*Review*

# Neural Correlates of Aberrant Salience and Source Monitoring in Schizophrenia and At-Risk Mental States - A Systematic Review of fMRI Studies

**Joachim Kowalski, Adrianna Aleksandrowicz, Małgorzata Dąbkowska and Łukasz Gawęda \***

Experimental Psychopathology Lab, Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland
\*  Correspondence: l.gaweda@psych.pan.pl.

**Abstract:** Cognitive biases are an important factor contributing to the development and symptom severity of psychosis. Despite that various cognitive biases are contributing to psychosis, they are rarely investigated together. In the current systematic review, we aimed at investigating specific and shared neural correlates of two important cognitive biases: aberrant salience and source monitoring. We conducted a systematic search of fMRI studies of said cognitive biases. Eight studies on aberrant salience and eleven studies on source monitoring were included in the review. We critically discussed behavioural and neuroimaging findings concerning cognitive biases. Various brain regions are associated with aberrant salience and source monitoring in individuals with schizophrenia and the risk of psychosis. Ventral striatum and insula contribute to aberrant salience. The medial prefrontal cortex, superior and middle temporal gyrus contribute to source monitoring. The anterior cingulate cortex and hippocampus contribute to both cognitive biases, constituting a neural overlap. Our review indicates that aberrant salience and source monitoring may share neural mechanisms, suggesting their joint role in producing disrupted external attributions of perceptual and cognitive experiences, thus elucidating their role in positive symptoms of psychosis. Account bridging mechanisms of these two biases is discussed. Further studies are warranted.

Keywords: aberrant salience; source monitoring; psychosis; cognitive biases; self-disturbance; neural; fMRI

---

## 1. Introduction

Cognitive biases are identified as an important factor in the development and sustainment of symptoms of schizophrenia spectrum disorders (SSD) and positive symptoms in particular [1]. Also, numerous studies link cognitive biases with the clinical and non-clinical risk of psychosis [2–5]. The contribution of the cognitive biases to positive symptoms have been confirmed by clinical studies showing a beneficial impact reduction of cognitive biases on the severity of symptoms [6]. There is no single cognitive bias that is responsible for the development of psychotic symptoms [7]. To date, several cognitive biases have been found to be related to hallucinations [8–10] and delusions [11–13]. It is then likely that several cognitive biases play a role in the etiology of psychotic symptoms. In line with cognitive approaches to psychosis [14,15] as well as with the observations that different factors may have additive effect on the risk of psychosis [16,17], different combinations of cognitive biases may also play a role. And yet, various cognitive biases are relatively rarely investigated simultaneously, despite that such studies could advance understanding of how these biases interact and how they lead to the development of psychotic symptoms in combination. Such studies could help to find answers for another vital question - whether cognitive biases share similar mechanisms. One of the proposed levels of understanding of the mechanisms of cognitive biases is brain functioning [18–20]. In the present review, we summarised the results of studies on functional and structural neuroimaging results concerning aberrant salience and source monitoring biases, i.e. two

cognitive biases that have a prominent role in theoretical accounts [14,21] and empirical studies of psychotic symptoms. We aimed to try to elucidate their shared and unique brain correlates.

### 1.1. *Role of Aberrant Salience in Psychosis*

On a phenomenological level, aberrant salience is described as interpreting irrelevant as relevant, highly familiar as novel, having sharpened senses, keenness for things normally deemed negligible or a sense of newly found insight or understanding [18,22]. The persistent state of heightened and aberrant salience leads to the formation of delusional beliefs as a way to make sense of meanings found in this state. In his recognized paper, Kapur [18] has conceptualized "psychosis as a state of aberrant salience" giving priority to the process of searching the (ir)relevant meanings of stimuli. On a perceptual level, aberrant salience is associated with strong prior predictions or underweighted prediction error, which leads to excessive reliance on priors and undervaluation of sensory input [23]. This account explains how aberrant salience, understood as one of the disrupted precision mechanisms, contributes to positive symptoms of psychosis, e.g. delusions [24] and auditory hallucinations [25]. However, this model is still developed, with important questions yet to be answered [23]. On the behavioural level, one of the most prominently used markers of aberrant salience is a speeding of responses to irrelevant stimuli in undirected or probabilistic motivational tasks, like the Salience Attribution Test. Paradigms using different types than reward salience, like salience associated with novelty [26], are much less popular.

On the self-description level, aberrant salience is consistently linked with psychotic symptoms [27,28]. For a review of studies on non-clinical populations, see [2]. However, the results of behavioural studies are ambiguous in this regard. Studies show greater implicit and explicit aberrant salience in SSD [29–31] and UHR individuals [32,33] in comparison to healthy controls. Notably, these studies are inconsistent in obtained effects in different types of aberrant salience - either implicit or explicit. Some studies could not confirm differences between clinical and healthy control groups at all [34–36]. Inconsistency in reports of behaviorally measured aberrant salience is an argument for a systematic review of such studies. In the current study, behavioural effects will be discussed along with reviewing neuroimaging results.

### 1.2. *Role of Source Monitoring in Psychosis*

Source monitoring is another cognitive bias that has been early linked to psychotic symptoms in schizophrenia in theoretical accounts [37,38] that gained a lot of interest over the past decades as one of the cognitive biases that could provide a potential insight into the mechanisms in the development of psychosis. It refers to a set of processes involved in making attributions about the origins of memories, knowledge, beliefs and perceptual experiences [38]. According to the source-monitoring framework proposed by Jonhson and colleagues [38], source monitoring is based on characteristics of perception and memories in combination with the judgment process and consists of different types of discrimination. There are at least three types of source monitoring: external source monitoring, internal source monitoring and external-internal reality monitoring. Source monitoring processes serve to distinguish between various types of cues associated with different types of sources: sensory/perceptual information, contextual (spatial or temporal) information, semantic detail, affective information and cognitive operations.

Bentall and colleagues [39] hypothesized that reality monitoring errors might play a role in the development of hallucinations. When patients with active hallucinations are uncertain of the source of a perceived event, they have a tendency to attribute internal events to an external source. Since then, there have been a growing number of studies investigating its role in psychotic symptoms for reviews: [5,40] and ultra-high risk states for psychosis [3].

A recent meta-analysis [40] on behavioural results showed that patients with schizophrenia have a significantly greater tendency to misperceive inner states as originating

from external sources and it has been more prominent in patients with auditory hallucinations. Importantly, the effect is evident across studies regardless of the adopted paradigm (action modality, time delay and design). Furthermore, the tendency in patients with schizophrenia to misidentify their own speech as that of another person appears to be related to delusions and positive symptoms generally [41].

### *1.3. Possible Integration of Aberrant Salience and Source Monitoring Deficits and Aims of this Review*

There are not many theoretical and empirical accounts that integrate aberrant salience and disrupted source monitoring into a comprehensive model of positive symptoms of psychosis. Griffin and Fletcher [42] presented an account joining source monitoring and predictive processing in a hierarchical model. In the predictive coding account of psychoses [23,25,43] aberrant salience is seen as a type of disruption in precision-weighing mechanisms. Proneness to strong prior beliefs (or underweighted prediction error [23]) about the experiences and expected shape of reality results in a top-down influence that takes precedence or outweighs perceptual input, resulting in false inferences. The resulting prediction errors and biased response criteria are the basis for source monitoring processes, which are taking place in suboptimal circumstances, leading to greater source monitoring bias and deficit [42]. In the minimal self-disturbance account [22,44–46] aberrant salience and source monitoring deficits are conceptualized as neurocognitive processes underlying a fragile or unstable minimal self. On a phenomenological level, this translates to a disturbed sense of ownership of cognitive processes and the body and a disturbed sense of agency (as in attributing own actions to oneself). However, preliminary empirical studies have found that source monitoring deficits may have more linkages to self-disturbances at very early clinical manifestations of the risk of psychosis, as compared to aberrant salience [45]. Both presented accounts assume that aberrant salience and deficits in the source monitoring are complementary and lead, more or less directly, to positive psychotic symptoms.

The aim of the current article was to systematically review fMRI studies investigating neuronal correlates of source monitoring and aberrant salience in schizophrenia and clinical high risk states of psychosis. The main focus was to better understand their brain-level characteristics and to investigate whether these two cognitive biases may share neuronal correlates. We conducted a systematic search of comprehensive databases to find fMRI studies of these biases in said groups. We extracted and analysed behavioural and neuroimaging results, mainly considering the effects of comparative analyses between clinical and control groups, as an indicator of specific brain correlates of the biases. Additionally, we also systematically analysed other results presented by authors, like the correlational analyses of symptom severity or clinical measurements with a BOLD response to a certain bias or bias-related within-group effects.

## 2. Materials and Methods

### *2.1. Literature Search Strategy*

The search was a part of a larger project on systematically reviewing associations between schizophrenia or psychosis risk and various cognitive biases. We systematically searched four comprehensive medical and psychological databases: PubMed, EMBASE, Scopus and PsycInfo. The search was conducted up to the end of January 2021, without specific criteria for a publication date. Articles in press were also considered if they were accepted for publication and available online.

The search was conducted by combining keywords regarding schizophrenia, risk of psychosis and keywords related to aberrant salience or source monitoring. Searched strings were as follows. For schizophrenia and risk of psychosis: (psychosis OR psychotic OR schizophrenia OR schizophrenic OR UHR OR "ultra-high risk" OR "ultra-high-risk" OR "ultra high risk" OR CHR OR "clinical high risk" OR "clinical high-risk" OR ARMS OR "at-risk mental state" OR "at risk mental state" OR "at risk mental states" OR "at-risk mental states" OR "psychosis risk" OR "risk of psychosis" OR "schizophrenia risk" OR

"risk of schizophrenia" OR "risk of schizophrenic" OR "schizophrenic risk"). For aberrant salience: ("aberrant salience" OR "abnormal salience" OR salience OR salient OR "salience dysregulation" OR "disrupted salience" OR "incentive salience" OR "salience attribution test" OR "salient stimuli" OR "emotional salience"). For source monitoring: ("source monitoring" OR "source-monitoring" OR "self-monitoring" OR "self monitoring" OR "externalizing bias" OR "self recognition" OR "reality monitoring" OR "source memory" OR "source-monitoring processes" OR "source monitoring processes" OR "internal source monitoring" OR "false memories").

A complementary reverse search was performed. We data-scraped references of found studies, removed duplicates and manually examined these references for any papers not identified by the computerised literature search.

### 2.2. *Literature Selection*

We deemed studies suitable for systematic review if they met the following inclusion criteria. First, we included only studies assessing cognitive biases with experimental tasks, measuring indicators like reaction times (RTs) or error rates. Second, we included only studies investigating individuals with schizophrenia, schizophrenia spectrum disorders (e.g. schizoaffective disorder, delusional disorder), at-risk mental states or clinical high risk of psychosis. Also, these conditions were to be assessed by a clinician (opposed to a so-called psychometric diagnosis). We did a broad search for comparative, correlational and longitudinal studies. Third, for the aberrant salience bias, studies should employ tasks measuring aberrant salience where meaning or saliency of the stimuli was undirected, subjected to a degree of probability, or participants were blinded as for the function of cues. This excluded simpler paradigms of reinforcement learning, like monetary incentive delay task [47,48]. Four, for the source monitoring bias, we included only these studies that used the task directly targeting the recognition of the source of information (source monitoring decision, e.g., self-other recognition). Studies that were focused on different patterns of neuronal activation during the processing of self vs. other with no source monitoring decision performance [49,50] were excluded from the review. Fifth, we included only studies in English.

Results of the database's search were scanned by title and abstract at first to exclude articles without a connection to the review topic. The second search identified articles that were relevant for the review. The second search was conducted independently by two researchers, JK and MD for aberrant salience and AA and ŁG for source monitoring. Any inconsistencies in selected papers were resolved by a discussion.

### 2.3. *Computation and Interpretation of Results*

A large proportion of researchers do not provide estimates for interpretable effect size. Where possible, we calculated or estimated an effect size for significant results and presented them as Cohen's d's. For interpretation of effect size, we choose cut-off points proposed by Sawilowsky [51]. The second calculator from the Psychometrica.de site (https://www.psychometrica.de/effect_size.html) was used whenever possible. When authors did not provide suitable data, calculators 5 and 6 were used to estimate effect sizes from values of statistical tests. To broaden our search, we also decided to include and describe results with p values that were borderline significant, i.e. $0.10 > p > 0.05$. We treated between-groups comparisons in experimental tasks of aberrant salience and source monitoring as results of primary interest, and all other types of analyses, like correlational analyses, as secondary ones.

### 3. Results

#### 3.1. Results of Systematic Review

A total of 1363 unique records for aberrant salient articles search were identified. After removing reviews, articles without an experimental measure of aberrant salience, not in English, without participants with schizophrenia or risk of psychosis, a total of 18 articles remained. Among these articles, we identified 7 fMRI studies of aberrant salience.

Reverse search resulted in an additional study, giving 8 in total. Of these, 7 studies were comparative and cross-sectional and one study was a longitudinal study with a mean of 17 months between measurements. Four studies concerned individuals with schizophrenia or SSD (one of these studies also concerned a group of healthy controls with high scores on a delusions inventory), 1 study concerned patients with FEP and risk of psychosis, and 3 studies concerned individuals with risk of psychosis. Five studies employed the Aberrant Salience Test. Three other used Implicit Salience Paradigm, Salience Integration Task and an Implicit Salience Attribution Task. Seven studies employed fMRI for aberrant salience measures and one study employed fMRI for measurement of self-referential processing, which was then correlated with the aberrant salience measure.

A total of 1083 unique records for source monitoring articles search were identified. After removing reviews, articles without an experimental measure of source monitoring, not in English, without participants with schizophrenia or risk of psychosis, a total of 113 (u mnie -3) articles remained. Among these articles, we identified 8 fMRI studies of source monitoring or self-monitoring. Reverse search resulted in three additional studies, giving 11 in total. Of these, 10 studies were comparative and cross-sectional and one study was an intervention study, where intensive training of cognitive functions was administered with 6 months between measurements. Ten studies concerned individuals with schizophrenia or SSD and 1 study concerned patients with FEP. We decided to include all types of source monitoring paradigms in the current review due to proven consistency in results reported in the previous meta-analysis on source monitoring [40]. Five studies used the source monitoring paradigm with instant discrimination of the source of the speech. The remaining 6 studies implemented the source monitoring task with an encoding phase and source memory identification phase with between 6 to 45 min delay. In most of the paradigms with two phases participants underwent fMRI scanning only during the second phase with retrieval. In all of those paradigms, the task was to discriminate between perceived/imagined or/and self/other generated stimuli or discriminate between different forms of stimuli presentation (self/other speech, sentences, word puzzles, pictures or pictures with labels etc.). Six studies used auditory stimuli or auditory stimuli accompanied with visual cues, in forms of either pre-recorded speech or speaking while undergoing fMRI. Additionally, one study manipulated participants expectancies using visual cues that were either congruent or incongruent with the speech as well as ambiguity (discrimination between distorted and undistorted speech).

### 3.2. Aberrant Salience

Table 1. presents details of 8 fMRI articles on aberrant salience.

### 3.2.1. Behavioral Results for Aberrant Salience

As for implicit aberrant salience (IAbS), only 3 of 8 studies showed significant differences between clinical and control groups [30,33,52]. The study by Katthagen and colleagues showed significantly higher IAbS in SCH patients than HC with a small effect size. The study by Pankow and colleagues [30] showed a significantly higher IAbS in SCH patients than HC with medium effect size, but not significantly higher than subclinical delusions group, although with a relatively similar medium effect size. The longitudinal study by Schmidt and colleagues [33] showed a significantly higher IAbS in UHR participants across both measurements with a large effect size. Although, this effect was only significant in the second measurement, where the difference between groups was very large in effect size and driven by a drop of IAbS score in the HC group compared to the first measurement.

For explicit aberrant salience (EAbS), only 2 of 5 studies reporting this parameter showed significant differences between clinical and control groups [32,33] and one on a trend level   [53]. In a study by Roiser and colleagues [32], UHR participants had higher EAbS than HC with a large effect size. In the study by Schmidt and colleagues [33], UHR participants had higher EAbS with a medium effect size in the first measurement but not in the second. The study by Smieskova and colleagues [53] showed a trend-level result for

comparing four groups - HC, ARMS, FEP with and without medication. Interestingly, in this study, HC had the second-highest indicator of EAbS. Comparison of the two extreme groups, FEP with medication and ARMS, showed a difference with a large effect size.

### 3.2.2. Behavioral Results for Adaptive Salience

For implicit and explicit adaptive salience (IAdS and EAdS, respectively) only 2 of 5 eligible studies using the Salience Attribution Test showed significant differences between clinical and control groups [30,33]. Notably, a study by Walter and colleagues [35] used Salience Attribution Test but reported only results for aberrant salience. In the Study by Pankow and colleagues [30] subclinical delusions group had significantly lower IAdS in comparison to HC with a large effect size but not in comparison to SCH group (small effect size). There was also a medium, but not significant, difference between SCH group and HC. Additionally, the SCH group showed significantly decreased EAdS in comparison to HC and subclinical delusions groups, both with a very large effect size. Schmidt and colleagues [33] showed significantly lower IAdS in UHR participants with a very large effect size across both measurements. The difference in the first measurement was significant and very large in effect size, and the difference in the second measurement was at a trend level with a medium effect size. In EAdS the results showed a significant difference with a large effect size across both measurements, the difference in the first measurement was significant and very large in effect size, and the difference in the second measurement was not significant with a medium effect size.

### 3.2.3. Neuroimaging Results for Group Differences in Salience and Reward Prediction

For group differences in the BOLD responses, there were 6 studies that reported such analyses. The study by Katthagen and colleagues [52] focused on modelling subjective relevance and did not present group comparisons in fMRI measurement. The study by Pankow and colleagues [30] used fMRI for a self-referential processing task, which results were correlated with results of the Aberrant Salience Test. Of these 6 studies, 4 showed significant group differences. Studies employing SAT used two contrasts of interest. Aberrant reward prediction - contrast of irrelevant cues assessed as high-probability of reward vs those assessed as low-probability of reward and adaptive reward prediction - contrast of relevant cues with high-probability of reward vs those with low-probability of reward [54], though they can be described as aberrant and adaptive salience contrasts [53]. The study by Walter and colleagues [35] showed that aberrant salience contrast was associated with greater BOLD response in L-insula in SCH group with higher levels of positive symptoms. In the study by Smieskova and colleagues [53], there were two significant differences in BOLD response in the aberrant salience condition. FEP patients without medication showed higher activation in comparison to HC in the R-cuneus and R-middle occipital gyrus, and HC showed higher activation in comparison to ARMS and FEP in L-inferior parietal lobule. Numerous effects for adaptive salience from this study are detailed in Table 1. Study by Schmidt and colleagues [33] found no significant effects for BOLD responses associated with aberrant reward prediction. However, they found significant effects for adaptive reward prediction - lower BOLD responses in UHR group's ventral striatum, calcarine sulcus and midbrain bilaterally and in the L-cuneus and L-middle temporal gyrus across both measurements. In the first measurement, such an effect was found in the ventral striatum bilaterally and the L-parahippocampal and L-middle temporal gyrus and cerebellum. And in the second measurement in both ventral striata. In the study by Winton-Brown and colleagues [26], there was a higher bold response for reward prediction cues in the UHR group in L-ventral pallidum and L-midbrain.

**Table 1.** Studies on aberrant salience.

1

| Study | Study type | Sample sizes | Clinical sample(s) characteristics | Experimental task | Aberrant Salience behavioural results | Main neuroimaging results - group comparisons | Additional neuroimaging results |
|---|---|---|---|---|---|---|---|
| **Schizophrenia participants** | | | | | | | |
| Esslinger et al., 2012 | A comparative study of FEP and HC; an fMRI study | FEP = 27 HC = 27 | Convenience sample from an admissions centre for a mental health hospital; never medicated | Implicit Salience Attribution Task (famous and non-famous faces) | - no significant effect reported for face x group interaction<br>- "Three way interaction indicated that patients profited less from famousness once the stimuli were colourful [control task condition, colourful vs dull]" - Esslinger et al., 2012, p. 117. | - "There were no significant differences in activation between patients and controls regarding the contrasts of interest in the whole brain or ROI analyses" - Esslinger et al.., 2012, p. 118 | - A correlation in the bilateral ventral striatum ROI between activation in two tasks contrasts, (monetary incentive vs control in monetary incentive task) vs (famous vs non-famous face presentation). Correlations for peak voxels revealed significant relationships between contrasts in FEP, right ventral striatum: r = .58, left ventral striatum: r = .48. Correlations in the HC were not significant. Additionally, correlation coefficients were not significantly different between groups. |
| Katthagen et al., 2018 | A comparative study of SCH and HC; an fMRI study | SCH = 42 HC = 42 | Convenience sample from in-patient and out-patient units; all participants had antipsychotic medication | Implicit Salience Paradigm | -SCH showed increased implicit aberrant salience compared to HC (d = 0.39)<br>- SCH showed greater irrelevance bias (described in the paper) compared to HC (d = 0.55) | not applicable | - relevance weighted prediction error (described in the paper) correlated with increased response in the anterior cingulate cortex in all participants<br>- relevance weighted absolute prediction errors (described in the paper) correlated with decreased response in the L-hippocampus in all participants<br>- irrelevance bias correlated with decreased relevance weighted prediction errors in bilateral nucleus accumbens response |
| Pankow et al., 2016 | A comparative study of SCH, subclinical delusions and HC; an fMRI study | SCH = 29 subclinical delusions = 24 HC = 50 | SCH: a convenience sample from a hospital department of psychiatry and psychotherapy. Most of the patients were medicated; subclinical delusions: people with results in 4Q of PDI from a large internet sample | Salience Attribution Test | - Implicit Aberrant Salience: SCH showed significantly greater IAbS in comparison to HC (d = 0.68), but not subclinical delusions group (d = 0.56<br>- Explicit Aberrant Salience: no significant differences<br>- Implicit Adaptive Salience: subclinical delusions group showed decreased IAdS salience in comparison to HC (d = -0.85), but not SCH (d = -0.14); difference between SCH and HC was also not significant (d = -0.52)<br>- Explicit Adaptive Salience: | not applicable | - a significant relationship between aberrant salience score in SCH related to self-referential task activation in the ventromedial prefrontal cortex (r = -0.60)<br>- no such correlation was observed for subclinical delusions group and HC |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | SCH showed significantly decreased EAdS in comparison to HC (d = -1.40) and subclinical delusions group (d = -1.24) | |
| Walter et al., 2016 | A comparative study of SSD with lower and higher levels of positive symptoms; an fMRI study | SSD with higher positive symptoms = 21 SSD with lower positive symptoms = 21 | A convenience sample of patients with FEP and chronic SCH patients with a history of violent offense; most of the patients were medicated | Salience Attribution Test | - Implicit Aberrant Salience: no significant differences<br>- Explicit Aberrant Salience: no significant differences<br>- Implicit Adaptive Salience: no data<br>- Explicit Adaptive Salience: no data | - aberrant salience was associated with increased BOLD response in left insula in the higher positive symptoms group | - aberrant reward prediction signals in the left posterior insula correlated negatively with the chlorpromazine-equivalent doses (r = −0.31) |
| **Schizophrenia + UHR participants** | | | | | | |
| Smieskova et al., 2015 | A comparative study of ARMS, FEP and HC; an fMRI study | FEP with medication = 12 FEP without medication = 17 ARMS = 34 HC = 19 | A convenience sample of patients with FEP (medicated and unmedicated) and ARMS patients | Salience Attribution Test | - Implicit Aberrant Salience: no significant differences<br>- Explicit Aberrant Salience: a trend result (p=0.096) for group effect (FEP with medication>HC>FEP without medication>ARMS) (d for FEP with medication>ARMS = 0.88)<br>- Implicit Adaptive Salience: no significant differences<br>- Explicit Adaptive Salience: no significant differences | - in aberrant salience FEP without medication showed higher BOLD response in comparison to HC in the R-cuneus and R-middle occipital gyrus<br>- in the aberrant salience HC showed higher BOLD response in comparison to ARMS and FEP in L-inferior parietal lobule<br>- in adaptive salience ARMS showed lower BOLD response in comparison to HC in R-supramarginal gyrus and R-inferior parietal lobule<br>- in adaptive salience FEP without medication showed lower BOLD response in comparison to HC in dorsal part of L-anterior cingulate gyrus, L-middle frontal gyrus and L-precentral gyrus<br>- in adaptive salience FEP with medication showed lower BOLD response in comparison to HC in R-insula and R-precentral gyrus, R-paracingulate gyrus and R-anterior cingulate gyrus<br>- in adaptive salience combined FEP groups showed | - Negative correlation between R-insular adaptive salience activation and hallucinations in FEP without medication (r = -0.64) |

|  |  |  |  |  |  |  | lower BOLD response in comparison to HC in R-pre-central gyrus and R-insula<br>- in adaptive salience FEP with medication showed lower BOLD response in comparison to ARMS in bi-lateral paracingulate gyri and L-anterior cingulate gyrus |  |
|---|---|---|---|---|---|---|---|---|

**UHR participants**

| Roiser et al., 2013 | Comparative study of UHR and HC; an fMRI and PET study | UHR = 18 HC = 18 | A convenience sample; 2 participants with anti-psychotic medication (authors checked that data from these participants did not alter the results) | Salience Attribution Test | - Implicit Aberrant Salience: no significant differences<br>- Explicit Aberrant Salience: UHR showed greater EAbS in comparison to HC (d = 0.88)<br>- Implicit Adaptive Salience: no significant differences<br>- Explicit Adaptive Salience: no significant differences | - no significant group differences | - in both groups collapsed there was a positive relationship between aberrant reward prediction response in the ventral striatum and explicit aberrant salience, "however, the slope of the regression line was significantly flatter and nonsignificant in UHR" (Roiser et al., 2013, p. 1331)<br>- aberrant reward prediction in the hippocampus was positively correlated with dorsal striatal dopamine synthesis capacity in HC (r = 0.65) and negatively in the UHR (r = -0.52) in a peak voxel<br>- adaptive reward prediction response in the caudate was positively correlated with ventral striatum dopamine levels in HC (r = 0.63) and negatively in UHR (r = -0.63) in a peak voxel |
| Schmidt et al., 2017 | A longitudinal study of UHR and HC; an fMRI study | UHR = 23 HC = 13 | A convenience sample from a clinical service for people at high risk of psychosis | Salience Attribution Test | - Implicit Aberrant Salience: in both timepoints UHR showed significantly higher IAbS than HC (d = 0.93); at follow-up UHR group showed significantly higher IAbS (d = 1.25) (drop of IAbS in HC), but not at baseline<br>- Explicit Aberrant Salience: at baseline UHR showed significantly higher EAbS (d = 0.78), but not at follow-up<br>- Implicit Adaptive Salience: in both timepoints UHR showed lower IAdS (d = -1.21); at baseline significantly lower IAdS (d = -1.33) and at follow-up at trend level (p = 0.062, d = -0.69)<br>- Explicit Adaptive Salience: | - no significant effects of group or time, and no group x time interactions, significant group differences in aberrant reward prediction at either baseline or follow-up<br>- no significant group × time interactions for adaptive reward prediction<br>- in both timepoints UHR showed weaker BOLD response than HC in the ventral striatum, calcarine sulcus and midbrain bilaterally and in the left cuneus and middle temporal gyrus in adaptive reward prediction<br>- at baseline UHR showed significantly less activation in the ventral striatum bilaterally and the left parahippocampal and middle temporal | - no significant relationships between changes in clinical features and longitudinal changes in brain activation during aberrant reward prediction<br>- improvement in abnormal beliefs over time was associated with the increase in activation during adaptive reward prediction in the right ventral striatum and in the supplementary motor cortex bilaterally (peak-voxel correlation) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | UHR showed lower EAdS (d = -0.84) in both timepoints; significantly lower EAdS (d = -1.07) at baseline, but not at follow-up (d = -0.42) | gyrus, and cerebellum during adaptive reward prediction<br>- at follow-up the UHR showed significantly less activation in the ventral striatum bilaterally during adaptive reward prediction | |
| Winton-Brown et al., 2017 | A comparative study of UHR and HC; an fMRI study | UHR = 29<br>HC = 32 | A convenience sample from a clinical service for people at high risk of psychosis | Salience Integration Task | - no significant differences for reward, novelty and aversion conditions | - UHR showed significantly greater BOLD response to reward-predicting cues than HC in the L-ventral pallidum and L-midbrain; no areas where HC showed greater activation than UHR<br>- no significant differences for novelty condition<br>- no significant differences for aversion condition | - reward-induced modulation of ventral striatum/pallidum to midbrain connectivity was significantly greater in UHR than in HC<br>- described above modulation was correlated with CAARMS unusual thought content in UHR (r=0.50), but not other positive symptoms subscales<br>- no significant correlations between group differences in activation related to reward processing and CAARMS positive symptoms |

FEP – first episode of psychosis group, HC – healthy controls group, SCH – schizophrenia or schizophrenia spectrum disorder group, ARMS – at risk mental states group, UHR – clinical ultra-high risk of psychosis group, IAbS – implicit aberrant salience condition, EAbS – explicit aberrant salience condition, IAdS – implicit adaptive salience condition, EAdS – explicit adaptive salience condition, CAARMS - Comprehensive Assessment of at Risk Mental States.

### 3.2.4. Additional Neuroimaging Results

Besides group comparisons, all of the studies report other types of analyses, like correlations between BOLD response from aberrant salience and other clinical features, e.g. psychopathological symptoms. Details of these results are presented in Table 1.

Interestingly, apart from main effects in the salience network regions reported in Smieskova and colleagues paper [53], there was a negative correlation between BOLD response in R-insula and hallucinations in a drug naive FEP group (r = -0.64). Another study [35] showed that BOLD response associated with aberrant reward prediction in the left posterior insula correlated negatively with the dose of antipsychotic medication. The activity of a salience-related brain structure was also associated with relevance weighted prediction error, i.e. increased activation in the anterior cingulate cortex in both clinical and control groups [52].

There were also several studies reporting effects in the basal ganglia, apart from the between-group effects reported in the papers of Schmidt and colleagues [33] and Winton-Brown and colleagues [26]. In the study of Roiser and colleagues [32] there were associations of dopamine synthesis capacity in the ventral striatum and adaptive reward prediction in caudate - positive in HC and negative in UHR, and analogous results for aberrant reward prediction signal in the hippocampus and dorsal striatal synthesis capacity. However, the authors reported only correlations for peak voxels, rendering results hard to interpret [55]. The study by Esslinger and colleagues [56] reported a difference in correlations in the bilateral ventral striatum between a monetary incentive vs famous and non-famous faces contrasts. However, they reported only peak-voxel correlations. Additionally, FEP and HC groups did not differ in the strength of their correlations. The longitudinal study [33] reported that improvement in abnormal beliefs over time was associated with the increase in activation during adaptive reward prediction in the right ventral striatum, however, they reported correlation coefficients only for the peak voxel, undermining the reliability of the analysis [55]. Another study [26] showed that reward-induced changes in connectivity between the ventral striatum and midbrain were significantly greater in UHR than in HC. Also, this change in connectivity correlated with CAARMS unusual thought content subscale in UHR participants.

### 3.3. *Source Monitoring*

Table 2. presents details of 11 fMRI studies on source monitoring.

### 3.3.1. Behavioral Results for Source Monitoring

Out of 11 studies considered in the current review, 10 implemented external-internal source monitoring paradigm. Of those, most studies showed significant differences between clinical and control groups. Results [57,58] where SSD patients were divided into those with active auditory verbal hallucinations (AVH) and those that did not experience auditory verbal hallucinations in their lifetime (non-AVH), patients with AVH made significantly more external misattributions while performing a source monitoring task than non-AVH patients and HC group. The study by Garrison et al., [59] showed that SCH had significantly less general correct responses than HC in reality monitoring with a large effect size. Study by Kumari and collaborators [60] also showed that the SCH group had impaired performance in source monitoring task in comparison with the HC group with a large effect size. In the study by Thoresen and colleagues [61], patients made more misattributions where they classified the imagined condition as a viewed condition with a very large effect size. Vinogradov and collaborators [62] demonstrated that the SCH group was significantly less accurate in self-generated items when compared with the HC (independently of age or poorer item recognition memory) with a very large effect size (although the sample size was relatively small). Study by Subramaniam and colleagues [63] investigated changes in the reality monitoring task performance after 16 weeks of cognitive training (participants trained cognitive functions such as basic emotion

recognition, theory of mind, memory, executive functions). Results revealed that patients who took part in the training compared to patients who played computer games showed a significant improvement in their accuracy performing a reality monitoring task. Another study [64] showed that HC performed better than the SCH group at overall source-memory identification across both self-generated and externally presented word items. In another study [65] there was a significant group effect in the overall task performance accuracy, that is patients revealed deficits in source monitoring as compared to HC. A study conducted on the FEP group [66] showed that patients made significantly more total errors than HC when listening to their own voice preceded by an invalid cue with a large effect size. Remaining one study implemented an external-external source monitoring paradigm [67]. Results showed no differences in the response bias index (tendency to re-member word-items as pictures) between SCH and HC groups. However, authors demon-strated that greater severity of visual hallucinations was associated with increased rates of false memories of pictures, but there was no relation with the verbal hallucination score. Results are presented in Table 2. in detail.

55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70

**Table 2.** Studies on source monitoring.

| Study | Study type | Sample size | Sample characteristics | Experimental task | Behavioural results | Main neuroimaging results - group comparison | Additional neuroimaging results |
|---|---|---|---|---|---|---|---|
| Allen et al., 2007 | Comparative study of SCH sample and HC, an fMRI study | SCH = 20 (division on AVH = 10; non-AVH = 10) HC = 11 | A convenience sample of psychiatric patients recruited through the South London and Maudsley National Health Service Trust | Source monitoring task (discrimination between self/researcher and distorted/undistorted speech ) | - a significant interaction between the effects of source of speech and group - a significant group difference in the self speech condition, AVH group made significantly more misattribution errors than the participants in both the non-AVH and HC | - a significant interaction between response accuracy (correct/misattribution) and group in the left middle temporal gyrus; in both HC and non-AVH groups there was greater activation for correct responses (correct attribution of either self or other speech) than for misattributions, whereas there was no significant difference in the AVH group; similar patterns of activation occurred when only self speech condition (i.e. the correct identification of self speech v. its misattribution to an external source) was analysed | - |
| Garrison et al., 2017 | Comparative study of SCH sample and HC, an fMRI study | SCH = 20 HC = 20 | A convenience sample | Reality monitoring task (discrimination between perceived/imagined and self/researcher word-pairs) | - SCH group had significantly less correct responses than HC (ηp2 = 0.175; d=0.92*) - no significant interaction between group and reality monitoring condition - SCH group had significantly slower RTs than HC (ηp2 = 0.239; d=1.12) | - no significant group differences in BOLD response - a significant three-way interaction between group, task, and region (left and right medial anterior PFC, left and right dorsolateral PFC) - a between groups comparison indicated a trend towards a significant reduction in reality monitoring activity in SCH compared with HC in the left medial anterior PFC region of interest only (p = 0.057) | - HC exhibited significant activity associated with the reality monitoring contrast in the left medial anterior PFC region of interest and in the left and right dorsolateral PFC regions of interest, and at a whole-brain corrected voxel-wise height threshold of p < .05, in the occipital lobe - in SCH group, significant reality monitoring-related activity was detected only in left dorsolateral PFC - no significant correlation between the percentage signal change for each participant in any of the a priori left and right medial anterior and dorsolateral PFC voxels, for the reality monitoring contrast and the working memory contrast, for either SCH or HC |
| Kambeitz-Ilankovic et al., 2013 | Comparative study of FEP and HC; an fMRI study | FEP = 20 HC = 20 | A convenience sample of psychiatric patients recruited through Maudsley National Health Service Trust | Source attribution task (discrimination between self-/other speech, ambiguity: undistorted/distorted speech and validity: valid/invalid cues) | - all interactions were nonsignificant, although there was a trend for an interaction between validity, source, and group (p = 0.063) - post hoc tests revealed that FEP made significantly more total errors than HC when listening to their own voice preceded by an invalid cue (non-self-cue) (d = 0.8) | - a significant interaction between group, validity, and source in the right MTG and in the left precuneus; in both these regions, HC showed greater activation for self-speech relative to other speech condition during invalidly cued (but not validly cued) trials; unaltered activation during both source and validity manipulations in these regions in FEP patients - A post hoc analysis showed that in | - a negative relationship in the FEP group between the activation in the right MTG and a severity of positive psychotic symptoms measured by PANSS positive symptoms subscale (r = −0.62) and the PSYRATS delusion items (r = −0.45) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | HC, relative to valid self-trials, invalid self-trial condition were associated with greater activation in the right Pc, MTG, and left insula. In FEP, there were no areas more active during invalid relative to valid trials.<br>- nonsignificant interactions between group and validity and between group, validity, source, and distortion | |
| Kumari et al., 2008 | Comparative study of SCH and HC; an fMRI study | SCH = 63<br>HC = 20 | A convenience sample | - HC showed more accurate performance than SCH group (d = 1.15)<br>- a significant group x source x distortion interaction (d = 0.54), but after complementing the results for percentage of correct answers the interaction changed to a trend level (p = 0.06)<br>- SCH group made more misattributions than HC (d = 0.98) | - SCH group showed more activity than HC in the ventral striatum, hypothalamus, and part of the thalamus in the other > self-contrast | - Within the SCH group, greater ventral striatal-hypothalamic activity during other > self contrast correlated with higher negative symptoms PANSS score (r = 0.29) |
| Mechelli et al., 2007 | Comparative study of SCH and HC; an fMRI study | SCH = 20 (division on AVH = 11; non-AVH = 10)<br>HC = 10 | A convenience sample of psychiatric patients recruited via the South London and Maudsley National Healthy Trust (SLAM) | Source monitoring task (discrimination between self/researcher and disorted/undistorted speech ) | - a significant interaction between source of speech and group; post-hoc tests revealed a significant difference between groups in the self-speech condition, patients with AVH misidentified their own speech as other speech more often than both HC and non-AVH patients | - a significant effect of source on the left superior temporal cortex to the anterior cingulate that differed across the three experimental groups; intrinsic connection was stronger for alien- than self-generated speech in HC and in non-AVH patients but not in patients with AVH, who expressed the reverse trend; the effect of source on this connection was significantly stronger in HC relative to AVH and in non-AVH relative to AVH, but did not differ between HC and non-AVH; thus, it is specifically impaired in patients with AVH<br>- a significant connection from right superior temporal cortex to the anterior cingulate that was stronger for alien- than self-generated speech in all three experimental groups, but no significant differences between the groups | - |
| Stephan-Otto et al., 2017 | Comparative study of SCH sample and HC, an fMRI study | SCH = 23<br>HC = 26 | A convenience sample of psychiatric patients recruited from the Parc Sanitari Sant Joan de Deu network | Reality monitoring task (discrimination between pictures and picture labels) | - The response bias index Br* was equivalent in the SCH and HC groups | - not applicable | '- no significant differential activation was observed for any of the contrasts studied when the patients with visual hallucinations were compared to the |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | of mental health services in Barcelona | | | other patients. The patients with verbal hallucinations presented increased bilateral activation in the thalamus and in the precuneus relative to the other patients when correctly remembering pictures (pictures remembered as pictures > pictures remembered as words)<br>- comparison of patients with low Br versus high Br showed a differential activation at a trend level of the left superior temporal gyrus when erroneously remembering pictures (words remembered as pictures > words remembered as words) |
| Subramaniam et al., 2012 | An intervention study of SCH; an fMRI study | SCH (division into AT: active training and CG: computer games training groups) = 31 (15 + 14 and 2 dropouts at baseline; 13 and 12 returned 6 month later)<br>HC = 15 | A convenience sample of psychiatric patients recruited from community mental health centres and outpatient clinics | Reality monitoring task (discrimination between self-generated or externally presented words) | Baseline assessment:<br>'- a significant group x condition interaction (d = 0.59); driven by the HC subjects who identified significantly more self-generated items than SCH subjects ( d= 0.84) , but not more externally presented items<br>- signal detection theoretic analysis confirmed that HC subjects performed significantly better than SCH subjects during overall source memory identification of word items (d = 0.66)<br>- The effect size of the overall source memory accuracy difference between HC and SCH subjects at baseline was 0.65.<br>Post training:<br>- a significant group x session (baseline and after 16 weeks) interaction in d-prime scores for overall source memory identification of word items (d = 0.92); a significant group x session effect for self-generated word items (d = 0.87) but not for externally presented word items. The SCH-AT subjects, when compared to the SCH-CG subjects, identified the source of significantly more word items overall at 16 weeks compared to baseline (d | Baseline assesment:<br>'- a significant group effect in mPFC activity for self-generated minus externally presented items. This group effect at baseline was driven by the HC subjects, who revealed significantly more mPFC activity for self-generated items than externally presented items when compared to the SCH-CG and SCH-AT, there was no significant difference in mPFC activity between SCH-CG and SCH-AT subjects at baseline<br>Post training:<br>- a significant group x session interaction in mPFC reality monitoring activity, driven by the SCH-AT subjects, who had significantly more mPFC signal activation after the intervention than the SCH-CG subjects than the HC subjects; no differences between sessions for HC or SCH-CG subjects in mPFC signal for the self-generated item minus externally presented item comparison. | - a significant correlation between mPFC signal and verbal memory scores at post-training in SCH-AT (r=0.51); no significant correlations at baseline nor in the SCH-CG group after intervention<br>- no significant correlation between mPFC activity and executive functions at post-training<br>- a significant correlation between the level of reality monitoring signal within the a priori spherical mPFC ROI immediately after training and ratings of social functioning at the 6 month follow-up (r=0.55) |

|  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|
|  |  |  |  | = 1.1) and also specifically identified more self-generated items (d = 1.01), with a trend effect for externally presented items (p = 0.07). The SCH-AT subjects, when compared to the HC subjects, identified the source of more word items overall at 16 weeks compared to baseline (d = 0.88), identifying more self-generated (d = 1.01) but not more externally presented items; no differences between sessions for HC or SZ-CG subjects on overall source-memory accuracy, on self-generated items or externally presented items |  |  |
| Subramaniam et al., 2017 | A comparative study of SCH; an fMRI study | SCH = 20 HC = 20 | A convenience sample of psychiatric patients recruited from a double-blind randomized clinical trial of cognitive training in schizophrenia (ClinicalTrials.gov NCT02105779). | Reality monitoring task (discrimination between self-generated or externally presented words) | - a significant main effect of group (d = 2.02); no significant interaction between mood and group, task accuracy and group, or between mood and task accuracy and group - between-group contrasts revealed HC performed significantly better than SCH at overall source-memory identification across both self-generated and externally presented word items (d = 0.87) | - ROI contrasts showed a significant between-group differences in which HC showed greater signal activation during source monitoring task than SCH group within the mPFC ROI for both positive and neutral mood states conditions and within the left putamen ROI for positive mood states; no significant between-group signal differences during source monitoring performance in the PCC for either positive mood or neutral mood states conditions or within the left putamen ROI for neutral mood - whole-brain analysis showed no between-group differences during negative mood states, when compared to the neutral mood condition (after FWE cluster corrections) | - when HC participants were in the negative mood, signal within the left dorsal region of the mPFC negatively correlated with externally-presented identification (r = -0.48), and with overall reality-monitoring performance (r = -0.44) - when SCH were in the negative versus neutral mood, signals in both left and right dorsal mPFC ROIs correlated with better externally-presented identification (left dorsal mPFC: r = 0.54; right dorsal mPFC: r = 056), and with overall reality-monitoring performance (left dorsal mPFC: r = 0.54; right dorsal mPFC: r = 0.56), despite the fact that patients did not show increased signal within these ROIs during the negative MI when compared to the neutral MI. |
| Thoresen et al., 2014 | Comparative study of SCH sample and HC, an fMRI study | SCH = 19 HC = 20 | A convenience sample | Reality monitoring task (discrimination between presented or an imagined object/scene) | - no significant difference between the groups for item recognition - a significant difference between the two groups in the imagined condition (IC); patients showed lower accuracy compared to HC (d = 0.62) - no significant difference in the viewed condition (VC) - patients made more misattributions where they classified the IC as VC (d = 1.30*) | - HC showed greater activity in left DLPFC for the contrast IC>Baseline and VC>Baseline and left hippocampus for the contrast IC>Baseline compared to the SCH group - BOLD response in the regions that displayed significant voxel-wise activation in conditions and between groups showed that both groups activated left DLPFC in IC compared to the baseline task and left hippocampus in IC compared to the baseline task | - a significant negative correlations with degree of delusions and left hippocampal activity in the IC in SCH group (ρ = 0.54); the relationship was still significant after controlling for dose of antipsychotic medication (r = 0.49), and PANSS general psychopathology symptom score, (r = 0.48), but not significant after controlling for Information, Digit Span, or Digit Symbol-Coding (WAIS-III subscales). - left hippocampal activity in the VC correlated significantly with delusions, |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | - SCH group also classified IC as new to a greater extent compared to the HC (d = 0.79)<br>- main effect of group in RTs; patients were responding slower than HC (d = 1.45)<br>- no significant interactions between source and group | | (ρ = 0.53)<br>- no significant association between activity in the regions of interest and the degree of hallucinations |
| Vinogradov et al., 2008 | Comparative study of SCH sample and HC, an fMRI study | SCH = 8<br>HC = 8 | A convenience sample | Source memory task (discrimination between the self-generated nouns and presented nouns in the sentence completion paradigm) | - no significant interaction effects for hit rates; however there was a significant effect of condition on source memory hit rate (with age as a covariant)  (d = 1.56); SCH group had a significantly lower hit rate for self-generated items compared with the HC (d = 3.04); no significant differences between SCH and HC groups on hit rates for external items; old items; or new items<br>- no interaction effects for RTs; there was a group difference at the trend level, SCH showed longer RTs on correct identification of self-generated items   (p = 0.08) and old items (p = 0.06) compared with HC group | -    Between-group analyses for the contrast of self-generated > externally presented items revealed that HC had significantly greater activity than SCH subjects in right rostral mPFC and in right superior frontal gyrus; SCH had significantly greater activity than HC in scattered areas that included left supplementary motor area, occipital cortex, anterior cingulate, and basal ganglia<br>- The SVC of mPFC on the group level revealed a significant cluster of 216 voxels for the HC, with a corrected P < 0.001, and no significant clusters for the SCH. For the between-group contrast of HC > SCH, the SVC revealed a significant cluster of 72 voxels with a corrected P < 0.001 and no significant clusters for the SCH > HC | - no significant correlation between age and percent signal change in mPFC during the self-generated condition |
| Wang et al., 2011 | Comparative study of SCH sample and HC, an fMRI study | SCH = 23<br>HC = 33 | A convenience sample of psychiatric patients recruited from from psychiatric hospitals and community health agencies in and around Vancouver, British Columbia, Canada | Self–other source monitoring task (discrimination between self-generated and other-generated words when solving a puzzle) | - no significant interaction between group and condition (p = 0.1)<br>- significant group effect was observed in performance accuracy (d = 0.81)<br>- no significant interaction between group and condition and no significant group effect in RTs (correct trials only) | - interaction analysis on the source-memory contrast of self-generated (SG) greater than other-generated (OG) found a significant between-group difference in the left superior temporal gyrus; HC subjects showed significantly higher mPFC-LSTG connectivity during OG than SG conditions, and SCH group showed significantly higher connectivity during SG than OG conditions. | - |

FEP – first episode of psychosis group, HC – healthy controls group, SCH – schizophrenia or schizophrenia spectrum disorder group, AVH – auditory verbal hallucinations group, RTs – reaction times, PFC – prefrontal cortex, mPFC – medial prefrontal cortex, MTG – middle temporal gyrus, DLPFC – dorsolateral prefrontal cortex, FWE – family-wise error correction, MI – mood induction, PCC – posterior cingulate cortex, SVC – small-volume correction, *index Br - the propensity to report words as pictures in case of uncertainty

### 3.3.3. Neuroimaging Results for Group Differences in Source Monitoring

For the external-internal source monitoring paradigm 9 studies reported group differences in the BOLD response. The one study that implemented the external-external source monitoring paradigm did not provide group comparison in the BOLD response. The study by Allen and colleagues [57] compared brain activation during source monitoring task in patients with auditory verbal hallucinations (AVH), patients without auditory verbal hallucinations (non-AVH) and HC group. Results showed that in HC and AVH patients greater activation was seen in the middle temporal gyrus for correct responses than for misattributions, while there was no significant difference in the AVH group. Another analysis of above mentioned data [58] revealed that intrinsic connection between left superior temporal and anterior cingulate cortex was modulated by source of speech in patients without AVH and healthy controls but not in patients with AVH, where the reverse trend was found. The study by Garrison and colleagues [59] showed that performance in the reality monitoring task is associated with the differential BOLD response in the medial anterior prefrontal cortex between investigated groups. In the study by Vinogradov et al. [62] results revealed that the SCH group manifested a deficit in the rostral medial prefrontal cortex while undergoing a source monitoring task. On the other hand, SCH had significantly greater activation in the left supplementary motor area, occipital cortex, anterior cingulate, and basal ganglia. Intervention study by Subramaniam and colleagues [63] showed that during baseline assessment HC group revealed significantly more mPFC activity for self-generated items than externally presented items when compared to the patients with schizophrenia that underwent computer games condition (SCH-CG) and patients that underwent active training condition (SCH-AT). While, after intervention SCH-AT, had significantly more mPFC activity than SCH-CG, no differences were found for HC and SCH-CG. In another study [64] in addition to the previously used paradigm [62,63] mood state conditions were implemented, where positive, negative or neutral mood states were induced prior to the task completion. Results showed differential activation of BOLD response in the right dorsal mPFC during the task performance when HC and SCH groups were compared. Study conducted by Thoresen and colleagues [61] displayed that in SCH patients reduced accuracy in the imagined condition was accompanied with reduced activity in the left dorsolateral prefrontal cortex as well as in the left hippocampus. Other results [60] demonstrated that the SCH group showed more activity than HC in the ventral striatum, hypothalamus, and part of the thalamus, when the other versus self contrast was applied.

### 3.3.4. Additional Neuroimaging Results

Additionally to group comparisons, 8 studies report other types of analyses, like correlations between BOLD response from source monitoring and clinical features, e.g. psychopathological symptoms or cognitive functions. Details of these results are presented in Table 2.

In the context of the clinical features, like psychopathological symptoms, one study [66] showed that severity of positive psychotic symptoms measured by PANSS positive symptoms subscale (r = −0.620) and the PSYRATS delusion items (r = −0.451) negatively correlated with the activation in the right MTG in the FEP group. Kumari and collaborators [60] conducted correlations between BOLD response during source monitoring task performance and PANSS scores. Results showed that within the SCH group, greater ventral striatal-hypothalamic activity during other versus self contrast correlated with higher negative symptoms PANSS score (r = 0.29). Another study [61] demonstrated that in the patient group there is a negative relationship between the degree of delusions and left hippocampal activity in the imagined condition (IC) (ϱ = 0.54). Moreover, left hippocampal activity in the viewed condition (VC) correlated significantly with delusions (ϱ = 0.53). Stephan-Otto and collaborators [67] showed that the patients with verbal hallucinations presented increased bilateral activation in the thalamus and in the precuneus relative to the other patients when correctly remembering pictures. On the other hand, when the patients with visual hallucinations were compared to the other patients, no differences in

brain activation was seen. The study by Subramaniam and colleagues [64], in which different mood states were induced, results showed that when HC were in a negative mood there was a negative relationship between the left dorsal region of the mPFC and externally-presented identification (r = -0.48) as well as overall reality-monitoring performance (r = -0.44). On the other hand, when SCH were in a negative mood versus neutral mood there was a positive relationship between signals in both left and right dorsal mPFC ROIs and better externally-presented identification (left dorsal region of the mPFC: r=0.54; right dorsal region of the mPFC: r=056), and overall reality-monitoring performance (left dorsal region of the mPFC: r=0.54; right dorsal region of the mPFC: r=0.56).

Subramaniam et al. [63] compared the mPFC activity with different cognitive exercises that they applied during the intervention. Results show a significant correlation between mPFC signal during reality monitoring task and verbal memory scores at 16 weeks training in SCH-AT (r = 0.51). Moreover, there was a significant correlation between the level of reality monitoring mPFC signal measured immediately after training and ratings of social functioning at the 6 month follow-up. Whereas, no significant relationship was found in mPFC activity and executive functions.

## 4. Discussion

In the current systematic review, we found eight fMRI studies of aberrant salience and eleven fMRI studies of source monitoring in schizophrenia and clinical risk of psychosis. As stated in the introduction, rarely cognitive biases are studied simultaneously, and the results of this review seem to confirm that. To our knowledge, there are no neuroimaging studies linking paradigms measuring aberrant salience and source monitoring. We aimed to review the neuronal correlates of aberrant salience and source monitoring in the spectrum of psychoses and elucidate specific and possible shared neuronal mechanisms of these cognitive biases. We also reviewed behavioural results obtained in the found studies.

### 4.1. Behavioral Results for Aberrant Salience

Results show that the most common paradigm was Salience Attribution Test, with five out of eight studies using it, followed by single studies using other types of tasks. Most of the studies researched schizophrenia spectrum or/and clinical risk of psychosis compared to a non-clinical control group. One study [35] studied schizophrenia participants with high and low positive symptoms. Regardless, obtained results were diversified, with only about one-third of the effects concerning aberrant salience, implicit or explicit, being reported as significant. These results may posit a question about reliable differences in aberrant salience between individuals on the psychosis continuum and healthy controls, as measured by behavioural tasks. Alternatively, perhaps, methods and conditions in which tests of aberrant salience were conducted could have influenced results obtained in the reviewed studies.

The three significant effects for implicit aberrant salience were varied, spanning from small (d = 0.39) (largest sample size, schizophrenia patients) [52], through medium (d = 0.68) [30] to large (d = 0.93) (longitudinal study in an UHR sample) [33]. The two effects for explicit aberrant salience were more consistent, with effects in the large range (d = 0.78 - 0.88) [32,33]. We estimated sample sizes required to obtain a significant and suitably powered effect in GPower [68]. We assumed effect sizes of d = 0.6 for implicit aberrant salience and d = 0.8 for explicit aberrant salience, $\alpha$ = 0.05 (two-tailed), power of $\beta$ = 0.95 (as avoiding false negatives at this point seems crucial, [cf. 65], and a design comparing two independent groups. These analyses indicated single group sample sizes of n = 74 and n = 42 (n = 45 and n = 26 for the often assumed value of $\beta$ = 0.80) for implicit and explicit aberrant salience, respectively. This would indicate that the majority of the fMRI studies of aberrant salience up to date were underpowered in the context of behavioural results and thus might have had problems with detecting effects of interest. Hence, future studies should consider employing larger samples.

Additionally, the effects of aberrant salience in the schizophrenia spectrum may be hard to observe, and studies on aberrant salience in schizophrenia patients may be hard to interpret. This happens in the context of dopamine antagonist medication used in the treatment of this disorder. Some studies point to the possibility that antipsychotic medication modulates reward and salience processing in the brain and that these effects may be heterogeneous depending on the type of medication [53,69]. In the current review, in most studies, schizophrenia participants were taking antipsychotic medications. Controlling for the effects of medication in future studies seems a viable methodological consideration. Another possibility is that fMRI setting somehow influenced tasks measuring aberrant salience, rendering them not as sensitive or specific as in a typical "desk" setting. A review of all behavioural studies of aberrant salience in schizophrenia and clinical risk of psychosis is warranted to discuss the reliability of aberrant salience tasks.

### *4.2. Neuroimaging Results for Aberrant Salience*

Reviewed studies show that neural processes related to aberrant and adaptive salience are linked to two major brain systems, primarily the subcortical dopamine system and the salience network. Though, it is important to notice that the insula, anterior cingulate cortex, and subcortical structures, amygdala, ventral striatum and ventral tegmental area are regarded as part of the salience network [70]. Therefore, there is a substantial overlap or even equality in the two discussed systems. Nevertheless, as parts of this network are often considered in relation to their specific functional mechanisms, we will discuss obtained results with a division into cortico-striatal-thalamic circuitry associated with dopamine signalling and salience network as composed of the insulae and anterior cingulate cortices.

The first of discussed systems is a cortico-striatal-thalamic circuitry responsible for reward and salience processing, adaptive and aberrant alike, especially striatum, hippocampus, amygdala and prefrontal cortex [54,71]. See Howes et al., [23] for a review of reward and salience processing studies in UHR participants. A meta-analysis of studies on reward processing in schizophrenia shows a general effect of hypoactivation in patients' ventral striatum [72]. This system is associated with motivational salience and its disruption [18]. The aberrant salience model assumes that the process of developing psychosis has a basis in the dysregulation of subcortical dopaminergic pathways in the brain [18,73,74].

In the current review, only two studies [26,33] showed direct effects of group comparisons in the mesolimbic pathway, both concerning UHR participants. One study [33] revealed diminished neural response in subcortical brain regions, ventral striatum and parahippocampal gyrus, and the other [26] showed an increased response in the ventral striatum. However, correlational results also found effects associated with these brain areas. Unfortunately, three of five studies reported such effects as correlations of peak voxels signal and behavioural results [32,33,56], rendering these results non-independent and hard to interpret in terms of actual effects [55]. In another study [52] low relevance weighted prediction error in the salience task was associated with high activation in the hippocampus, which was deemed not in line with previous results [75], where large prediction errors are associated with hippocampal response. This unexpected result was interpreted as an effect of higher-order processes regulating memory encoding when the prediction error is low and memorizing is adaptive [52]. Winton-Brown and colleagues [26] found that reward-induced connectivity between ventral striatum/pallidum and midbrain was significantly greater in UHR than in HC. Also, greater connectivity was correlated with unusual thought content in UHR. Authors interpreted these results as evidence for a disrupted circuitry, where enhanced hippocampal activity influences ventral striatum, pallidum, and midbrain functioning [76]. These findings tentatively suggest the engagement of disrupted ventral striatum and hippocampus processing in the aberrant salience tasks.

The other major brain system is the salience network, primarily associated with higher cortical structures, insulae and anterior cingulate cortex. The salience network is

seen as a network whose function is to detect and "sort" salient stimuli [77] and to regulate and switch between other, task-oriented and default mode networks [78]. It is implicated as the source of disrupted neural processing and symptoms of psychosis in the aberrant proximal salience concept [79]. On the symptoms level, disrupted salience network processing, along with its disrupted connectivity to task-positive and default mode networks, is associated with cognitive deficits, reality distortions [80] and hallucinations [81,82]. It is further hypothesized that the processing of the salience network is associated with aberrant attribution of salience and meaning to stimuli [81], which is a key process in aberrant salience cognitive bias. At the same time, it is associated with diminished capacity to differentiate between internal and external states [80,83], which is key in the source monitoring cognitive bias.

In the current review, Walter and colleagues [35] reported an increased bold response in the left insula in the schizophrenia group with high positive symptoms compared to the group with low positive symptoms. Additionally, response in the left insula correlated negatively with medication dose. Another study [53] found differences in bold response to adaptive salience processing in the anterior cingulate and right hemisphere insula cortices between FEP patients and healthy controls. Additionally, in FEP patients without medication, there was a negative correlation between right hemisphere insula activation during adaptive salience and hallucinations score. Reviewed studies showed differences in the salience network functioning of individuals with psychosis on the static contrast estimates of salience tasks. However, Smieskova and colleagues [53] and others point out that future research of salience network in psychosis may benefit from adopting a connectivity-based approach, allowing to study more intricate functional disruptions in this network [e.g. 80,81]. Current findings suggest the engagement of disrupted insula and anterior cingulate cortices processing in the aberrant salience tasks.

To sum up, the role of excessive release of dopamine in psychotic disorders is well documented [84,85], and associated with predictive processing account [23]. Notably, both these systems, the cortico-striatal-thalamic circuitry and the salience network, are modulated by dopamine functioning [23,79,86]. Reviewed studies indicate that individuals with psychotic disorders and at risk of psychosis have altered salience processing which is associated with disrupted functioning of the subcortical dopaminergic pathway and salience network in comparison to healthy controls [23,69]. However, the heterogeneity of the studied populations and methods, methodological issues, like reporting correlations for peak voxels, and most notably variability and inconsistency in results, means that more research is needed to replicate these findings reliably.

### 4.3. *Behavioral Results for Source Monitoring*

Studies considered in the current review show that the most common source monitoring paradigm is the external-internal paradigm, which was implemented in ten out of eleven studies. Only one study employed the external-external source monitoring paradigm [67]. In most of the studies, patients with schizophrenia spectrum disorder were compared to a non-clinical control group. Two papers [57,58] described a population of schizophrenia participants divided into groups with and without hallucinations. Despite the fact that most of the studies used an external-internal paradigm, the stimulus and task design usually differed. Nonetheless, a substantial part of the reviewed articles revealed significant differences between patients with SSD and non-clinical control groups in the source monitoring task performance. The obtained results are in line with the previous meta-analyses [3,40].

Ten studies that implemented the external-internal paradigm. Most of them reported significant differences in the overall source monitoring performance [57–60,63–65], where patients with schizophrenia made more misattributions than the control group. Significant effects for overall source monitoring task accuracy varied from medium (d = 0.66) [63] to large (d = 0.98) [60]. Thoresen et al. [61] showed that patients made more misattributions where they classified the imagined actions as viewed with a very large effect size (d = 1.30). Moreover, Vinogradov et al., [62] demonstrated that the schizophrenia group

had a significantly lower hit rate for self-generated items than the healthy controls (d = 3.04). However, in this case, the effect size should be interpreted with caution due to the considerably small sample size. In the study by Kambeitz-Ilankovic et al., [66], there was a significant difference between FEP and HC group in the self speech conditions while it was preceded by an invalid cue (picture of another person's face) with a large effect size (d = 0.8). On the other hand, one study [67] that implemented the external-external source monitoring paradigm has not found significant differences in the task performance between groups.

To sum up, for internal-external source monitoring paradigm results appear to be in the direction of previous meta-analyses [3,5,40], showing that patients with schizophrenia have a general tendency to commit more misattributions in the source monitoring tasks. Contrarily, the same effect has not been observed in the study [67] that investigated the external-external discrimination paradigm. Therefore, in the current review, the results for this type of discrimination are inconclusive since the considered paradigm is substantially underrepresented. Thus, more research is needed to add new insight into the topic. Moreover, most of the above studies considered patients with schizophrenia and one with the first episode of psychosis. Thus, more research is needed on ultra high risk of psychosis samples to exclude the potential impact of medication, impaired general cognitive functioning in schizophrenia patients, motivational aspects and other factors that could potentially influence performance in all sorts of different behavioural tasks in patients with schizophrenia [87,88].

The current review provides additional evidence to already existing knowledge on source monitoring characteristics across psychotic patients. In future studies, direct comparison in source monitoring performance between schizophrenia patients and participants without full-blown psychotic symptoms is needed, as it would widen the knowledge on source monitoring deficits across the continuum for psychosis.

### 4.4. *Neuroimaging Results for Source Monitoring*

Reviewed studies used a variety of contrasts in the neuroimaging analyses. Some of the studies concentrated on the effects concerning the difference between reality monitoring and a control task [59], other on more specific effects associated with a type of memory or stimuli source [41,61], and there were also studies which searched for effects associated with misidentifying memory source [67]. This heterogeneity in sought out effects may translate into the variability of obtained brain activations. Nevertheless, we were able to identify some of the neural effects that were shared across different studies, like the medial prefrontal cortex (mPFC), hippocampus, middle and superior temporal gyri and anterior cingulate cortex (ACC).

The brain structure most often identified in the found studies was the medial prefrontal cortex [59,60,62–65]. In the majority of cases, this part of the prefrontal lobes was under investigation as a region of interest, as it was implicated in the studies of self-referential and memory processing [62,89] and reality monitoring in healthy subjects (for review on reality monitoring see: [90]). Most of the studies point to diminished activity in the mPFC in patients during contrasts associated with source monitoring, except Kumari and colleagues [60] study, where patients with worse results in the verbal monitoring task had higher activations in mPFC than patients with better results. Wang and colleagues [65] used mPFC as a seed in a connectivity analysis. They found an increased effect between the medial frontal cortex and superior temporal gyrus in individuals with schizophrenia during self-monitoring condition relative to controls. Authors interpreted this result as a brain circuitry abnormality that contributes to the self–other confusion, which later manifests as positive symptoms. Taken together, findings from these studies suggest that diminished activity of the mPFC is associated with the disrupted source monitoring, facilitating confusion in attributing the source of perceptions or memories from the self to others and contributing to the "externalization bias" [10].

The current review also implicated the hippocampus in the source monitoring deficits. The study by Thorensen and colleagues [61] found diminished activity in this region

during source monitoring of imagined stimuli. Additionally, diminished hippocampal activity was associated with higher levels of delusions. Kumari and colleagues [60] found greater response in the hippocampus during source monitoring of self-generated stimuli in the patient's group with better results in this task. Medial parts of the temporal lobes are associated with memory encoding and retrieval [91,92], which is a crucial aspect of the source monitoring processes. For an extensive review of the role of the hippocampus in the source monitoring, see: [19]. One of the recent findings points out that mPFC modulates the hippocampus's functioning, minimising interference that would otherwise impede discrimination between similar memory representations [93]. This could suggest an impaired top-down circuitry in schizophrenia, where disruption of mPFC affects the functioning of the hippocampus during memory tasks.

Several of the reviewed studies also found effects in the superior and middle temporal gyri. This was expected in the context of the majority of the found studies using speech-based stimuli. Effects of source monitoring were primarily associated with diminished response in these parts of the temporal cortex [41,58,60,66]. Also, Wang and colleagues [65] and Mechelli and colleagues [58] found a disrupted connectivity pattern between the superior temporal gyrus and the mPFC and ACC, respectively. These temporal regions are associated with sensory processing and crossmodal integration of stimuli, but also self-referential processing and disrcrimination of source of speech [41,58,60,66]. A similar function of reality monitoring is associated with the paracingulate sulcus. This tertiary structure is morphologically diverse in the human population and located between the mPFC and the ACC [94,95]. In the current review, apart from studies concentrating on the mPFC, three studies found significant effects in the proximal vicinity of the paracingulate gyrus, namely in the anterior cingulate cortex [57,58,62]. However, these results are heterogeneous, showing diminished ACC activity in the patients in one study [57] and increased activity in another [62] during the processing of self-generated stimuli. Additionally, Mechelli and colleagues reported diminished connectivity between the ACC and the superior temporal gyrus in hallucinating patients. Nonetheless, these results suggest that structures placed in the broad anterior cingulum area of the brain are functionally relevant to the source monitoring bias [90].

Noteworthy, several studies reported effects localized in the basal ganglia, i.e. the striatum and the putamen [60,62,64]. These structures were not extensively discussed in the literature on source monitoring. Subramaniam and colleagues [64] noted that perhaps in patients with schizophrenia there is a disruption in a dopaminergic frontostriatal pathway associated with motivational and goal-directed behaviour. It may be suggested that functional abnormalities of the basal ganglia, especially striatum, is a reflection of disrupted predictive processing during the source monitoring. This would coincide with aberrant salience account, where striatum is modulated by the excessive hippocampus activity, but further studies are warranted to elaborate on the role of basal ganglia in disruptions of source monitoring processes.

To sum up, schizophrenia patients show disruptions in several brain regions during source monitoring tasks. Obtained results are in line with proposed neural correlates of reality monitoring in the context of the mPFC, medial temporal lobe, and paracingulate gyrus and the ACC [90]. Also, significant between-group differences during source monitoring were found for middle and superior temporal gyri. Notably, these effects are also consistent with neural correlates of hallucinations in schizophrenia, underlying the role of disrupted source monitoring in positive symptoms of psychosis [96].

### 4.5. *Neuronal Overlaps*

The review of the found studies pointed to two key brain regions implicated in both aberrant salience and source monitoring biases. The first region is the hippocampus, which is associated with memory processing [97] but also the regulation of the dopamine functioning in the mesolimbic pathways [76,98]. The other one was the anterior cingulate cortex, which is a key structure in the brain salience network, responsible for regulating internal states and self-referential processing [70,99]. Figure 1 demonstrates the most

common neuronal differences between patients and controls in aberrant salience and source monitoring biases. Additionally, it marks overlaps and specific effects for aberrant salience and source monitoring.
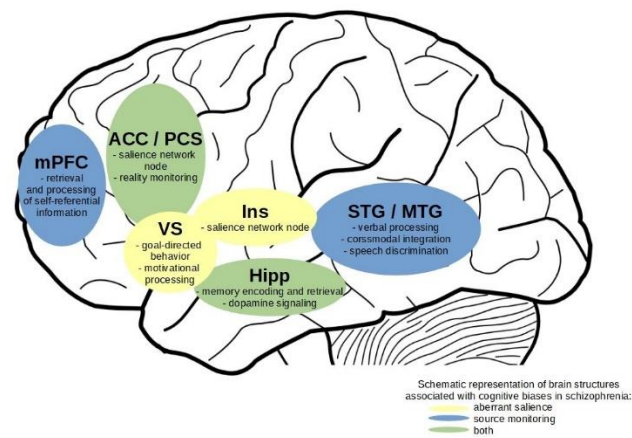


**Figure 1.** Schematic representation of brain regions with effects detected between schizophrenia patients and non-clinical controls in fMRI studies of aberrant salience and source monitoring. For display purposes effects are presented only on the left hemisphere of the brain. mPFC - medial prefrontal cortex, ACC / PCS - anterior cingulate cortex and paracingulate sulcus, VS - central striatum, Ins - insula, Hipp - hippocampus, STG / MTG - superior and middle temporal gyri**.**

Evidence points to the hippocampus as the structure responsible for regulating dopamine functioning [100–102]. Metabolism of this neurotransmitter is disrupted in schizophrenia [85,103] and states of clinical risk of psychosis [23,104] and is associated with positive symptoms of psychosis [76]. Also, studies are pointing to the structural and functional disruption in the hippocampus in schizophrenia and the risk of psychosis [102,105]. States of psychosis risk and transition to psychosis are associated with structural and functional abnormalities in the medial temporal cortex, i.e. the hippocampus [106–108]. Also, schizophrenia is associated with abnormality in the hippocampal function [105,109]. Study on schizophrenia patients' brain response to the dopamine-antagonist medication showed that it affected hippocampus and ventral striatum in the early stages of treatment and that normalization of hippocampal function was associated with a better response to treatment [110]. Lodge and Grace [111] described the hippocampus as a focal point of influence of various risk factors, like substance abuse, genetic risk or psychological stress, contributing to the development of psychotic symptoms.

Studies in the current review are in line with these results, showing disruption of hippocampus functioning in patients with schizophrenia and risk of psychosis when exhibiting aberrant salience and disrupted source monitoring. However, in both these cognitive biases, the hippocampus is seen as a part of two partially distinct circuitries associated with different cognitive processes. In aberrant salience it is seen as a regulator of dopaminergic neurons in the ventral striatum and the midbrain [26,32,33,76]. In the source monitoring it is seen as a part of circuitry engaged in recollection and source identification of memories, where prefrontal regions, associated with self-referential processing and cognitive control, have the regulating role [19,61,90]. There is also light to be shed on the relationship between the hippocampus and the basal ganglia during source monitoring. To sum up, despite its diverse functions, the hippocampus seems to be a key brain structure in aberrant salience and disrupted source monitoring in individuals with schizophrenia and at risk of psychosis.

Another structure associated with both cognitive biases was the anterior cingulate cortex. Studies show that disruption of function and structure of the anterior cingulate cortex is associated with psychosis [112–115]. The high risk of psychosis and transmission to psychosis is associated with reduced grey matter in the anterior cingulate cortex [106]. A study on patients' response to dopamine-antagonist treatment showed an effect in the anterior cingulate and medial prefrontal cortex [110].

The current review is in line with studies linking the ACC with psychosis. We demonstrated that both aberrant salience and source monitoring are linked to abnormal functioning of this brain structure. However, in the aberrant salience studies, it is interpreted in the context of disrupted processing of the brain salience network [52,79]. In this approach, the salience network is responsible for regulating cognitive processing and reallocation of attention, switching between task-positive and task-negative brain networks. In the source monitoring studies, observed disruptions in the ACC functioning are identified as associated with the dysfunction of the paracingulate sulcus, primarily as it seems due to the anatomical proximity and relationship [94]. This argument appears to be pervasive due to the association of paracingulate sulcus, reality monitoring [90,95] and hallucinations [116,117].

To sum up, both brain regions identified in this review as associated with both aberrant salience and disrupted source monitoring are discussed in their respective bodies of literature in the context of various neural mechanisms, corresponding to their cognitive and phenomenological mechanisms and character. This apparent dissociation calls for an attempt in bridging these two cognitive biases into an account that will encompass their characteristics and enable future studies on their possible joint neural mechanisms.

### 4.6. *Integration of Results on Aberrant Salience and Source Monitoring*

A number of studies and reviews discussed the possible integration of aberrant salience and source monitoring [42,44,45,118]. However, our study is the first to systematically review neural correlates of aberrant salience and source monitoring in functional MRI studies. Our findings seem to be in line with these previous accounts. Observed neural differences between individuals with schizophrenia or the risk of psychosis during source monitoring and aberrant salience may indicate a disturbance in hierarchical predictive processing and disturbances in minimal self [42,45].

Obtained overlaps between neural correlates of these cognitive biases suggest that they may function in a feedback loop. In such an account, aberrant salience and disrupted source monitoring lead to an experience of externalization and emergent experience of aberrant attributional processes of cognitive functions and one's own experiences as originating externally. Aberrant salience is associated with strong priors or underweighted prediction errors [23]. They lead to erroneous prediction updates, which in turn may lead, on the next hierarchical level of predictive processing, to disrupted priors in processes of source monitoring and attribution [42]. On a phenomenological level, aberrant salience translates to a sense of heightened yet erratic motivational and attentional functioning. A state of perceptual and self-referential catch-as-catch-can and disrupted attentional control [46]. As Mitchell and Johnson [19] noted in their account of cognitive mechanisms of source monitoring, memory is selective and contextual, influenced by meta-memory, knowledge, schemas, biases and temporarily elicited states. In a state of aberrant salience, inferences about the origin of perceptual and cognitive phenomena are burdened simultaneously with erroneous predictions and with disrupted self-referential processing and stimuli discrimination, leading to a disturbed sense of ownership of cognitive processes and agency. In this approach state of aberrant salience would be deemed a context-modifier for more meta-level processes of source monitoring. Nevertheless, their interaction would be necessary to produce an external attribution of internal phenomena, like sensory experiences or thoughts. This attribution, resulting in a qualitative change of how reality is perceived, could lead in turn to positive symptoms of psychosis [119,120].

However, as we assume a hierarchical type precedence of aberrant salience over source monitoring deficits [42], we cannot certainly presuppose temporal precedence of these two cognitive biases. Indeed, there are studies indicating the stronger predictive potential of source monitoring deficits in the early psychosis' self-disturbances than that of aberrant salience [45,121]. Our assumption is supported by neural correlates of researched cognitive biases, where source monitoring is associated more with higher-order cortical structures. Nonetheless, an account where a state of aberrant salience occurs earlier and

contributes to the source monitoring disruptions warrants further investigation, especially with the use of neuroimaging studies.

## 5. Conclusions

To sum up, we conducted a systematic review of fMRI studies on source monitoring and aberrant salience. There are several areas of concern or questions that could be addressed by future research. The first one is the reliability of available fMRI studies of aberrant salience, where the majority of expected effects were non-significant. A systematic review of behavioural tasks of aberrant salience is warranted, and future neurocognitive studies of this cognitive bias should employ larger sample sizes. Another question is about neural correlates of source monitoring in participants with a high clinical risk of psychosis. To our knowledge, there are no fMRI studies that would elucidate neural correlates of disrupted source monitoring in such a sample.

Furthermore, we found that several studies of source monitoring found effects associated with basal ganglia, yet the role of this structure was not extensively discussed in the previous literature. Future accounts of source monitoring could explore how disrupted motivational and salience processing adds to disrupted source monitoring. This could also be an incentive to explore the role of predictive processing in the source monitoring accounts. Additionally, we found overlap in the neural correlates of aberrant salience and source monitoring, namely the anterior cingulate cortex. However, effects found in this structure in the source monitoring accounts are associated with morphological diversity of the paracingulate sulcus, which is responsible for self-monitoring functions. Future research could shed additional light on how specific structures in the anterior cingulate area contribute to aberrant salience and source monitoring.

Finally, we reviewed and discussed neural correlates of aberrant salience and source monitoring in patients with schizophrenia or at clinical risk of psychosis. We point to several regions associated with aberrant salience - ventral striatum and the insula cortex - and with source monitoring - the medial prefrontal cortex and superior and middle temporal gyri. There was a neural overlap in the anterior cingulate cortex and paracingulate sulcus and hippocampus. Our review indicates that aberrant salience and source monitoring may share neural mechanisms, suggesting their joint role in producing disrupted external attributions of perceptual and cognitive experiences, thus elucidating their role in positive symptoms of psychosis. Future studies, both behavioural and fMRI, may benefit from investigating source monitoring and aberrant salience simultaneously. This may provide empirical verification of similarities and overlaps of these two important cognitive biases that contribute to psychosis.

**Institutional Review Board Statement:** "Not applicable."

**Informed Consent Statement:** NA

**Data Availability Statement:** In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Please refer to suggested Data Availability Statements in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics. You might choose to exclude this statement if the study did not report any data.

## Appendix A

The appendix is an optional section that can contain details and data supplemental to the main text—for example, explanations of experimental details that would disrupt the flow of the main text but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data is shown in the main text can be added here if brief, or as Supplementary data. Mathematical proofs of results not central to the paper can be added as an appendix.

## Appendix B

All appendix sections must be cited in the main text. In the appendices, Figures, Tables, etc. should be labeled starting with "A"—e.g., Figure A1, Figure A2, etc.

## References

1.  McLean, B.F.; Mattiske, J.K.; Balzan, R.P. Association of the Jumping to Conclusions and Evidence Integration Biases with Delusions in Psychosis: A Detailed Meta-Analysis. Schizophr. Bull. 2017, 43, 344–354.
2.  Livet, A.; Navarri, X.; Potvin, S.; Conrod, P. Cognitive Biases in Individuals with Psychotic-like Experiences: A Systematic Review and a Meta-Analysis. Schizophr. Res. 2020, 222, 10–22.
3.  Lavallé, L.; Dondé, C.; Gawęda, Ł.; Brunelin, J.; Mondino, M. Impaired Self-Recognition in Individuals with No Full-Blown Psychotic Symptoms Represented across the Continuum of Psychosis: A Meta-Analysis. Psychol. Med. 2020, 1–11.
4.  Dudley, R.; Taylor, P.; Wickham, S. Psychosis, Delusions and the "jumping to Conclusions" Reasoning Bias: A Systematic Review and Meta-Analysis. Schizophrenia 2016.
5.  Brookwell, M.L.; Bentall, R.P.; Varese, F. Externalizing Biases and Hallucinations in Source-Monitoring, Self-Monitoring and Signal Detection Studies: A Meta-Analytic Review. Psychol. Med. 2013, 43, 2465–2475.
6.  Sauvé, G.; Lavigne, K.M.; Pochiet, G.; Brodeur, M.B.; Lepage, M. Efficacy of Psychological Interventions Targeting Cognitive Biases in Schizophrenia: A Systematic Review and Meta-Analysis. Clin. Psychol. Rev. 2020, 78, 101854.
7.  Moritz, S.; Veckenstedt, R.; Hottenrott, B.; Woodward, T.S.; Randjbar, S.; Lincoln, T.M. Different Sides of the Same Coin? Intercorrelations of Cognitive Biases in Schizophrenia. Cogn. Neuropsychiatry 2010, 15, 406–421.
8.  Gawęda, L.; Woodward, T.S.; Moritz, S.; Kokoszka, A. Impaired Action Self-Monitoring in Schizophrenia Patients with Auditory Hallucinations. Schizophr. Res. 2013, 144, 72–79.
9.  Ochoa, S.; Haro, J.M.; Huerta-Ramos, E.; Cuevas-Esteban, J.; Stephan-Otto, C.; Usall, J.; Nieto, L.; Brebion, G. Relation between Jumping to Conclusions and Cognitive Functioning in People with Schizophrenia in Contrast with Healthy Participants. Schizophr. Res. 2014, 159, 211–217.
10. Woodward, T.S.; Menon, M.; Whitman, J.C. Source Monitoring Biases and Auditory Hallucinations. Cogn. Neuropsychiatry 2007, 12, 477–494.
11. Moritz, S.; Woodward, T.S. Jumping to Conclusions in Delusional and Non-Delusional Schizophrenic Patients. Br. J. Clin. Psychol. 2005, 44, 193–207.
12. Gawęda, Ł.; Staszkiewicz, M.; Balzan, R.P. The Relationship between Cognitive Biases and Psychological Dimensions of Delusions: The Importance of Jumping to Conclusions. J. Behav. Ther. Exp. Psychiatry 2017, 56, 51–56.
13. So, S.H.-W.; Chau, A.K.C.; Peters, E.R.; Swendsen, J.; Garety, P.A.; Kapur, S. Moment-to-Moment Associations between Negative Affect, Aberrant Salience, and Paranoia. Cogn. Neuropsychiatry 2018, 23, 299–306.
14. Larøi, F.; Woodward, T.S. Hallucinations from a Cognitive Perspective. Harv. Rev. Psychiatry 2007, 15, 109–117.
15. Beck, A.T.; Rector, N.A. A Cognitive Model of Hallucinations. Cognit. Ther. Res. 2003, 27, 19–52.
16. Frydecka, D.; Misiak, B.; Kotowicz, K.; Pionke, R.; Krężołek, M.; Cechnicki, A.; Gawęda, Ł. The Interplay between Childhood Trauma, Cognitive Biases, and Cannabis Use on the Risk of Psychosis in Nonclinical Young Adults in Poland. Eur. Psychiatry 2020, 63, e35.
17. Gawęda, Ł.; Göritz, A.S.; Moritz, S. Mediating Role of Aberrant Salience and Self-Disturbances for the Relationship between Childhood Trauma and Psychotic-like Experiences in the General Population. Schizophr. Res. 2019, 206, 149–156.
18. Kapur, S. Psychosis as a State of Aberrant Salience: A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia. Am. J. Psychiatry 2003, 160, 13–23.
19. Mitchell, K.J.; Johnson, M.K. Source Monitoring 15 Years Later: What Have We Learned from fMRI about the Neural Mechanisms of Source Memory? Psychol. Bull. 2009, 135, 638–677.

20. Insel, T.; Cuthbert, B.; Garvey, M.; Heinssen, R.; Pine, D.S.; Quinn, K.; Sanislow, C.; Wang, P. Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. Am. J. Psychiatry 2010, 167, 748–751.

21. Waters, F.; Allen, P.; Aleman, A.; Fernyhough, C.; Woodward, T.S.; Badcock, J.C.; Barkus, E.; Johns, L.; Varese, F.; Menon, M.; et al. Auditory Hallucinations in Schizophrenia and Nonschizophrenia Populations: A Review and Integrated Model of Cognitive Mechanisms. Schizophr. Bull. 2012, 38, 683–693.

22. Nelson, B.; Whitford, T.J.; Lavoie, S.; Sass, L.A. What Are the Neurocognitive Correlates of Basic Self-Disturbance in Schizophrenia? Integrating Phenomenology and Neurocognition: Part 2 (aberrant Salience). Schizophr. Res. 2014, 152, 20–27.

23. Howes, O.D.; Hird, E.J.; Adams, R.A.; Corlett, P.R.; McGuire, P. Aberrant Salience, Information Processing, and Dopaminergic Signaling in People at Clinical High Risk for Psychosis. Biol. Psychiatry 2020, 88, 304–314.

24. Corlett, P.R.; Taylor, J.R.; Wang, X.-J.; Fletcher, P.C.; Krystal, J.H. Toward a Neurobiology of Delusions. Prog. Neurobiol. 2010, 92, 345–369.

25. Corlett, P.R.; Horga, G.; Fletcher, P.C.; Alderson-Day, B.; Schmack, K.; Powers, A.R. Hallucinations and Strong Priors. Trends Cogn. Sci. 2019, 23, 114–127.

26. Winton-Brown, T.; Schmidt, A.; Roiser, J.P.; Howes, O.D.; Egerton, A.; Fusar-Poli, P.; Bunzeck, N.; Grace, A.A.; Duzel, E.; Kapur, S.; et al. Altered Activation and Connectivity in a Hippocampal--Basal Ganglia--Midbrain Circuit during Salience Processing in Subjects at Ultra High Risk for Psychosis. Transl. Psychiatry 2017, 7, e1245–e1245.

27. Cicero, D.C.; Becker, T.M.; Martin, E.A.; Docherty, A.R.; Kerns, J.G. The Role of Aberrant Salience and Self-Concept Clarity in Psychotic-like Experiences. Personal. Disord. 2013, 4, 33–42.

28. Pelizza, L.; Azzali, S.; Garlassi, S.; Scazza, I.; Paterlini, F.; Chiri, L.R.; Poletti, M.; Pupo, S.; Cicero, D.C.; Preti, A.; et al. Assessing Aberrant Salience in Young Community Help-Seekers with Early Psychosis: The Approved Italian Version of the Aberrant Salience Inventory. J. Clin. Psychol. 2021, 77, 782–803.

29. Roiser, J.P.; Stephan, K.E.; den Ouden, H.E.M.; Barnes, T.R.E.; Friston, K.J.; Joyce, E.M. Do Patients with Schizophrenia Exhibit Aberrant Salience? Psychol. Med. 2009, 39, 199–209.

30. Pankow, A.; Katthagen, T.; Diner, S.; Deserno, L.; Boehme, R.; Kathmann, N.; Gleich, T.; Gaebler, M.; Walter, H.; Heinz, A.; et al. Aberrant Salience Is Related to Dysfunctional Self-Referential Processing in Psychosis. Schizophr. Bull. 2016, 42, 67–76.

31. Katthagen, T.; Dammering, F.; Kathmann, N.; Kaminski, J.; Walter, H.; Heinz, A.; Schlagenhauf, F. Validating the Construct of Aberrant Salience in Schizophrenia - Behavioral Evidence for an Automatic Process. Schizophr. Res. Cogn. 2016, 6, 22–27.

32. Roiser, J.P.; Howes, O.D.; Chaddock, C.A.; Joyce, E.M.; McGuire, P. Neural and Behavioral Correlates of Aberrant Salience in Individuals at Risk for Psychosis. Schizophr. Bull. 2013, 39, 1328–1336.

33. Schmidt, A.; Antoniades, M.; Allen, P.; Egerton, A.; Chaddock, C.A.; Borgwardt, S.; Fusar-Poli, P.; Roiser, J.P.; Howes, O.; McGuire, P. Longitudinal Alterations in Motivational Salience Processing in Ultra-High-Risk Subjects for Psychosis. Psychol. Med. 2017, 47, 243–254.

34. Tessier, C.; Sweers, K.; Frajerman, A.; Bergaoui, H.; Ferreri, F.; Delva, C.; Lapidus, N.; Lamaziere, A.; Roiser, J.P.; De Hert, M.; et al. Membrane Lipidomics in Schizophrenia Patients: A Correlational Study with Clinical and Cognitive Manifestations. Transl. Psychiatry 2016, 6, e906.

35. Walter, A.; Suenderhauf, C.; Smieskova, R.; Lenz, C.; Harrisberger, F.; Schmidt, A.; Vogel, T.; Lang, U.E.; Riecher-Rössler, A.; Eckert, A.; et al. Altered Insular Function during Aberrant Salience Processing in Relation to the Severity of Psychotic Symptoms. Front. Psychiatry 2016, 7, 189.

36. Spark, J.; Gawęda, Ł.; Allott, K.; Hartmann, J.A.; Jack, B.N.; Koren, D.; Lavoie, S.; Li, E.; McGorry, P.D.; Parnas, J.; et al. Distinguishing Schizophrenia Spectrum from Non-Spectrum Disorders among Young Patients with First Episode Psychosis and at High Clinical Risk: The Role of Basic Self-Disturbance and Neurocognition. Schizophr. Res. 2021, 228, 19–28.

37. Johnson, M.K.; Raye, C.L. Reality Monitoring. Psychol. Rev. 1981, 88, 67.

38. Johnson, M.K.; Hashtroudi, S.; Lindsay, D.S. Source Monitoring. Psychol. Bull. 1993, 114, 3–28.

39. Bentall, R.P.; Baker, G.A.; Havers, S. Reality Monitoring and Psychotic Hallucinations. Br. J. Clin. Psychol. 1991, 30, 213–222.

40. Waters, F.; Woodward, T.; Allen, P.; Aleman, A.; Sommer, I. Self-Recognition Deficits in Schizophrenia Patients with Auditory Hallucinations: A Meta-Analysis of the Literature. Schizophr. Bull. 2012, 38, 741–750.

41. Allen, P.; Aleman, A.; McGuire, P.K. Inner Speech Models of Auditory Verbal Hallucinations: Evidence from Behavioural and Neuroimaging Studies. Int. Rev. Psychiatry 2007, 19, 407–415.

42. Griffin, J.D.; Fletcher, P.C. Predictive Processing, Source Monitoring, and Psychosis. Annu. Rev. Clin. Psychol. 2017, 13, 265–289.

43. Fletcher, P.C.; Frith, C.D. Perceiving Is Believing: A Bayesian Approach to Explaining the Positive Symptoms of Schizophrenia. Nat. Rev. Neurosci. 2009, 10, 48–58.

44. Nelson, B.; Whitford, T.J.; Lavoie, S.; Sass, L.A. What Are the Neurocognitive Correlates of Basic Self-Disturbance in Schizophrenia?: Integrating Phenomenology and Neurocognition. Part 1 (Source Monitoring Deficits). Schizophr. Res. 2014, 152, 12–19.

45. Nelson, B.; Lavoie, S.; Gawęda, Ł.; Li, E.; Sass, L.A.; Koren, D.; McGorry, P.D.; Jack, B.N.; Parnas, J.; Polari, A.; et al. The Neurophenomenology of Early Psychosis: An Integrative Empirical Study. Conscious. Cogn. 2020, 77, 102845.

46. Nelson, B.; Sass, L.A. Towards Integrating Phenomenology and Neurocognition: Possible Neurocognitive Correlates of Basic Self-Disturbance in Schizophrenia. Current Problems of Psychiatry 2017, 18.

47. Juckel, G.; Friedel, E.; Koslowski, M.; Witthaus, H.; Ozgürdal, S.; Gudlowski, Y.; Knutson, B.; Wrase, J.; Brüne, M.; Heinz, A.; et al. Ventral Striatal Activation during Reward Processing in Subjects with Ultra-High Risk for Schizophrenia. Neuropsychobiology 2012, 66, 50–56.

48. Wotruba, D.; Heekeren, K.; Michels, L.; Buechler, R.; Simon, J.J.; Theodoridou, A.; Kollias, S.; Rössler, W.; Kaiser, S. Symptom Dimensions Are Associated with Reward Processing in Unmedicated Persons at Risk for Psychosis. Front. Behav. Neurosci. 2014, 8, 382.

49. Shergill, S.S.; Bullmore, E.; Simmons, A.; Murray, R.; McGuire, P. Functional Anatomy of Auditory Verbal Imagery in Schizophrenic Patients with Auditory Hallucinations. Am. J. Psychiatry 2000, 157, 1691–1693.

50. Simons, C.J.P.; Tracy, D.K.; Sanghera, K.K.; O'Daly, O.; Gilleen, J.; Dominguez, M.-G.; Krabbendam, L.; Shergill, S.S. Functional Magnetic Resonance Imaging of Inner Speech in Schizophrenia. Biol. Psychiatry 2010, 67, 232–237.

51. Sawilowsky, S.S. New Effect Size Rules of Thumb. J. Mod. Appl. Stat. Methods 2009, 8, 597–599.

52. Katthagen, T.; Mathys, C.; Deserno, L.; Walter, H.; Kathmann, N.; Heinz, A.; Schlagenhauf, F. Modeling Subjective Relevance in Schizophrenia and Its Relation to Aberrant Salience. PLoS Comput. Biol. 2018, 14, e1006319.

53. Smieskova, R.; Roiser, J.P.; Chaddock, C.A.; Schmidt, A.; Harrisberger, F.; Bendfeldt, K.; Simon, A.; Walter, A.; Fusar-Poli, P.; McGuire, P.K.; et al. Modulation of Motivational Salience Processing during the Early Stages of Psychosis. Schizophr. Res. 2015, 166, 17–23.

54. Roiser, J.P.; Stephan, K.E.; den Ouden, H.E.M.; Friston, K.J.; Joyce, E.M. Adaptive and Aberrant Reward Prediction Signals in the Human Brain. Neuroimage 2010, 50, 657–664.

55. Vul, E.; Harris, C.; Winkielman, P.; Pashler, H. Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. Perspect. Psychol. Sci. 2009, 4, 274–290.

56. Esslinger, C.; Englisch, S.; Inta, D.; Rausch, F.; Schirmbeck, F.; Mier, D.; Kirsch, P.; Meyer-Lindenberg, A.; Zink, M. Ventral Striatal Activation during Attribution of Stimulus Saliency and Reward Anticipation Is Correlated in Unmedicated First Episode Schizophrenia Patients. Schizophr. Res. 2012, 140, 114–121.

57. Allen, P.; Amaro, E.; Fu, C.H.Y.; Williams, S.C.R.; Brammer, M.J.; Johns, L.C.; McGuire, P.K. Neural Correlates of the Misattribution of Speech in Schizophrenia. Br. J. Psychiatry 2007, 190, 162–169.

58. Mechelli, A.; Allen, P.; Amaro, E., Jr; Fu, C.H.Y.; Williams, S.C.R.; Brammer, M.J.; Johns, L.C.; McGuire, P.K. Misattribution of Speech and Impaired Connectivity in Patients with Auditory Verbal Hallucinations. Hum. Brain Mapp. 2007, 28, 1213–1222.

59. Garrison, J.R.; Fernandez-Egea, E.; Zaman, R.; Agius, M.; Simons, J.S. Reality Monitoring Impairment in Schizophrenia Reflects Specific Prefrontal Cortex Dysfunction. Neuroimage Clin 2017, 14, 260–268.

60. Kumari, V.; Fannon, D.; Ffytche, D.H.; Raveendran, V.; Antonova, E.; Premkumar, P.; Cooke, M.A.; Anilkumar, A.P.P.; Williams, S.C.R.; Andrew, C.; et al. Functional MRI of Verbal Self-Monitoring in Schizophrenia: Performance and Illness-Specific Effects. Schizophr. Bull. 2010, 36, 740–755.

61. Thoresen, C.; Endestad, T.; Sigvartsen, N.P.B.; Server, A.; Bolstad, I.; Johansson, M.; Andreassen, O.A.; Jensen, J. Frontotemporal Hypoactivity during a Reality Monitoring Paradigm Is Associated with Delusions in Patients with Schizophrenia Spectrum Disorders. Cogn. Neuropsychiatry 2014, 19, 97–115.

62. Vinogradov, S.; Luks, T.L.; Schulman, B.J.; Simpson, G.V. Deficit in a Neural Correlate of Reality Monitoring in Schizophrenia Patients. Cereb. Cortex 2008, 18, 2532–2539.

63. Subramaniam, K.; Luks, T.L.; Fisher, M.; Simpson, G.V.; Nagarajan, S.; Vinogradov, S. Computerized Cognitive Training Restores Neural Activity within the Reality Monitoring Network in Schizophrenia. Neuron 2012, 73, 842–853.

64. Subramaniam, K.; Ranasinghe, K.G.; Mathalon, D.; Nagarajan, S.; Vinogradov, S. Neural Mechanisms of Mood-Induced Modulation of Reality Monitoring in Schizophrenia. Cortex 2017, 91, 271–286.

65. Wang, L.; Metzak, P.D.; Woodward, T.S. Aberrant Connectivity during Self–other Source Monitoring in Schizophrenia. Schizophr. Res. 2011, 125, 136–142.

66. Kambeitz-Ilankovic, L.; Hennig-Fast, K.; Benetti, S.; Kambeitz, J.; Pettersson-Yeo, W.; O'Daly, O.; McGuire, P.; Allen, P. Attentional Modulation of Source Attribution in First-Episode Psychosis: A Functional Magnetic Resonance Imaging Study. Schizophr. Bull. 2013, 39, 1027–1036.

67.  Stephan-Otto, C.; Siddi, S.; Senior, C.; Cuevas-Esteban, J.; Cambra-Martí, M.R.; Ochoa, S.; Brébion, G. Remembering Verbally-Presented Items as Pictures: Brain Activity Underlying Visual Mental Images in Schizophrenia Patients with Visual Hallucinations. Cortex 2017, 94, 113–122.

68.  Faul, F.; Erdfelder, E.; Lang, A.-G.; Buchner, A. G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. Behav. Res. Methods 2007, 39, 175–191.

69.  Winton-Brown, T.T.; Fusar-Poli, P.; Ungless, M.A.; Howes, O.D. Dopaminergic Basis of Salience Dysregulation in Psychosis. Trends Neurosci. 2014, 37, 85–94.

70.  Toga, A.W. Brain Mapping: An Encyclopedic Reference; Academic Press, 2015;.

71.  Grace, A.A.; Floresco, S.B.; Goto, Y.; Lodge, D.J. Regulation of Firing of Dopaminergic Neurons and Control of Goal-Directed Behaviors. Trends Neurosci. 2007, 30, 220–227.

72.  Radua, J.; Schmidt, A.; Borgwardt, S.; Heinz, A.; Schlagenhauf, F.; McGuire, P.; Fusar-Poli, P. Ventral Striatal Activation During Reward Processing in Psychosis: A Neurofunctional Meta-Analysis. JAMA Psychiatry 2015, 72, 1243–1251.

73.  Howes, O.D.; Kapur, S. The Dopamine Hypothesis of Schizophrenia: Version III—the Final Common Pathway. Schizophr. Bull. 2009, 35, 549–562.

74.  Heinz, A.; Schlagenhauf, F. Dopaminergic Dysfunction in Schizophrenia: Salience Attribution Revisited. Schizophr. Bull. 2010, 36, 472–485.

75.  Fernández, R.S.; Boccia, M.M.; Pedreira, M.E. The Fate of Memory: Reconsolidation and the Case of Prediction Error. Neurosci. Biobehav. Rev. 2016, 68, 423–441.

76.  Lodge, D.J.; Grace, A.A. Hippocampal Dysregulation of Dopamine System Function and the Pathophysiology of Schizophrenia. Trends Pharmacol. Sci. 2011, 32, 507–513.

77.  Menon, V.; Uddin, L.Q. Saliency, Switching, Attention and Control: A Network Model of Insula Function. Brain Struct. Funct. 2010, 214, 655–667.

78.  Goulden, N.; Khusnulina, A.; Davis, N.J.; Bracewell, R.M.; Bokde, A.L.; McNulty, J.P.; Mullins, P.G. The Salience Network Is Responsible for Switching between the Default Mode Network and the Central Executive Network: Replication from DCM. Neuroimage 2014, 99, 180–190.

79.  Palaniyappan, L.; Liddle, P.F. Does the Salience Network Play a Cardinal Role in Psychosis? An Emerging Hypothesis of Insular Dysfunction. J. Psychiatry Neurosci. 2012, 37, 17–27.

80.  Wotruba, D.; Michels, L.; Buechler, R.; Metzler, S.; Theodoridou, A.; Gerstenberg, M.; Walitza, S.; Kollias, S.; Rössler, W.; Heekeren, K. Aberrant Coupling within and across the Default Mode, Task-Positive, and Salience Network in Subjects at Risk for Psychosis. Schizophr. Bull. 2014, 40, 1095–1104.

81.  Mallikarjun, P.K.; Lalousis, P.A.; Dunne, T.F.; Heinze, K.; Reniers, R.L.; Broome, M.R.; Farmah, B.; Oyebode, F.; Wood, S.J.; Upthegrove, R. Aberrant Salience Network Functional Connectivity in Auditory Verbal Hallucinations: A First Episode Psychosis Sample. Transl. Psychiatry 2018, 8, 69.

82.  Manoliu, A.; Riedl, V.; Zherdin, A.; Mühlau, M.; Schwerthöffer, D.; Scherr, M.; Peters, H.; Zimmer, C.; Förstl, H.; Bäuml, J.; et al. Aberrant Dependence of Default Mode/central Executive Network Interactions on Anterior Insular Salience Network Activity in Schizophrenia. Schizophr. Bull. 2014, 40, 428–437.

83.  Bolton, T.A.W.; Wotruba, D.; Buechler, R.; Theodoridou, A.; Michels, L.; Kollias, S.; Rössler, W.; Heekeren, K.; Van De Ville, D. Triple Network Model Dynamically Revisited: Lower Salience Network State Switching in Pre-Psychosis. Front. Physiol. 2020, 11, 66.

84.  Howes, O.D.; Murray, R.M. Schizophrenia: An Integrated Sociodevelopmental-Cognitive Model. Lancet 2014, 383, 1677–1687.

85.  Maia, T.V.; Frank, M.J. An Integrative Perspective on the Role of Dopamine in Schizophrenia. Biol. Psychiatry 2017, 81, 52–66.

86.  McCutcheon, R.A.; Nour, M.M.; Dahoun, T.; Jauhar, S.; Pepper, F.; Expert, P.; Veronese, M.; Adams, R.A.; Turkheimer, F.; Mehta, M.A.; et al. Mesolimbic Dopamine Function Is Related to Salience Network Connectivity: An Integrative Positron Emission Tomography and Magnetic Resonance Study. Biol. Psychiatry 2019, 85, 368–378.

87.  Moritz, S.; Silverstein, S.M.; Beblo, T.; Özaslan, Z.; Zink, M.; Gallinat, J. Much of the Neurocognitive Impairment in Schizophrenia Is due to Factors Other than Schizophrenia Itself: Implications for Research and Treatment. Schizophrenia Bulletin Open 2021, 2, sgaa034.

88.  Moritz, S.; Klein, J.P.; Desler, T.; Lill, H.; Gallinat, J.; Schneider, B.C. Neurocognitive Deficits in Schizophrenia. Are We Making Mountains out of Molehills? Psychol. Med. 2017, 47, 2602–2612.

89.  Gilbert, S.J.; Spengler, S.; Simons, J.S.; Steele, J.D.; Lawrie, S.M.; Frith, C.D.; Burgess, P.W. Functional Specialization within Rostral Prefrontal Cortex (area 10): A Meta-Analysis. J. Cogn. Neurosci. 2006, 18, 932–948.

90.  Simons, J.S.; Garrison, J.R.; Johnson, M.K. Brain Mechanisms of Reality Monitoring. Trends Cogn. Sci. 2017, 21, 462–473.

91. Rugg, M.D.; Vilberg, K.L. Brain Networks Underlying Episodic Memory Retrieval. Curr. Opin. Neurobiol. 2013, 23, 255–260.

92. Moscovitch, M.; Cabeza, R.; Winocur, G.; Nadel, L. Episodic Memory and beyond: The Hippocampus and Neocortex in Transformation. Annu. Rev. Psychol. 2016, 67, 105–134.

93. Guise, K.G.; Shapiro, M.L. Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. Neuron 2017, 94, 183–192.e8.

94. Fornito, A.; Wood, S.J.; Whittle, S.; Fuller, J.; Adamson, C.; Saling, M.M.; Velakoulis, D.; Pantelis, C.; Yücel, M. Variability of the Paracingulate Sulcus and Morphometry of the Medial Frontal Cortex: Associations with Cortical Thickness, Surface Area, Volume, and Sulcal Depth. Hum. Brain Mapp. 2008, 29, 222–236.

95. Buda, M.; Fornito, A.; Bergström, Z.M.; Simons, J.S. A Specific Brain Structural Basis for Individual Differences in Reality Monitoring. J. Neurosci. 2011, 31, 14308–14313.

96. Zmigrod, L.; Garrison, J.R.; Carr, J.; Simons, J.S. The Neural Mechanisms of Hallucinations: A Quantitative Meta-Analysis of Neuroimaging Studies. Neurosci. Biobehav. Rev. 2016, 69, 113–123.

97. Lisman, J.E.; Grace, A.A. The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. Neuron 2005, 46, 703–713.

98. Lodge, D.J.; Grace, A.A. Aberrant Hippocampal Activity Underlies the Dopamine Dysregulation in an Animal Model of Schizophrenia. J. Neurosci. 2007, 27, 11424–11430.

99. Perini, I.; Gustafsson, P.A.; Hamilton, J.P.; Kämpe, R.; Zetterqvist, M.; Heilig, M. The Salience of Self, Not Social Pain, Is Encoded by Dorsal Anterior Cingulate and Insula. Sci. Rep. 2018, 8, 6165.

100. Floresco, S.B.; Todd, C.L.; Grace, A.A. Glutamatergic Afferents from the Hippocampus to the Nucleus Accumbens Regulate Activity of Ventral Tegmental Area Dopamine Neurons. J. Neurosci. 2001, 21, 4915–4922.

101. Lodge, D.J.; Grace, A.A. The Hippocampus Modulates Dopamine Neuron Responsivity by Regulating the Intensity of Phasic Neuron Activation. Neuropsychopharmacology 2006, 31, 1356–1361.

102. Sonnenschein, S.F.; Gomes, F.V.; Grace, A.A. Dysregulation of Midbrain Dopamine System and the Pathophysiology of Schizophrenia. Front. Psychiatry 2020, 11, 613.

103. Howes, O.D.; Kambeitz, J.; Kim, E.; Stahl, D.; Slifstein, M.; Abi-Dargham, A.; Kapur, S. The Nature of Dopamine Dysfunction in Schizophrenia and What This Means for Treatment: Meta-Analysis of Imaging Studies. Arch. Gen. Psychiatry 2012, 69, 776–786.

104. Egerton, A.; Chaddock, C.A.; Winton-Brown, T.T.; Bloomfield, M.A.P.; Bhattacharyya, S.; Allen, P.; McGuire, P.K.; Howes, O.D. Presynaptic Striatal Dopamine Dysfunction in People at Ultra-High Risk for Psychosis: Findings in a Second Cohort. Biol. Psychiatry 2013, 74, 106–112.

105. Nakahara, S.; Matsumoto, M.; van Erp, T.G.M. Hippocampal Subregion Abnormalities in Schizophrenia: A Systematic Review of Structural and Physiological Imaging Studies. Neuropsychopharmacol Rep 2018, 38, 156–166.

106. Pantelis, C.; Velakoulis, D.; McGorry, P.D.; Wood, S.J.; Suckling, J.; Phillips, L.J.; Yung, A.R.; Bullmore, E.T.; Brewer, W.; Soulsby, B.; et al. Neuroanatomical Abnormalities before and after Onset of Psychosis: A Cross-Sectional and Longitudinal MRI Comparison. Lancet 2003, 361, 281–288.

107. Wood, S.J.; Kennedy, D.; Phillips, L.J.; Seal, M.L.; Yücel, M.; Nelson, B.; Yung, A.R.; Jackson, G.; McGorry, P.D.; Velakoulis, D.; et al. Hippocampal Pathology in Individuals at Ultra-High Risk for Psychosis: A Multi-Modal Magnetic Resonance Study. Neuroimage 2010, 52, 62–68.

108. Allen, P.; Luigjes, J.; Howes, O.D.; Egerton, A.; Hirao, K.; Valli, I.; Kambeitz, J.; Fusar-Poli, P.; Broome, M.; McGuire, P. Transition to Psychosis Associated with Prefrontal and Subcortical Dysfunction in Ultra High-Risk Individuals. Schizophr. Bull. 2012, 38, 1268–1276.

109. Medoff, D.R.; Holcomb, H.H.; Lahti, A.C.; Tamminga, C.A. Probing the Human Hippocampus Using rCBF: Contrasts in Schizophrenia. Hippocampus 2001, 11, 543–550.

110. Lahti, A.C.; Weiler, M.A.; Holcomb, H.H.; Tamminga, C.A.; Cropsey, K.L. Modulation of Limbic Circuitry Predicts Treatment Response to Antipsychotic Medication: A Functional Imaging Study in Schizophrenia. Neuropsychopharmacology 2009, 34, 2675–2690.

111. Lodge, D.J.; Grace, A.A. Developmental Pathology, Dopamine, Stress and Schizophrenia. Int. J. Dev. Neurosci. 2011, 29, 207–213.

112. White, T.P.; Joseph, V.; Francis, S.T.; Liddle, P.F. Aberrant Salience Network (bilateral Insula and Anterior Cingulate Cortex) Connectivity during Information Processing in Schizophrenia. Schizophr. Res. 2010, 123, 105–115.

113. Yan, H.; Tian, L.; Yan, J.; Sun, W.; Liu, Q.; Zhang, Y.-B.; Li, X.-M.; Zang, Y.-F.; Zhang, D. Functional and Anatomical Connectivity Abnormalities in Cognitive Division of Anterior Cingulate Cortex in Schizophrenia. PLoS One 2012, 7, e45659.

114. Rametti, G.; Junqué, C.; Bartrés-Faz, D.; Zubiaurre-Elorza, L.; Catalán, R.; Penadés, R.; Bargalló, N.; Bernardo, M. Anterior Cingulate and Paracingulate Sulci Morphology in Patients with Schizophrenia. Schizophr. Res. 2010, 121, 66–74.

115. Ellison-Wright, I.; Bullmore, E. Anatomy of Bipolar Disorder and Schizophrenia: A Meta-Analysis. Schizophr. Res. 2010, 117, 1–12.

116. Garrison, J.R.; Fernyhough, C.; McCarthy-Jones, S.; Simons, J.S.; Sommer, I.E.C. Paracingulate Sulcus Morphology and Hallucinations in Clinical and Nonclinical Groups. Schizophr. Bull. 2019, 45, 733–741.

117. Garrison, J.R.; Fernyhough, C.; McCarthy-Jones, S.; Haggard, M.; Australian Schizophrenia Research Bank; Simons, J.S. Paracingulate Sulcus Morphology Is Associated with Hallucinations in the Human Brain. Nat. Commun. 2015, 6, 8956.

118. Sass, L.; Borda, J.P.; Madeira, L.; Pienkos, E.; Nelson, B. Varieties of Self Disorder: A Bio-Pheno-Social Model of Schizophrenia. Schizophr. Bull. 2018, 44, 720–727.

119. Pienkos, E.; Giersch, A.; Hansen, M.; Humpston, C.; McCarthy-Jones, S.; Mishara, A.; Nelson, B.; Park, S.; Raballo, A.; Sharma, R.; et al. Hallucinations Beyond Voices: A Conceptual Review of the Phenomenology of Altered Perception in Psychosis. Schizophr. Bull. 2019, 45, S67–S77.

120. Rosen, C.; Jones, N.; Chase, K.A.; Gin, H.; Grossman, L.S.; Sharma, R.P. The Intrasubjectivity of Self, Voices and Delusions: A Phenomenological Analysis. Psychosis 2016, 8, 357–368.

121. Nelson, B.; Lavoie, S.; Gaweda, L.; Li, E.; Sass, L.A.; Koren, D.; McGorry, P.D.; Jack, B.N.; Parnas, J.; Polari, A.; et al. Testing a Neurophenomenological Model of Basic Self Disturbance in Early Psychosis. World Psychiatry 2019, 18, 104–105.