusions: Much remains to be elucidated regarding the neural substrate of delusions. Functional MRI provides a useful though imperfect modality for their study. Given the difficult-to-define nature of delusions, careful consideration must be given to the implications of our assumptions on study design and data interpretation in order for the field to progress.

TABLE OF CONTENTS

INTRODUCTION

Delusions represent a core symptom of schizophrenia and related psychotic disorders, and a major cause of morbidity and mortality associated with psychiatric illness. Classically defined as fixed false beliefs held in spite of overt evidence to the contrary, contemporary definitions of delusion have abandoned the requirement of falsehood in recognition of the difficulty in establishing what constitutes “truth”, and included instead a requirement that delusional beliefs not be endorsed by members of patients’ cultures (Cummings & Mega 2003). While hallucinations in schizophrenia have been extensively investigated, the neural bases of delusions remain relatively uncharacterized. The difficulty of defining delusions, and that they may exist on a spectrum with normal beliefs, complicates their study.

As they represent such a significant cause of morbidity, delusions are something psychiatrists would like to treat. Antipsychotic medications and reality checking techniques and are the current options. The medications have transformed the treatment of psychosis, but remain of mixed efficacy, with high patient-to-patient variability and undesirable side effect profiles. The history of antipsychotic drug development began with the development of chlorpromazine while searching for a more sedating version of promethazine for use in anesthesia; subsequent antipsychotics, including the atypical neuroleptics, have built off of this or other similar unintended drug-effect discoveries. We have then worked backwards from the drug’s effects to, for example, deduce that the psychosis of schizophrenia has something to do with dopamine in the brain. Ideally, psychiatry would move toward a more targeted approach to treatment development, as has, for example, the field of immunology. There is little hope of getting to that point, however, unless the processes that underlie psychiatric pathology are better elucidated. So it is for delusions.

A systems-level neuroscience lens is a useful perspective from which to study delusions. Beliefs, not to mention delusions, are inherently complex. They might be seen as a composite of sensory experiences, memories, and other internal signals, all of which combine to create thoughts, which in turn might connect to other thoughts, emotions, and memories of previous experiences, that must then undergo a process of gaining special importance and established, in our neural systems, as what we colloquially think of as a “belief”; i.e., a mode of thinking from which we derive our experience of how the world works. When a delusion is formed, therefore, something has gone awry, likely on multiple levels, such that this interpretation of how the world works or relates to the individual is not in keeping with the culture that surrounds the person. That person’s entire lens of reality has shifted from the (relative) norm. The implications are that any modality we use to begin to study the etiology of delusions should be capable of casting a broad enough net to capture something of these processes.

TECHNOLOGICAL METHODOLOGIES USED TO STUDY DELUSIONS: FMRI

Functional magnetic resonance imaging (fMRI) provides such a systems level view of the brain. With fMRI, changes in blood oxygenation and perfusion are detected to noninvasively map whole brain activity with high spatial resolution. While the ways in which these measures and neural activity relate are not fully understood, fMRI remains the non-invasive technique of

choice for measuring brain activity similarities and differences between different individuals and/or in response to a variety stimuli.

Blood oxygenation level dependent (BOLD) contrast and arterial spin labeling (ASL) are the two primary contrast mechanisms used in fMRI. BOLD contrast relies on the magnetic properties of the iron molecule in deoxygenated hemoglobin (Chow 2017). Signal reflects decreases in regional deoxyhemoglobin, meaning it tends to arise from venous structures. This decrease in regional deoxyhemoglobin (which is paramagnetic) has been attributed to increases in cerebral blood flow (CBF) that exceed metabolic demand, though the mechanism has not been fully elucidated. Of note is that BOLD signal changes are not quantifiable in absolute physiological units and are therefore expressed as a percentage signal change or a statistical significance level based on a statistical model. The practical implications for this are that absolute or resting activity are not easily assessed with BOLD fMRI, meaning that it may be difficult to distinguish absolute baseline from effects in BOLD fMRI studies. ASL measures CBF more directly by using a radiofrequency inversion pulse to magnetically label the water molecules in arterial blood before it enters an area of interest, then measuring magnetic changes induced by the influx of the labeled water molecules. The advantages of ASL over BOLD are that absolute CBF measurements (in units of ml 100 g(-1) min(-1)) are possible, and since arterial flow rather than venous flow is distinguished, the result is a more direct measure of CBF. The main disadvantage of ASL fMRI is a lower temporal resolution relative to BOLD fMRI (one image per 4 seconds, compared to an image every 1-2 seconds (accounting for a 4-6 second time lag), or even as frequently as every 100ms using BOLD). Because of the better temporal resolution and because it has been around for longer (and there is thus increased familiarity and capacity for the technique), BOLD remains the method most commonly used in contemporary studies.

Functional MRI is by no means perfect for measuring the neural substrate of delusions. When analyzing the implications of fMRI data it is important to remember that one starts by taking multiple images of a three dimensional object (the brain) in snippets of time, while a subject is performing a certain task, or presented with a certain stimulus. The technique has a spatial resolution of 1-2 mm, which is impressive, but does not provide micro neural circuit data that can be highly significant in neural processes. The four dimensional raw data are then fed into highly complex image processing and statistical modeling systems that attempt to build a profile of neural activity.

There are several statistical modeling methods used across different studies. All of these models seek to find the areas of the brain that have statistically significant variation in levels of blood flow over time during the relevant task. These models must be designed to analyze data so as to eliminate confounding factors such as small movements of the head, which might otherwise render the data unreadable. The models themselves are converging, yet differences need to be taken into account when interpreting results.

The older, and still most commonly used model is univariate analysis. Univariate analysis parses data by examining individual voxels in the brain within different slices of time, searching for areas that have statistically higher signal (using a predefined statistical model), with many adjustments made to account for innate variability and noise. Given the size of the data and the amount of processing required to search through every voxel in order to derive the bases for the statistical model used to analyze the rest of the voxels, researchers often decide in advance of the data analysis which areas of the brain are relevant, and build the model through which the

rest of the data are analyzed based on these predetermined regions. This addresses multiple comparisons, but heightens the potential for confirmation bias.

Other methods use multivariate approaches such as principal component analysis (PCA) to facilitate the construction of this statistical model from whole brain data (Zhong et al 2009). PCA uses whole brain data to identify the regions with a variability relationship to the variability of other regions, and uses these areas to construct the model. In this way, the intrinsic structure of the data are emphasized, rather than fitting the data into a preconceived structure. They also enable the mapping of networks of brain activity that vary together, and, because they create models of variability rather than processing voxel by voxel, they eliminate much of the noise inherent to univariate analysis. This mitigates the problems of researchers looking preferentially to areas they expect to be relevant.

Thus, while the method is imperfect, it continues to improve, and fMRI currently represents the best non-invasive technique available to examine brains “in action”. As such, this review will look to the data of fMRI studies for clues surrounding the neural substrate of delusions, while keeping a simultaneous eye to the underlying potential problems researchers face when tackling a complex psychological concept/construct like delusions with the somewhat limited tools currently available to them.

THE CHALLENGE OF CLASSIFYING DELUSIONS

After establishing that fMRI may the best modality available, we can consider ways in which we might study delusion. A question that immediately arises is how or whether to distinguish different types of delusions. If different mechanisms lead to their differing manifestations, but we do not take care to separate the types, we may lose our ability to detect unifying patterns driving specific delusion types. Dividing too specifically, on the other hand, would risk losing important distinctions, and mean more patients would be needed to gain sufficient study power. It is a challenge, then, that there is no single clear and widely accepted classificatory schema for delusions, and it remains unclear whether they should be considered unitary or diversified phenomena. Most of the current classification systems include as categories persecutory delusions and delusions of reference, which together (if accepted as categories) comprise the majority of delusions in schizophrenia. Both have also been explored to some extent using fMRI. Delusions of reference often co-occur with other “types” of delusion, including persecutory delusions, and have been theorized to relate or even underlie still other delusional types. Because they are the most common delusional types in schizophrenia and also the most studied, persecutory delusions and delusions of reference will be explored in this paper in order to try and gain a better understanding of what might underlie these symptoms. The implications of the classification of delusion will be discussed further in the conclusion.

THE ROLE OF ABERRANT SALIENCE IN DELUSIONS

A body of evidence points to certain regions of the brain which, in healthy subjects, have been found to play a role in marking stimuli as relevant, or warranting attention. Referred to as the 'salience network' (SN), these regions include the anterior cingulate and anterior insula, as well as 3 key subcortical structures (amygdala, ventral striatum, and substantia nigra/ventral tegmental area) (Menon 2015). The leading model to date for delusions in schizophrenia is based upon the aberrant salience hypothesis of psychosis, where inappropriate salience attribution might lead to an incidental stimulus acquiring unwanted causal and motivational significance, resulting in delusion formation, while continuing deficits in the recruitment of appropriate executive attentional systems hampers the correction of the belief in response to external stimuli (Palaniyappan & Liddle 2012). In spite of having its origins in discoveries driven somewhat by chance (as will be further discussed in the conclusion), the aberrant salience hypothesis of delusions is supported by a growing body of evidence.

Multiple studies have found abnormalities in salience network structures in patients with schizophrenia, some of which have been found to correlate with the severity of delusions. Palaniyappan and Liddle (2012) review these studies and propose a model highlighting the role of the insula, a region that has been linked with prediction error coding (Singer et al 2009, Bossaerts 2010, Murray et al 2008). Studies examining salience network structures in patients with schizophrenia compared to controls find increased activation of the insula without the normal accompanying cingular activation; other studies examining the functional connectivity of these structures find a reduction in the connectivity between both insula and ACC with other brain regions (Palaniyappan and Liddle 2012). In Palaniyappan and Liddle's model, abnormal insular activation corresponds to inappropriate salience generation during an otherwise normal activity or in response to internally generated stimuli, and an accompanying lack of cingulate activation demonstrates a disruption in normal SN activity that could reflect abnormal error-monitoring circuitry. In other words, abnormal insular activation leads an individual to perceive an incidental or internally generated stimulus as salient, which, in the context of abnormal error monitoring systems, results (perhaps with repetition, over time) in delusion formation and maintenance.

ABERRANT SALIENCE TO ABERRANT BELIEFS

Abnormal salience appears to play an integral role in the mechanisms underlying delusions, but it does not constitute the entire picture of the phenomena. What happens once a stimulus is marked as significant? How does this become an unusual belief? The "belief" aspect of delusions appears to involve specific content beyond the feeling that a stimulus is important, that survives as "meaningful" in spite of what others might consider evidence to the contrary. What attaches to that stimulus the idea content that a patient is being followed, or that a patient is famous and that's why everyone is talking about her? The case of persecutory delusions will be used to explore this next step.

PERSECUTORY DELUSIONS

Delusions of persecution are often considered a hallmark symptom of schizophrenia, affecting up to 70% of individuals carrying the diagnosis. They consist of preoccupations that an individual is being persecuted, conspired against, or potentially harmed (DSM 5). The current predominating theory of persecutory delusions is that they arise from a heightened perception of environmental threat. The amygdala has been implicated as playing a key role in threat and salience processing, and investigators studying paranoia naturally postulated a link between amygdalar activity and paranoia. Accordingly, patients with paranoid schizophrenia have been found to have an increase in baseline amygdalar activity (Pinkham 2015), as well as a heightened autonomic response (as measured by skin conductance) to facial expressions of fear (Russell 2007, Williams 2007). A growing body of evidence links persecutory delusions with abnormal threat processing, consistent with the concept of paranoia as continual perception of environmental threat (Holt et al 2009, Linnman et al 2013, Jensen et al 2008).

Recently, Perez et al (2015) correlated persecutory delusion severity in medicated patients with schizophrenia and schizoaffective disorder with BOLD fMRI of patients in an instructed fear paradigm, with conditions designed to elicit anticipation of threat vs safety. Briefly, the patients were told they might get a shock if they saw a particular color cue, while they would definitely not get a shock if they saw a different cue. Persecutory delusion severity was graded based on the persecution/suspiciousness item of the positive and negative syndrome scale (PANSS, Kay et al 1987). It was found that the primary/association visual cortex, lateral orbitofrontal cortex and left V4 color area all exhibited increasing activation correlated with increasing persecutory delusion severity in the fear vs safety and fear vs baseline conditions. The researchers formulated these findings as increased processing of threat signals, with the magnitude explained by increased activation in fear centers; they also had decreased bilateral posterior dorsal ACC and right insula activation in the fear conditions, explained as “aberrant salience processing”. Then, in the safety vs baseline condition, increasing severity of PDs was correlated with decreased lateral orbitofrontal cortex and ventral occipital temporal cortex as well as decreased ventral occipital temporal cortex (including the V4 color area), which the researchers formulated as less visual processing of safety cues, as well as an increase in bilateral dorsal ACC and anterior insula (components of the aforedescribed salience network). This study, then, illustrated a sensical picture of persecutory delusions being linked to both increased processing of potentially threatening stimuli and aberrant salience processing.

Taken together, the results of these studies could suggest that patients with persecutory delusions have high baseline amygdalar activity, and associated heightened processing of threat cues and reduced processing of safety cues, which combine to transform abnormal salience into delusions of persecution. We might imagine these patients as feeling constant heightened feelings of fear, coloring their perceptions of the world, paired with a reduced ability to perceive safety, meaning they never fully deactivate the fear, even in situations others would interpret as safe. This formulation would be consistent with the aberrant salience hypothesis of delusions.

DELUSIONS OF REFERENCE

Delusions of reference (DoR) are another of the most common positive symptoms in schizophrenia (Lariviere 2017, Fletcher and Frith, 2009), estimated as existing in 67% of patients with the disorder (Menon 2011, citing a WHO estimate from 1973), often co-existing with (and/or possibly underlying) other types of delusions. They are described as misinterpretations of external stimuli wherein neutral events trigger ideas or sensations of heightened self-relevance. DoR stand out from the other types of delusions in that they appear to occur in discrete moments, lending to their study in fMRI experiments. Because they occur "in the moment" and therefore can be triggered by various tasks and observed in process in the fMRI machine, DoR appear to be the delusion type most focused upon in research. Their broad manifestations and unique temporal nature may underlie why some classificatory systems for delusions categorize DoR a separate class of delusions, while others incorporate them as components of other types of delusions (for example, persecutory delusions and delusions of grandeur), and still others do both (Startup and Startup 2005). In some cases, DoR appear to play a role in the formation/maintenance of other delusion types (for example, persecution or grandeur), while in other cases they can appear in isolation without associated suspicion or belief that they are due to the individual's, say, special talents. After a deeper dive into what the current literature suggests about DoR, a potential causative link between DoR and other delusion types will be explored.

THE ANATOMY OF SELF REFERENCE

Meta analyses of contemporary structural and functional studies suggest that self referential processes are mediated in large part by the cortical midline structures (CMS) in the brain, comprised of the medial prefrontal cortex (mPFC, which can further be subdivided into ventromedial and dorsomedial, vmPFC, dmPFC) as well as the anterior and posterior cingulate cortices (ACC, PCC) (van der Meer et al2010, Murray et al 2012, Northof et al 2006). Several white matter structural studies highlight differences in these brain regions thought to be integral to self-referential processing in patients with delusions of reference. Fitzimmons et al (2015) examine the fractional anisotropy of the cingulum bundle and find that the magnitude of the radial diffusivity (water moving perpendicularly to axons) and trace diffusivity are both correlated with delusion of reference strength in first episode schizophrenics. These findings could be formulated as, for example, disruption in the cingulum as contributing to DoR. That these were first episode patients removes the possible confounding of medication effects. Tao et al (2015) examined gray matter density in the caudate head portion of the striatum (a key subcortical node in prefrontal networks), and found it to be lower in first episode psychosis patients with DoR, with a negative correlation between caudate head gray matter density and strength of DoR. Palaniyappan et al (2011) find that lower volume of anterior cingulate and insula in

medicated patients with schizophrenia on the L side showed some correlation to severity of DoR.

With regards to functional MRI, multiple studies of self-reference in schizophrenia reveal a wide range of altered CMS activation during self referential judgments compared with healthy controls. The results of two related studies by the same group (Menon et al 2011 and Lariviere et al 2017) show differences in self-reference task performance between patients and non patients. Both studies found that schizophrenia patients with DoR thought that more neutral sentences referred to them compared to healthy controls. The self-reference task in these studies sought to evoke self-referential ideation by mixing sentences that reflected particularities about each subject with sentences not designed to refer specifically to the subjects. Patients were shown a sentence and asked whether they "got the feeling" that a sentence was written specifically about them. One third of the sentences were "neutral" (e.g., "he collects CDs"), one third "emotionally salient" (e.g., "he was in a horrible accident"), and one third "personally salient" (subject-specific statements crafted based on screening interview performed one week previous to the study about the particular subject's current life circumstances, hobbies, interests, and symptomatology).

In the 2011 study, Menon and colleagues used univariate analysis of whole brain fMRI BOLD signal data comparing the subject groups and endorsement conditions, and found that while controls showed a sharp decrease in activity in the medial prefrontal cortex (mPFC) and ventral striatum when sentences were judged as non self referential (non-SR) versus self referential (SR), the delusional patients showed reduced differences (more activity in the mPFC when judging sentence to be SR). In other words, there was less differentiation in this region for the DoR patients between SR and non-SR processing. These structures were previously connected to self referential processing (as outlined above), and the authors formulated their findings as this reduced differentiation making it more difficult for patients to correctly reject information as non-self-relevant.

An important limitation of the 2011 study by Menon et al study was that the subjects with DoR were all medicated patients with schizophrenia, while the control group were unmedicated subjects without schizophrenia, meaning it was difficult to know whether effects observed were related to DoR specifically or simply due to medication effects or effects of schizophrenia in general. In recognition of this limitation, the same group (this time with first author Lariviere) performed essentially a repeat of the previous study, but with the addition of a non-delusional schizophrenic patient group to control for the effects of medication and non-DoR aspects of schizophrenia (Lariviere 2017). They also employed a different approach to identifying regions of interest in the fMRI images, using principal component analysis (PCA) to identify network of brain regions that explained the differences between the endorsed and non-endorsed conditions, and then comparing those regions between groups.

The researchers again found higher rates of self reference in the patients with DoR relative to controls, but this high rate was matched by the patients without DoR. Their imaging findings were also not precisely in support of their previous findings being attributable to DoR specifically (which could possibly be explained by the different method of statistical analysis used). They identified three principal components, where each component consisted of a network of regions that acted together in predictable ways to distinguish the brain activity during self referential processing. The networks were temporally staggered (during the task, component 1 parts activated first, then 2, then 3).

The first (component 1) consisted of what they called the “posterior cortical midline structures” (posterior cingulate cortex, precuneus) as well as subcortical structures such as the dorsal (caudate and putamen) and ventral striatum (the latter being part of salience network), and the thalamus. The researchers found that when subjects judged a statement to be non-self-referential, activity in component 1 decreased, while when self-reference was endorsed, there was less deactivation (i.e. more activity) in these posterior CMS and subcortical structures. Interestingly, the only significant difference between study groups noted here was that the non-delusional schizophrenic patients showed muted deactivation when they judged a sentence non-SR relative to the healthy controls (there was less difference in the activation of this network between SR and non-SR processing). The group’s previous study (Menon et al 2011) found this similar effect in the subcortical portions of these regions, but in patients with DoR compared to healthy controls, which was not noted here. Taken together, these data seem to suggest that the muted deactivation in the CMS and subcortical structures during self referential processing in one or the other group of patients could be secondary to either medications or non-DoR specific effects associated with schizophrenia.

Next, component 2, which they called the “salience/response” network, comprised the key salience network components of the dorsal ACC and bilateral anterior insula along with sensorimotor components (consistent with the subjects’ right handed response) that differed in its activation pattern between endorsed and non endorsed conditions. Coactivations of the SN with response-based network consistently reported in past (Ide 2013, Menon 2010)~~,~~ strongly implicate SN in influencing motor responses to salient stimuli. Interestingly, a significant difference between the two patient groups in this study (but not relative to the healthy controls) occurred in this component early in the post stimulus time course (at around one second) when subjects endorsed a statement as self-referent, with the non-delusional patients having a greater BOLD activation in this component. A greater BOLD signal for this component in this group relative to the patients with delusions was also observed later in the time course (by approximately two seconds) of the non-endorsed condition. Both patient groups demonstrated significantly reduced activity much later (at around 20 seconds) in the endorsed condition relative to healthy controls, while only the non-delusional patients exhibited significantly reduced activity here in the non-endorsed condition.

Then, occurring after components 1 and 2, was component 3, consisting of increased activations in anterior CMS regions like the orbital PFC and the mPFC, as well as lateral temporal regions, and decreased activations in some supplementary/sensorimotor areas when statements were judged as self-referential. It is notable that component 3 includes the mPFC, which, as aforementioned, is thought to be integral to SR processing. This anterior CMS network exhibited the highest BOLD signal in delusional patients, the lowest in controls, with non-delusional patients in between during SR processing, with this difference only reaching statistical significance between the delusional patients and the healthy controls (not between the two patient groups). Both patient groups exhibited a slower return to baseline relative to healthy controls that was statistically significant.

The findings are certainly not straightforward, particularly with regards to parsing differences between patient and control groups, and do not align exactly with the findings of the group’s previous study. The three components do, however, provide some support for the concept of distinct networks that we can see as mediating salience attribution and evaluation of stimuli for self reference, and a model for the relationship between these networks. That both patient groups exhibited reduced ability to suppress the posterior CMS network (though this was only

significant in the non delusional group) of component 1 could be seen as supportive of the salience hypothesis of delusions in schizophrenia, in that it could reflect a failure to dampen the salience of irrelevant external stimuli. That both patient groups showed hyperactivity in component 3 relative to healthy controls (and that this was only significant in the delusional patients) could be seen as reflecting increased activity in centers of self referential processing that led to the increased rate of endorsed self-reference in both patient groups relative to control. That component 2 consistently occurred between these two networks suggested a mediative role. The authors integrate these findings it into a model of SR processing whereby salience attribution to stimuli is first mediated by posterior regions and the ventral striatum (component 1), followed by reflective re-evaluation in the anterior prefrontal and lateral temporal regions (component 3), with the transition mediated by the salience network (component 2).

THE ROLE OF SELF-REFERENCE IN DELUSIONS

SImilar to aberrant salience, self-referential processing appears to play some role across the classes of delusions. Classificatory systems cannot seem to agree on whether delusions of reference constitute their own type of delusion or whether they should be seen to comprise a part of the other types. While they likely reflect a diathesis toward self referential processing, delusions of reference stand apart from other delusions in that they appear to occur in discrete moments in time, in relation to discrete perceptions, and do necessarily involve the complex emotionally-valenced explanatory models of persecutory delusions or delusions of grandiosity.

There has been much discussion relating aberrant salience attribution to self referential processing; these concepts seem to be related, with the processes conceptualized as underlying each both leading to a change in attentional allocation. A recent study by Pankow et al (2016), for example, sought to explore the relationship between aberrant salience processing and SR processing in schizophrenia. They found that schizophrenic patients had less activation in their vmPFC during SR processing, and that the degree of this decreased activation was correlated with rates of aberrant salience attribution. The authors formulated these findings as supportive of a connection between aberrant salience attribution and decreased neural self-referential processing, suggestive of disturbed attribution of relevance during self-reflection.

A detailed examination of the tasks, however, should raise concern for the validity of equating these tasks with what they purport to test.

The study compared 31 individuals with schizophrenia, 24 individuals with subclinical delusions, and 50 healthy controls. Their test of salience consisted of a task wherein subjects tried to win points by rapid responses to a target stimulus preceded by a conditioned stimulus. The subjects were then placed in an fMRI and imaged while they performed a task designed to induce self-referential processing. It was found that in the salience task, the subjects with schizophrenia had the highest implicit aberrant salience scores, followed by the delusion-prone subjects, with healthy controls exhibiting the lowest. However, the only statistically significant difference was between the patients and the healthy controls. This supports a dimensional (rather than categorical) approach to diagnosis. Explicit aberrant salience scores were similar across groups.

The fMRI during self-reference tasks revealed reduced ventromedial prefrontal cortex (vmPFC) activation in the patients with schizophrenia only (not the delusion prone individuals) during self-reference. This reduction in vmPFC activation was found to correlate with aberrant salience attribution (patients that thought more neutral sentences referred to themselves also had more aberrant salience attribution, as measured by the task). The task to measure salience and the task to measure self-reference both illustrate the importance of seemingly small elements of task design with respect to what is being probed and how the data are interpreted; as such, they will be analyzed in further detail.

For the salience task, the conditioned stimuli consisted of images of colored objects (red animals, blue animals, red household object, blue household object), wherein one dimension of the object (e.g., color) was associated with high or low likelihood of being followed by the target stimulus to which subjects were to react as quickly as possible in order to maximize their reward, while the other dimension had no association with the likelihood of the occurrence of the subsequent target stimulus (e.g, if the conditioned stimulus was red versus blue there might be an 87.5% versus 12.5% likelihood of being followed by the target stimulus, while it being an animal versus an object did not affect the likelihood of the target stimulus occurring). “Explicit” salience attribution was measured by asking participants twice during the task to rate the percentage of times they thought the four different stimulus types were associated with reward. Reaction times were used to measure “implicit” salience attribution. “Adaptive” salience was the difference between the reinforced and non reinforced stimuli of the relevant cue dimension (e.g., blue > red), while “aberrant” salience attribution was the difference between the two different irrelevant cue types (e.g., animals vs objects). The relevant vs irrelevant cues and cue dimensions were randomized across the patients, and each task consisted of 2 practice runs and 2 experimental blocks with 64 trials each.

Following this exercise, subjects were placed in an fMRI scanner and evaluated for self-referential processing with the performance of three yes/no tasks: to judge whether a neutral personality trait word was applicable to (1) themselves or to (2) Angela Merkel, or (3) whether the word had exactly 2 syllables (as a control task). These 3 conditions were presented in 18 alternating blocks of 5 words each, to which participants responded ‘yes’ or ‘no’.

It is notable that the methods for evaluating aberrant salience attribution - whether there was a significant difference in reaction time between animals vs objects in cases where the color was the relevant marker, for example - is designed to catch cases where subjects systematically misattribute salience to the irrelevant category. In other words, rather than measuring the rate of individual events of aberrant salience attribution, their metric for “aberrant salience” measures the degree to which subjects incorrectly associate a category (color or object type) with reward versus no reward. The aberrant salience attributed by a subject that reacts more quickly to animals than objects when the animals vs objects dimension was irrelevant would be captured in this measurement.

But if a second subject misattributed salience to particular individual images he or she saw, and this happened in a relatively even split across animals versus objects (e.g., incorrectly reacted quickly to blue animals 50% of the time and blue objects 50% of the time), though this individual may have had just as many aberrant salience attribution events as the previous individual, this second subject would be evaluated as having an aberrant salience score of zero. The implications of missing these patients is only relevant if this occurred at a different rate across groups. And even then, given the broader claims of the paper (that patients with schizophrenia

had higher rates of aberrant salience attribution), this would be most relevant in the case that this non-systematized aberrant salience attribution occurred more in the healthy control individuals. Still, the relationship between the “aberrant salience” that this study measures and the aberrant salience a patient with delusions assigns to a particular stimulus in the real world is not straightforward.

With regards what can be concluded based on the results of the self-reference tasks, the degree of neutrality of the words used becomes important. As long as some patient didn’t systematically perceive the neutral words as non-neutral, and have a relevant non-neutral feeling toward themselves and/or Angela Merkel, this task seems like it would capture a subject’s propensity for believing stimuli had specifically to do with them. A closer look at the way in which these words were chosen (Gruhn & Smith 2008) reveals that they were not found to be neutral in a systematic way to, for example, depressed versus non-depressed individuals. It is therefore not inconceivable that a patient with a depressed affect (as is not uncommon in schizophrenia) might perceive an inordinate number of these neutral words as having negative valence. If this patient also saw themselves in a negative light but liked Angela Merkel, they might be more likely to judge the ‘neutral’ words as referring to themselves than to Angela Merkel. This possibility could have been addressed by having the patients, for example, rate each word for its emotional valence, or by vetting the words beforehand with more diverse psychiatric populations to better ensure their “neutrality” vs. “non-neutrality”. But because the study at hand (Pankow 2016) does not control for these potential systematic biases regarding the neutrality of the words, and how this might interact with a non-neutral self-perception or view of Angela Merkel, we must exhibit caution in interpreting the experiment as reflective of self-reference. Thus, further research is needed to elucidate how, exactly, salience and self reference should be conceptualized to relate.

SALIENCE AND SELF REFERENCE AS DISCRETE ENTITIES

Even the degree to which salience and self-reference should be considered as distinct processes (or, conversely, whether they should be related at all) - remains an open question. It is important to acknowledge that even the notion of discrete networks responsible for mediating the processes of “salience” and/or “self reference” is arguable. The works of Menon et al (2011) and Lariviere et al (2017) provide evidence for a conceptualization of salience and self-reference as distinct but related subsystems involved in marking stimuli as self-relevant. This concept has been described in the literature prior to their work, with salience attribution and self-reference as two distinct but related subsystems that can function independently from or in conjunction with one another: one “reflexive” system for marking stimuli as warranting attention (similar to the SN as it has been described in this paper) and another “reflective” system that engages introspective processes (e.g., self-reflection, evaluation, recollection) (Menon et al 2011, Schmitz & Johnson 2007). If something like this model is accepted, delusions of reference might be conceptualized to occur when the reflexive system marks a stimulus as salient, and the reflective system connects this now-salient stimulus to memories or components of the individual’s concept of self.

This model could similarly be used to explore the relationship between self-reference and other delusion types. Integral to the model outlined for persecutory delusions, for instance, appears to be some component of self-reference, whether or not it be considered an independent component from fear and/or salience. A person with persecutory delusions doesn't just feel fear, but fears that particular events or stimuli are out specifically to get him/or her. How we conceive of the role of self reference in our model of persecutory delusions affects whether or not we might relate it to the abnormal self-referential processing that might be seen in patients with DoR. The question could be raised, for example, regarding whether the self reference component of persecutory delusions is integral to the feeling of fear itself (i.e., whether feeling fear necessitates self reference in itself, versus whether some distinct network for self-reference combines with a less specific anxiety to lead to the feeling of persecution). Further pointed studies will be required to address this question.

PATHOPHYSIOLOGICAL MODEL MODULARITY IN THE CLASSIFICATION OF DELUSIONS

There is no single clear and widely accepted classificatory schema for delusions, and it remains unclear whether they should be considered unitary or diversified clinical phenomena. The way in which we conceive of salience attribution and self-referential processing as distinct or overlapping processes affects the way in which we might classify delusions. If we consider SN involvement as its own component distinct from fear processing, we could say that aberrant salience combines with fear and self-reference to produce persecutory delusions, and relate the self-reference component to the self-reference seen in other delusion types. Here, the distinction between the SN, self-reference, and fear response functions would enable us to tie the pathophysiology of persecutory delusions with other types of delusions without the fear response. For example, an individual with persecutory delusions and another individual with grandiose delusions both assign meaning to things in the world which a normal person would probably find to be merely random or coincidental, such that such random, coincidental factors (a crow sitting on a fence, or a person who happens to be walking behind one in the subway station) are misinterpreted as being particularly meaningful. What would differentiate the persecutory from the grandiose patient, then, would be an additional factor of constant fear, or a constant sense of grandiosity, to combine with the repeated feelings of meaning into delusions of persecution or grandiosity. Delusions of reference, on the other hand, could be seen as the feelings of meaning, activations of the memory/self-reflection centers of the brain, without the associated fear or other valenced emotion to drive the creation of delusional explanatory models. This conception would also allow delusions of reference to contribute to persecutory or grandiose delusions with the addition of other components of the delusions.

In the case of grandiose delusions in particular, the complete paucity of fMRI studies examining this type of delusion specifically in schizophrenia makes it more difficult to comment on whether there are shared etiologic processes between these and delusions of persecution. One question for future research is whether a person with grandiose delusions has some of the same abnormalities as a person who has delusions of persecution, but the person with delusions of persecution has additional abnormalities, or whether the pathways are actually entirely different.

Future research could attempt to address this question more directly by, for example, comparing schizophrenic patients with persecutory versus grandiose delusions.

There is room for growth in the field of psychiatry with regards to acknowledging the assumptions made in the ways in which we categorize. For example, many of the studies used as evidence in this paper for the role of the salience network in the etiology of delusions do not explicitly distinguish between delusions and hallucinations, or between different types of delusions. Palaniyappan et al (2011), for example, correlated structural differences in the SN (volume reduction in the left anterior cingulate and anterior insula) of patients with schizophrenia relative to controls with the severity of 'reality distortion' as quantified with a single score from the Symptoms and Signs in Psychotic Illness (SSPI) scale, which includes both hallucinations and delusions (Liddle et al 2002). This was done under the assumption that delusions and hallucinations share similar etiologic bases with respect to the salience network, with delusions occurring when inappropriate salience is attached to external events, and hallucinations occurring when inappropriate salience is attached to internally generated events (Palaniyappan 2011). Though not explicitly stated in the study itself, this design also relies on shared etiologic roots across the range of delusion types experienced by the patients (which was not described). If either of these two conditions is not true, the results lose much of the specificity of their meaning with respect to the etiology of delusions.

Likewise, the majority of studies that probe the mechanism of persecutory delusions were designed to examine "paranoia", comparing patients with "paranoid" schizophrenia with "non-paranoid" schizophrenia. The studies using the DSM criteria for paranoid schizophrenia to distinguish these groups are of limited use for characterizing the mechanism of persecutory delusions, as the criteria require preoccupation with either delusions or frequent auditory hallucinations along with a lack of prominent negative or disorganized symptoms of schizophrenia (disorganized speech, disorganized or catatonic behavior, or flat or inappropriate affect). Williams et al in 2004 and again in 2007 use criteria that are slightly more specific in their study of a group of patients that they categorize as paranoid or non paranoid based on their scores in four areas of the positive and negative syndrome scale (PANSS, developed by Kay et al in 1987) - delusions, suspiciousness, grandiosity, and excitement. Unfortunately, this categorization still precludes distinguishing delusions of paranoia from, say, delusions of grandiosity, and/or an analysis of how they relate/interact.

If we do not acknowledge our assumptions, we run the risk of fitting data too vigorously into our pre-existing models and overlook possibly important findings. As an example, there were many aspects of the results from the study by Lariviere et al (2017) that are not accounted for in their interpretation of their data. The authors do not discuss the differences between groups found in component 2, for instance, in the discussion section of their paper. That certain statistically significant data are ignored while other data are highlighted to support a model that fits with current formulations tying salience, self-reference and delusions together raises some concern for possible confirmation bias confounding the conclusions drawn. If we attempt to look at the results without trying to pick and choose among them to fit them into the pre-existing, seemingly sensical models of brain network patterns in delusional patients, the lack of significant difference between patient groups combined with the fact that the non-delusional patients at times are further from controls than the delusional patients makes the entire framework conceptually not straightforward, suggesting that more thought may need to be given to the results themselves. While it is unavoidable that certain results be highlighted in order to synthesize information into an interpretable story, given that the sensical results are not consistently apparent across

studies and, in fact, are often seemingly contradictory, makes it difficult not to question the degree to which our bias towards what “makes sense” might actually limit our ability to understand the underlying processes. But again, since the studies, processes probed, data, and data analyses, are so complex, they are difficult to reproduce and therefore it becomes difficult to do more than cling to the few frameworks that seem to make sense with previous data surrounding seemingly related processes.

The origins of our explanatory structures provide all the more reason to be deliberate with our assumptions. The starting point for the aberrant salience hypothesis, for example, lies in the finding that antipsychotic medications that functioned to reduce the severity of the delusions and hallucinations of schizophrenia blocked dopamine receptors. This led to the theory of dopamine dysfunction as playing a major role in the pathogenesis of schizophrenia, which drove the imaging studies performed in the 90’s and early 2000s that confirmed this role (Heinz et al 2010). At the time, most of the foundational research surrounding dopamine had been performed in the study of drug addiction, where the dopaminergic signal had been conceptualized to represent a prediction error indicating the difference between predicted and received reward. This hypothesis was then transferred to a model of schizophrenic psychosis, wherein increased firing of dopaminergic neurons in the striatum of schizophrenia patients leads to the aberrant attribution of incentive salience to otherwise irrelevant stimuli (Kapur 2003). Much of the interpretation of the complex data from fMRI studies of delusions is performed with this in mind, either probing parts of the brain found in other studies to relate to salience attribution or highlighting findings involving these regions from whole brain data. Though much supporting data have since been elucidated, as the foundations of the basic structure of our model of delusions are based on a single finding that arguably arose by chance, it remains all the more important that we keep our sights broad as we interpret data resulting from studies with designs based on their resultant assumptions.

CONCLUSIONS

PRACTICAL CHALLENGES: METHODS, AND HOW WE SYNTHESIZE FINDINGS FROM MULTIPLE STUDIES

Even once we have examined our assumptions surrounding the nature of the models we use to design our study questions, there remain challenges to studying delusions surrounding the methods we use and how we compare data across studies.

The tool of fMRI itself is still relatively crude compared to the complexity of neural networks in the brain, and different results can be derived depending on the imaging method, image processing method, task design, and the amount of other random or unintended variation captured (anatomical differences, more head motion, etc). There is no consistent method used to either obtain the original images, or to process the data and thus provide the results (Haynes 2015). As is discussed in the introduction, the statistical models used to analyze fMRI data can also create confounding or even contradictory conclusions (Suma and Murali 2007). The imaging techniques themselves also have significant impact in the resulting data.

An illustration of the importance of understanding the limitations of the fMRI modalities is evident in initial fMRI studies examining BOLD activation response to fearful facial expression in paranoid versus non-paranoid schizophrenic patients. Contrary to what might be expected, these studies consistently reported amygdalar *hypo*activation in paranoid patients, paired with a more intuitive increase in autonomic response, with paranoid patients exhibiting the most pronounced "disjunction" between the two (Russell 2007, Williams et al 2007). An explanatory model for these findings was created that postulated that the paranoid patients had an increased autonomic response to the expressions that they were unable to "centrally" process. It was not until Pinkham et al (2015) used ASL fMRI to image similar subject groups at baseline and while observing fearful vs neutral facial expressions that it was shown that this decrease in amygdalar BOLD signal was likely due to a smaller *relative* difference secondary to an abnormally high baseline amygdalar activity in patients with paranoid schizophrenia. This example illustrates the significance of the fMRI imaging techniques in the outcomes and interpretations of these studies. Regardless of its inability to take absolute measurements and its more indirect relationship to CBF (which is, already, a proxy for the neural activity we seek to measure), its higher temporal resolution combine with an increased familiarity with and availability of this older technique to make BOLD continue to be the most common fMRI technique in contemporary studies. The use of both techniques simultaneously, however, has recently been shown to be possible and to improve data quality for both series of images (Cohen & Nencka 2018). Such simultaneous studies, perhaps even with the addition of EEG for high resolution temporal information, will likely be integral to improving the quality of future research.

Next, the tasks that subjects are asked to perform in different studies are often quite different. For example, comparing a task testing whether a subject can quickly differentiate the "salience" of blue vs. red instead of mistakenly thinking the salient factor was animals vs. household objects, followed by a list of (possibly) neutral words and looking at whether the patient thought the word was more related to themselves or Angela Merkel, vs a study in which entire sentences (rather than words, which involves a different level of language processing) are also pre-formed to particularly relate to a patient are questionably comparable. The relevance of the tasks are different. For example, it matters whether one is simply testing cognitive salience or primarily testing emotional valence as it relates to salience. The latter seems quite important when it comes to, say, delusions of persecution. Simply thinking that a particular drone that happens to fly through one's yard might relate to oneself or not may or may not be damaging to the person's function. But if one attaches negative emotional valence to the event, the drone goes from being not only salient but dangerous and, in turn, the patient might him/herself become dangerous, which is the worst possible outcome for both the patient and society.

In addition, the effects of antipsychotic medication (and possible other medications) can be difficult to disentangle from the neural features of symptoms that set patient populations apart from healthy controls. It is clinically apparent that many of these medications have a sedative effect - is it not possible that that might also mean regions of the brains of treated individuals are activated and deactivated more slowly? Some studies circumvent this problem by using patients with first episode psychosis, but the clinical course of these patients is unclear at the outset and this limits the population from which to sample. There are studies which have investigated pharmacologic effects on brain function in fMRI; aside from appropriate control groups, careful synthesis of the findings from these studies with the findings of studies probing the neural substrate of delusion will aid in elucidating the role of medications in any changes observed in fMRI.

CONCLUDING REMARKS

Our goals as clinicians or scientists, or at least the way we present them in order to get funding for our research, have practical reasons founded in an underlying, broadly accepted aim of lessening the suffering of human beings. We must therefore consider carefully what it is about delusions that we seek to treat, and how that does and does not relate to the etiologic processes underlying the delusion. Different processes might be targeted if, for instance, what leads a person with delusions to be dysfunctional is a matter of degree of the same sort of process underlying the delusions in a healthy individual, versus, for instance, an added factor of an inability to adjust or maintain multiple paradigms for reality to enable functioning. Or perhaps they all contribute, and it matters less what we target. In any case, uncovering underlying pathological processes has been integral to developing targeted treatment approaches across the disciplines of medicine, and psychiatry should be no exception.

Delusions are by nature complex to study. Though the current technologies available to study delusions are imperfect, much can still be (and is yet to be) learned with their use. Functional fMRI, particularly using multivariate data analysis techniques, can provide beautiful systems level data that informs our understanding of these complex neural phenomena. These data are inherently complex, and there is a temptation to ignore results that don't make sense. We cannot do this. But we also must recognize that the brain is complex, and the tasks we are having subjects perform elicit complex responses that are not necessarily what it is we seek to study. Attempting to compare results from different studies using different task designs and different fMRI data processing methods, then, becomes a daunting process. Careful consideration of the models we use to shape our understanding of delusions is necessary for more targeted experimental design and less biased interpretation of data. It is with care that we will be able to maximally harness the power of current technology to uncover the mechanisms underlying the delusions of schizophrenia, which in turn will pave the way for more deliberate treatment approaches.

REFERENCES

Bossaerts P (2010). Risk and risk prediction error signals in anterior in- sula. *Brain Struct Funct*. 214:645-53.

Braun CMJ, Suffren S (2011). A general neuropsychological model of delusion. *Cognitive Neuropsychiatry.* 16(1):1-39.

Chow MSM, Wu SL, Webb SE, Gluskin K, Yew DT (2017). Functional magnetic resonance imaging and the brain: A brief review. *World J Radiology.* 9(1):5-9.

Cohen AD, Nencka AS, Wang Y (2018). Multiband multi-echo simultaneous ASL/BOLD for task-induced functional MRI. *PLOS ONE.* 13(2): e0190427.

Corlett PR, Taylor JR, Wang XJ, Fletcher PC, Krystal JH (2010). Toward a neurobiology of delusions. *Prog Neurobiol*. 92(3):345-369.

Cummings JL, Mega MS. Chapter 12: Psychosis, delusions, and schizophrenia. From Neuropsychiatry and behavioral neuroscience. *Oxford University Press.* New York, NY. 2003.

Fitzsimmons J, Schneiderman JS, Whitford TJ, Swisher T, Niznikiewicz MA, Pelavin PE, Terry DP, Meholam-Gately RI, Seidman LJ, Goldstein JM, Kubicki M (2014). Cingulum bundle diffusivity and delusions of reference in first episode and chronic schizophrenia. *Psychiatry Res.* 224(2):124-132.

Gruhn D, Smith J (2008). Characteristics for 200 words rated by young and older adults: Age-dependent evaluations of German adjectives (AGE). *Behavior Research Methods*. 40(4):1088-1097.

Haynes JD (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls and Perspectives. *Neuron*. 15:257-270.

Huang H, Shu C, Chen J, Zou J, Chen C, Wu S, Xiao L, Liu Z, Wang H, Zhou Y, Wang G, Jiang T (2018). Altered corticostriatal pathway in first-episode paranoid schizophrenia: T Resting-state functional and causal connectivity analyses. *Psychiatry Research: Neuroimaging.* 272:38-45.

Kapur S (2003). Psychosis as a state of aberrant salience: a framework linking biology , phenomenology , and pharmacology in schizophrenia. *Am J Psychiatry*. 160:13-23.

Kay SR, Fiszbein A, Opler LA (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 13(2):261-276.

Lariviere S, Lavigne KM, Woodward TS, Gerretsen P, Graff-Guerrero A, Menon M (2017). Altered functional connectivity in brain networks underlying self-referential processing in delusions of reference in schizophrenia. *Psychiatry Research Neuroimaging.* 263:32-43.

van der Meer L, Costafreda S, Aleman A, David AS (2010). Self- reflection and the brain: a theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. Neurosci Biobehav Rev. 34:935–946.

Murray GK, Corlett PR, Clark L, et al (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry*.13:239, 267-76.

Murray RJ, Schaer M, Debbane M (2012). Degrees of separation: a quantitative neuroimaging meta-analysis inves- tigating self-specificity and shared neural activation between self- and other-reflection. Neurosci Biobehav Rev. 36:1043–1059.

Menon M, Schmitz TW, Anderson AK, Graff A, Korostil M, Mamo D, Gerretsen P, Addington J, Remington G, Kapur S (2011). Exploring the Neural Correlates of Delusions of Reference. *Biol Psychiatry.* 70:1127-1133.

Northoff G, Heinzel A, de Greck M, Bermpohl F, Dobrowolny H, Panksepp J (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. Neuroimage. 31:440–457.

Palaniyappan L, Liddle PF (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *J Psychiatry Neurosci*. 37(1).

Palaniyappan L, Mallikarjun P, Joseph V, White TP, Liddle PF (2017). Reality distortion is related to the structure of the salience network in schizophrenia. *Psychological Medicine*. 41:1701-1708.

Pankow A, Katthagen T, Diner S, Deserno L, Boehme R, Kathman N, Gleich T, Gaebler M, Walter H, Heinz A, Schlagenhauf F (2016). Aberrant Salience Is Related to Dysfunctional Self-Referential Processing in Psychosis. *Schizophrenia Bulletin.* 42(1):67-76.

Perez DL, Pan H, Weisholts DS, Root JC, Tuescher O, Fischer DB, Butler T, Vago DR, Isenberg N, Epstein J, Landa Y, Smith TE, Savitz AJ, Silbersweig DA, Stern E (2015). Altered threat and safety neural processing linked to persecutory delusions in schizophrenia: a two-task fMRI study. *Psychiatry Res*. 233(3):352-366.

Pinkham AE, Liu P, Hanzhang L, Kriegsman M, Simpson C, Tamminga C (2015). Amygdala Hyperactivity at Rest in Paranoid Individuals With Schizophrenia. *Am J Psychiatry.* 172: 8.

Schmitz TW, Johnson SC (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neurosci Biobehav Rev*. 31(4):585-596.

Singer T, Critchley HD, Preuschoff K (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn Sci*. 13:334-40.

Paolini E, Moretti P, Compton MT (2016). Delusions in First-Episode Psychosis: Principal Component Analysis of Twelve Types of Delusions and Demographic and Clinical Correlates of Resulting Domains. *Psychiatry Res.* 243:5-13.

Startup M, Startup S (2005). On two kinds of delusion of reference. *Psychiatry Research*. 137:87-92.

Suma HN, Murali S. Principal component analysis for analysis and classification of fMRI activation maps. *IJCSNS Int J Comp Sci and Network Security.* 7(11).

Tao H, Wong GHY, Zhang H, Zhou Y, Xue Z, Shan B, Chen EYH, Liu Z (2015). Grey matter morphological anomalies in the caudate head in first-episode psychosis patients with delusions of reference. *Psychiatry Research: Neuroimaging*. 233:57-63.

Thomas & Segal (2006). Comprehensive Handbook of Personality and Psychopathology: Personality and Everyday Functioning.

Thoresen C, Endestad T, Sigvartsen NPB et al (2014). Frontotemporal hypoactivity during a reality monitoring paradigm is associated with delusions in patients with schizophrenia spectrum disorders. *Cognitive Neuropsychiatry*. 19(2):97-115.

Williams LM, Das P, Harris AWF, Liddell BB, Bramer MJ, Olivieri G, Skerrett D, Phillips ML, David AS, Peduto A, Gordon E (2004). *Am J Psychiatry*. 161:3.

Williams LM, Das P, Liddell BJ, Olivieri G, Peduto AS, David AS, Gordon E, Harris AWF (2007). Fronto-limbic and autonomic disjunctions to negative emotion distinguish schizophrenia subtypes. *Psychiatry Research: Neuroimaging*. 155:29-44.

Zhong Y, Wang H, Lu G, Zhang Z, Jiao Q, Liu Y (2009). Detecting Functional Connectivity in fMRI Using PCA and Regression Analysis. *Brain Topogr*. 22: 134.

Zhou Y, Lian M, Tian L, Wang K, Hao Y, Liu H, Liu Z, Jiang T (2007). Functional disintegration in paranoid schizophrenia using resting-state fMRI. *Schizophrenia Research*. 97:194-205.