Supplementary materials

Search strategy

Development of the theoretical model and the accompanying discussion of relevant literature included the use of broad search string (i.e., free text) strategies using PudMed/MEDLINE, PsychInfo, and Google Scholar (a search engine that accesses additional databases). No sample population (humans vs. animal studies), publication date, or journal of publication restrictions were imposed; only articles written in English were included. Articles were also searched manually. Articles were analyzed and documented by 3 authors (MB, GL, LMT).
Note that this paper is not a systematic review, therefore exact numbers of papers found and filtered out are not given. The purpose of the search strategy was to build a theoretical model that integrates multiple disparate fields of study.
Specific search strings are below:

1. Search terms for literature relevant to sex effects on brain development and cognition in included: "sexual dimorphisms OR sex differences AND neurodevelopment OR brain development OR developmental trajectory OR frontal limbic OR frontal striatal OR emotion regulation OR reward processing OR cognition OR fMRI OR DTI OR grey matter volume"
2. Search terms for literature relevant to pubertal effects on brain development and cognition included: "adolescen* OR pubert* OR gonadal hormones OR estrogen OR testosterone OR DHEA AND neurodevelopment OR brain development OR developmental trajectory OR frontal limbic OR frontal striatal OR emotion regulation OR reward processing OR cognition OR fMRI OR DTI OR grey matter volume"
3. Search terms for literature specifically relevant to sex and pubertal effects on frontal limbic network included: "sex OR pubert* OR gonadal hormones OR estrogen OR testosterone OR DHEA AND *prefrontal* OR amygdala OR anterior cingulate OR emotion processing OR emotion regulation OR frontal limbic OR fMRI OR DTI OR grey matter volume"
4. Search terms for literature specifically relevant to sex and pubertal effects on frontal striatal network included: "sex OR pubert* OR gonadal hormones OR estrogen OR testosterone OR DHEA AND *prefrontal* OR *orbitofrontal* OR striat* OR reward* OR frontal striatal OR fMRI OR DTI OR grey matter volume"
5. Secondary search terms for literature relevant to the above topics *in schizophrenia spectrum disorders* included the above search terms plus "schiz* OR psychosis OR psychotic* OR high risk OR prodromal OR attenuated psychotic symptoms"

Annotated bibliography

This annotated bibliography summarizes and evaluates the papers covered in sections 2 to 4 of the main manuscript text. The annotations are grouped by where in the review they (first) appear. Each annotation mentions sample size (for empirical studies) or number of studies (for meta-analyses), sex, age range and predictors and outcomes to facilitate insight in methodological strengths and limitations across studies.

Topic (section of main text)	**Author and year**	**Study type**	**Sample size or no. of studies**	**Sex and species**	**Annotated bibliography**
Direct links between steroid hormones and SZ (section 2.1)	Bergeman n et al. 2007	Empirical paper	125	100% Female human	This study aimed to evaluate the association between the menstrual cycle or estrogen levels and symptoms in women with schizophrenia. The researchers measured timing of menses and assessed estradiol and positive and negative symptoms repeatedly in 125 adult women with schizophrenia (M age 35 years, SD=8). The results showed that psychotic, but not negative symptoms, were higher at times in the menstrual cycle that estradiol levels were lower. A strong aspect of the study is its examination of within-person changes in hormone levels and symptoms. However, its generalizability is somewhat limited by focusing on hospitalized patients and therefore on the higher end of the symptom range.
	Begeman n et al. 2012	Systematic review and/or meta-analysis	5	N/A	This systematic review and meta-analysis evaluated randomized controlled trials on the effect of estrogen supplementation on schizophrenia symptoms (5 studies). The effect of estrogen on reducing total, positive and negative symptoms in female patients was significant across studies. Their selection criteria ensured that only high-quality research designs on this topic were included in the meta-analysis.

					However, the results showed evidence of publication bias, which they don't acknowledge in the abstract.
	Huber et al. 2005	Empirical paper	68	100% Male human	This study aims to examine the hormonal differences between men with psychosis and healthy controls. Blood was drawn from 34 psychotic men and 34 healthy controls matched by age and BMI. Hormone levels ("testosterone, free testosterone, estradiol, oestrone, luteinizing hormone (LH) and follicle stimulating hormone (FSH)") were analyzed and compared using radioimmunoassay systems. Ages: psychotic men, 34.29 years (standard deviation (SD) 11.4); controls, 33.82 (SD 7.27) years. Psychotic men regardless of their medication exhibited significantly lower levels of all hormones, except LH and FSH which were within the healthy range, compared to healthy controls. A strong aspect of the study was their ability to separate the 34 psychotic men into groups based on their current medications and compare, in order to test assure that medications did not interfere with results; however, they had no hypothesis, their results section contained information that should have been in methods, and their one paragraph of a methods section left something to be desired.
	van Rijn et al. 2011	Empirical paper	42	100% Male human	This study aimed to examine the relationship between hormonal levels during different pubertal stages and the onset of psychotic symptoms in male adolescent patients. Researchers measured testosterone and oestradiol levels via saliva samples collected from male patients with prodromal symptoms (N=21) and from male controls without prodromal symptoms (N=21) between ages 12-18. Pubertal stage was assessed using Tanner stages criterion; prodromal symptoms and predictors of psychosis were assessed during SIPS and BSABS-P interviews. The results showed that patients with prodromal symptoms had lower levels of testosterone but equal levels of oestradiol compared to healthy controls throughout different pubertal stages. Results are consistent with current hypotheses regarding the relationship between hormonal levels and prepsychotic symptoms. However, sample size was relatively small

					and the study failed to consider the role of other hormones on the onset of psychotic symptoms.
	Markham 2012	Narrative review	N/A	N/A	This narrative review summarizes research on the role of estrogens and testosterone in schizophrenia from both animal and human studies. The review discusses the negative associations between estrogen and positive symptoms and between testosterone and negative symptoms; and it desribes influences of these hormones as well as sex differences in animal models of schizophrenia. The researchers acknowledge that the human research on hormone-brain associations has yet to be translated to preclinical schizophrenia research. It is a good starting point for understanding sex differences and relevance of sex steroids in schizophrenia, but is losing some of its relevance because of the research that has come out in the decade since this article's appearance.
	Ko et al. 2008	Empirical paper	26	100% Male human	The aim of this study was to determine the effects of a testosterone gel in addition to antipsychotic medication on schizophrenic symptoms in male patients. In this randomized, controlled trial males with schizophrenia (age 20-49) applied either a testosterone gel or placebo gel to four sites on the human body for 4 weeks (N=26), and then endured a 2 week washout period (N=17). Blood samples were used to assess hormone levels and PANSS, DIEPSS, and CDSS were used to analyze a patients' psychiatric symptoms. Testosterone treatment significantly decreased negative, but not positive or depressive, symptoms, more than placebo. Strengths of this study are its randomized, placebo-controlled design and its comparison of the effects of testosterone augmentation in patients taking varying antipsychotic medications or patients with varying baseline characteristics. However, it used a small sample of hospitalized patients, so generalizability may be limited.
	Ceskova et al. 2015	Empirical paper	498	40% Female human	This study aimed to investigate gender differences in responses to different treatments of first-episode schizophrenia by analyzing the results of the European First-Episode Schizophrenia Trial (EUFEST). A

					total of 498 participants (298 men, 200 women, ages 18-40) were randomly assigned to one of five different treatments being studied, and treatment response was evaluated using the positive and negative syndrome scale over the course of one year. Analysis of EUFEST data demonstrated a significantly higher response to antipsychotic treatment in female patients compared to male patients, and out of all five treatments studied, olanzapine showed the greatest difference in response with respect to gender. A strength of this study is its large sample size, and is a strong basis for the notion that gender differences should be considered when choosing treatments with antipsychotic drugs. A weakness of this study is its lack of placebo controls and allowance of simultaneous treatment by other medications.
Pubertal timing and SZ (section 2.2)	Cohen et al. 1999	Empirical paper	80	44% Female human	This study aims to examine the relationship between the onset age of schizophrenia and pubertal timing, hypothesizing an inverse relationship between the two. The study involved surveying 35 females and 45 males with DSM-IV defined schizophrenia or schizoaffective disorder as well as their mothers (68.6% of the females' mothers participated, 73.3% of males') about their first psychotic symptoms, pubertal timing (menarche age in females, age of voice change in males), DTREE to confirm diagnosis, substance abuse, family history, head injuries, and obstetric complications. Females with schizophrenia exhibited an inverse relationship between their menarche age and their schizophrenia onset, with younger ages of menarche corresponding to later onset of symptoms/hospitalization, suggesting the protective aspects of estrogen in schizophrenia. Retrospective assessment of puberty is a downside but the addition of the mothers helped verify the participants' survey results as well as provide insight to the obstetric complications scale.

	Galdos et al. 1993	Empirical paper	97	50% Female human	Researchers hypothesized that the onset of psychotic symptoms occurs earlier in females than in males due to earlier hormonal changes associated with puberty. Researchers used medical notes to assess psychotic symptoms and estimate the age of onset of illness for males (N=49) and females (N=48) ages 12-18 who were previously admitted in a South London hospital over a 13 year period. On average, females showed an earlier age of onset of psychotic symptoms than males. Only estimates of the age of onset could be made with the available information from previous medical notes. In addition, family history and other genetic factors were not taken into account.
	Hochman et al. 2004	Empirical paper	68	100% Female human	Researchers hypothesized that estrogen serves a protective role against the development of schizophrenia, so earlier ages of menarche should be correlated with later onsets of schizophrenia. Menarche age, clinical status, and time of schizophrenia onset were self-reported by a sample of 68 women (55 with schizophrenia, 13 with schizoaffective disorder; M age 45 years, SD=19). Results do not support the estrogen hypothesis that age of menarche and schizophrenia onset are negatively correlated. This study provides a basis for studying hormonal changes implicated in the early development of psychotic disorders in youth. Limitations of this study include its small sample size and data collection by self-report.
	Rubio-Abadal et al. 2016	Empirical paper	42	100% Female human	The aim of this study was to explore the relationship between menarche and psychosis onset, clinical severity and outcome in a population of females with first-episode psychosis. The researchers measured age at menarche, age at first-episode psychosis and clinical measures of positive and negative symptoms and cognitive functioning in 42 females with psychosis (M age 19 years, SD = 6) through self-report and report of a first-degree relative. The results showed that age of menarche was not significantly correlated with age at first-episode psychosis or

					symptom severity. Strengths of this study are the examination of clinical severity of positive, negative, and cognitive symptoms as a measure of comparison for clinical outcomes, and the methodological design that controlled for duration of untreated psychosis and pharmacological treatment. A weakness of this study is a small sample size and retrospective cross-sectional design.
	Ruiz et al. 2000	Empirical paper	105	100% Female human	The study hypothesizes that the relationship between the age of menarche and the age of one's first signs of schizophrenia are negatively correlated. The study involved 105 female patients from a Chile database of health records. 22 were randomly selected and clinically interviewed to verify that the data in the database were accurate. The mean of menarche start date in the 105 in the database was significantly lower than the rest of Chile, however, no correlation was found between the early onset of puberty and one's first SZ symptom. It has a large sample size due to it using a database, and I agree with the random sampling to see if the medical records were accurate.
	Moskow et al. 2016	Empirical paper	441	43% Female human	This study investigated the effect of age and the progression of puberty-associated stress on the risk for psychosis. Researchers hypothesized that there is a strong relationship between age/pubertal stage and stress in youth with high risk for psychosis. Salivary samples to measure cortisol levels and self-reported daily stress inventories were collected from 348 CHR (clinical high risk) youth and 93 healthy controls, ages 13-18. Pubertal stage was assessed using the Tanner scale. Prodromal symptoms and clinical risk were assessed via SIPS and SOPS. Although results indicate a positive relationship between age and stress/cortisol levels in both CHR and HC groups, the magnitude of relationship between age/pubertal stage and stress did not differ between the two groups,

					however stress levels were higher throughout all pubertal stages in the CHR group. Strengths of this study are the examination of the age-pubertal stage relationship, along with the age-stress and the pubertal stage-stress relationships, allowing for assessment of a number of factors on risk for psychosis. Limitations include the lack of longitudinal assessments and low representation in the pre-puberty CHR group.
	Ramanathan et al. 2015	Empirical paper	58	50% Female human	Researchers hypothesized that greater prodromal symptoms of psychosis would be associated with delayed onset of puberty in males and early onset of puberty in females. Researchers interviewed 58 child or adolescent family members of individuals with schizophrenia, using the Tanner Maturational Scale (TMS) to determine age of puberty and the Scale of Prodromal Symptoms (SOPS) to evaluate the presence of prodromal symptoms of psychosis. There was no significant correlation overall between being a later maturer and an individual's prodromal symptoms score, but when analyzing data separately among the two sexes, a later age of puberty was positively correlated with negative prodromal symptoms for males only. This study provides a basis for exploring how neuroendocrine markers of puberty can impact the onset of early psychosis. A limitation of this study is its small sample size but a strength is the focus on a child/adolescent sample and early signs of illness.
Frontal-limbic and frontal-striatal brain networks in	Modinos et al. 2015	Empirical paper	58	43% Female human	This study aimed to examine neural correlates of emotional salience in people with ultra high risk for psychosis (UHR), patients with first-episode psychosis (FEP), and healthy controls and the association with psychotic symptoms and level of global functioning. Researchers measured emotional arousal and corticolimbic response to emotional and neutral stimuli using fMRI in a sample of 18 FEP patients (M age 23

SZ (section 2.3)					years, SD = 4.6), 18 UHR patients (M age 24 years, SD = 4.1) and 22 healthy controls (M age 27 years, SD = 5). In the FEP and UHR groups, they measured global functioning, positive and negative symptoms, depression, anxiety, and stress. The results demonstrated reduced frontolimbic activation to emotional stimuli in FEP and UHR groups compared to healthy controls. Results also showed corticolimbic hyperactivity to neutral scenes in FEP and UHR participants that correlated with higher levels of positive symptoms and poorer levels of functioning. Strengths of this study include the comparison of both FEP and UHR groups with healthy controls, allowing for examination of frontolimbic activation and symptom severity in different early stages of illness. A weakness of this study is limited sample sizes.
	Vai et al. 2015	Empirical paper	54	48% Female human	This study aims to look at the difference between schizophrenic patients and controls in emotional circuit connectivity. In 25 SZ participants and 29 HC, they used a face-matching paradigm, fMRI, and DCM to study the connectivity between the dorsolateral prefrontal cortex, amygdala, visual cortex, fusiform gyrus, and ventral prefrontal cortex. Patients with SZ have lower top-down endogenous connectivity from DLPFC to Amy,a higher connectivity from Amy to VPFC and a lower driving input to Amy of affective stimuli compared to HC. The use of DCM to study connectivity is a strength and unique component. The sample is small, however, especially for a neural connectivity study.
	Park et al. 2018	Empirical paper	37	57% Female human	This study aimed to examine how abnormal amygdala-prefrontal cortex connectivity impacts the role of negative emotions during cognitive conflict control tasks in patients with schizophrenia. Researchers performed fMRI imaging on schizophrenic patients (N=17) and healthy controls (N=20) to identify abnormal activity in different brain regions as

					participants completed a modified Simon task in which conflict-resolution mechanisms were tested using emotional images. Results showed reduced right amygdala activity in schizophrenic patients compared to healthy controls during cognitive conflict under negative emotion, suggesting the role of a dysfunctional amygdala-prefrontal cortex neural circuit in abnormal cognitive conflict resolution occurring in patients with schizophrenia. In addition to a relatively small sample size, this study failed to account for the potential effects of medications, such as antipsychotic drugs, on the amygdala-prefrontal neural circuit. Despite this, researchers were able to attribute abnormal cognitive control resolution to a specific neural circuit.
	Anticevic et al. 2013	Empirical paper	165	54% Female human	This study aimed to determine if differences in amygdala whole-brain functional connectivity exist across phases of schizophrenia development. Researchers performed fMRI imaging and then compared whole-brain amygdala connectivity in healthy subjects (N=96), high-risk subjects (N=21), early schizophrenic patients (N=28), and chronic schizophrenia patients (N=20). Results showed that amygdala-orbitofrontal connectivity is reduced in patients with early or chronic schizophrenia, but amygdala-brainstem connectivity is increased in individuals that are at high risk for developing schizophrenia. This study is limited by its cross-sectional design, which prevents researchers from differentiating the stages of illness among high risk patients. Despite this, the study was able to replicate prior studies on amygdala connectivity in schizophrenia patients.
	Bjorkquist et al. 2016	Empirical paper	28	33% Female human	This study investigated the underlying affective impairment in schizophrenia using psychophysiological interaction analyses to examine functional connectivity between the amygdala and medial prefrontal

					cortex (mPFC) during an emotion perception task. Fourteen participants with schizophrenia (Mean age 31, SD = 7.5) and fourteen healthy controls (Mean age 31, SD = 11.3) viewed and rated positive, negative, and neutral images while undergoing functional neuroimaging. In the schizophrenia group, clinical ratings of positive and negative symptoms severity, motivation, and work and social functioning were collected. Results revealed a significant group difference in right amygdala-mPFC connectivity during perception of negative versus neutral images. Weaker right amygdala-mPFC coupling during negative compared to neutral image perception was associated with poorer social functioning in the schizophrenia group. Weaknesses include that functional outcome results approached significance after correcting for multiple comparisons, and the heterogeneity in antipsychotic medication usage in the schizophrenia group. However, the latter is reflective of patients receiving psychiatric care.
	Hooker et al. 2011	Empirical paper	58	26% Female human	This study tests the hypothesis that individuals with schizophrenia exhibit an exaggerated effect of emotional information on social judgments and this is associated with positive symptom severity. 23 individuals with SZ and 35 non-psychiatric controls completed a social judgment task in which they rated the trustworthiness of unfamiliar faces following a negative, neutral, or positive prime (IAPS). Relative to non-psychiatric controls, SZ participants judged faces as less trustworthy following the negative affective prime, but not the positive or neutral primes. Second, in SZ individuals the greater the influence of the negative affective primes on trustworthiness ratings, the greater the severity of positive symptoms, particularly suspiciousness/paranoia. Strengths include 1) the examination of both symptoms and internal affective state and/or irrelevant emotional information as potential

					influences on social judgements – associations not previously examined in the literature and 2) a task design that supports the delineation of these effects. Limitations include 1) the authors did not examine emotional state after presenting affective primes limiting interpretations of their precise effect on trustworthiness ratings and 2) the samples are small and unmatched on key demographics.
	Swartz et al. 2014	Empirical paper	71	68% Female human	The study aims to assess the relevance of the amygdala connectivity and response within adolescents during a face matching task. 45 patients with GAD, SP, and SAD; and 26 controls underwent an fMRI scan while they completed an emotional face matching task where participants chose which of two faces showed the same emotion as a given one. There was a heightened level of amygdala and prefrontal cortex-amygdala activity at the start of the task in the patient population compared to the controls, and activity in the amygdala that dwindled as the task continued significantly. The study was well written, with a well-established fMRI paradigm, but a small and heterogeneous clinical sample.
	Bertocci et al. 2017	Empirical paper	41	44% Female human	This study aimed to identify the effect of longitudinal changes in neural circuitry on the progression of psychiatric symptoms in youth. Researchers collected neuroimaging data during a modified emotional n-back test and assessed symptom severity in youth (N=41) during 2 different sessions, averaged 21 months apart. Results indicate longitudinal changes in neural activity in the amygdala and prefrontal cortex, as well as longitudinal changes in symptom severity, suggesting the role of changes in neural mechanisms in determining symptom severity. Findings have strong implications for identifying appropriate treatment options for specific psychiatric symptoms. However, relatively

					high participant dropout between scanning sessions may have impacted results.
	Gard et al. 2018	Empirical paper	167	100% Male human	The aim was to investigate socioemotional processing in low-income urban males in relation to internalizing symptoms. Researchers performed fMRI scans while a sample of 167 male participants (age 20) completed an implicit emotional face processing task and shape-matching task, to identify regions of the brain that demonstrated strongly positive or negative connections to the amygdala. Weaker amygdala–middle frontal gyrus negative connectivity to fearful faces and stronger amygdala–inferior frontal gyrus negative connectivity when viewing neutral faces predicted increases in internalizing behaviors. Strengths of this study include its examination of within-subject neuronal connectivity, a relatively large sample size, and focus on a historically underrepresented demographic group. However, the validity of the fMRI emotional face-processing task is limited due to missing participants' ratings of facial stimuli valence.
	Lin et al. 2015	Empirical paper	226	55% Female human	The study examined clinical outcomes for individuals identified as ultra-high risk for psychosis who did not develop a psychotic disorder. The researchers used a retrospective cohort design to measure DSM-IV diagnoses, negative and positive symptoms, attenuated psychotic symptoms, IQ, and global functioning in 226 participants (125 F, 101 M, M age 25.5, SD = 4.8) from a specialist clinic for young people at ultra-high risk for psychosis. The results demonstrated that individuals at ultra-high risk for psychosis who do not transition to psychosis are at significant risk for continued attenuated psychotic symptoms, persistent or recurrent disorders, and incident disorders. The authors were able to examine clinical outcomes in a large sample of the UHR diagnostic

					group, where few follow-up studies have been performed. A weakness, however, is the high variability in follow-up time since initial diagnosis.
	Yang et al. 2018	Empirical paper	104	90% Female human	In accordance with the RDoC conceptualization of psychiatric disorders, the authors test the hypothesis that symptoms of depression are dimensionally related to disrupted cognitive control neural circuitry in both MDD and PTSD and that hypoactivity in cognitive control circuitry will improve in both MDD and PTSD following cognitive behavioral therapy treatment. This is a longitudinal treatment study with two timepoints. At baseline, 28 individuals with MDD, 53 individuals with PTSD, and 23 non-psychiatric controls completed an emotional conflict task during fMRI. A subset of individuals (16 MDD; 20 PTSD then completed a 12-week course of CBT and returned for follow up fMRI. At baseline, cognitive control related neural activation was negatively associated with depression symptoms in both MDD and PTSD. Following CBT treatment, cognitive control regions (VLPFC, DLPFC) showed increased activation compared to baseline. Results indicate common neural mechanisms underlying symptoms in MDD and PTSD. Strengths include 1) dimensional examination of associations between symptoms and neural activity in the frontal-limbic brain network across psychiatric disorders, challenging the categorical distinction between disorders and 2) examination of effect of evidence-based therapy on neural circuitry. Limitations include 1) many participants were lost to follow up / did not complete CBT treatment limiting sample size and statistical power and 2) there was no association between change in neural activation and change in symptoms at follow up.
	van der Gaag et	Systematic review	10	N/A	This meta-analysis evaluated randomized controlled trials on the effectiveness of antipsychotics, omega-3 fatty acids, and cognitive

	al. 2013	and/or meta-analysis			behavioral therapy (CBT) for preventing transition to psychosis in those at clinically high risk for psychosis (10 studies). They found 54% risk reduction after 1 year across the preventions, and significant risk reductions for antipsychotics and CBT specifically. The authors clearly outline their search strategy and included evaluation of the quality of the empirical studies as well as examination of heterogeneity and publication bias.
	Kindler et al. 2016	Empirical paper	21	38% Female human	This study tested whether the estrogen receptor modulator raloxifene restores neuronal activity during emotional response inhibition in patients with schizophrenia. 21 adults with schizophrenia (13 male; mean age=35.5) were blindly randomized to either placebo first and then 120mg of raloxifene for 6 weeks each, or raloxifene and then placebo. fMRI with an emotional go/no-go task and self-report of symptoms was completed at baseline and after each treatment. Raloxifene increased dorsolateral PFC activity during the inhibition of response to negative words, but did not change performance on the fMRI task or symptom levels. The randomized, double-blind, placebo-controlled, cross-over design is a strength of the study, but the sample size is small, especially for a neuroimaging study, and the fMRI thresholds were lenient.
	Hubbard et al. 2019	Systematic review and/or meta-analysis	26	N/A	This meta-analysis summarized 26 studies comparing blood cortisol levels between first-episode psychosis patients and controls. They report significantly higher basal cortisol levels in patients than controls, but with high heterogeneity of results, partially explained by when the blood sample was taken (i.e. time of day, after fasting) and how it was analyzed. The meta-analysis clearly outlines inclusion criteria and examines heterogeneity and publication bias. The main limitation is that

					they focus on basal cortisol level based on a single sample, even though the majority of variation in cortisol is actually within-person fluctuations.
	Zorn et al. 2017	Systematic review and/or meta-analysis	25	N/A	This review and meta-analysis examined if cortisol stress reactivity in patients with various psychiatric disorders may vary by sex and symptomatic state. A total of 25 unique studies containing cortisol stress data for varying psychiatric disorders (14 studies on MDD, 9 studies on anxiety disorders, 4 studies on schizophrenia) were included in the meta-analysis and were analyzed based on sex and symptomatic state. Results showed that in response to psychosocial stress, women with MDD or an anxiety disorder exhibited a lower cortisol response compared to healthy controls while men with MDD or a social anxiety disorder exhibited a higher cortisol response. This study was the first meta-analysis to evaluate cortisol stress reactivity for different psychiatric disorders using standardized cortisol outcomes. A limitation of this study is its relatively low number of schizophrenia papers included compared to the other psychiatric disorders reviewed, and its simplification of the cortisol stress response (only comparing the overall increase in cortisol levels while excluding other factors of the response) in the meta-analysis.
	Walter et al. 2016	Systematic review and/or meta-analysis	8	N/A	This systematic review evaluates studies reporting hippocampal volume in subjects with clinical high-risk compared to healthy controls. In the meta-analysis review comprising 8 studies and 1429 subjects, authors found no significant difference in hippocampal volume for clinical high risk patients in comparison to healthy controls, and no reduction in hippocampal volume before transition to psychosis. Investigators clearly outline inclusion and exclusion criteria of empirical studies of hippocampal volume in high-risk populations that use the same

					methodology for assessing ROI, and explain exclusion of outliers in later analysis.
	Radua et al. 2012	Systematic review and/or meta-analysis	43	N/A	The authors use a novel meta-analytic approach to examine multi-modal (function and structural) brain abnormalities in first episode psychosis (FEP). The primary hypothesis was that FEP individuals would show functional and structural abnormalities in the same regions. Secondary analyses examined the potential confounding effects of a variety of covariates, including antipsychotic medications. 43 studies (25 structural, 18 functional across fMRI, PET, and SPECT studies) were included in the meta-analysis. FEP individuals showed both structural and brain activation abnormalities in the medial and anterior cingulate cortex and in the insula. Antipsychotic medications were associated with some reductions in gray matter volume and could explain some of the observed effects. This study uses a novel meta-analytic method and multi-modal approach to better delineate the neural mechanisms of psychosis. Limitations include standard challenges with meta-analyses (e.g., study methodological differences) and possible false negatives due to low sample size.
	Waltz et al. 2009	Empirical paper	36	25% Female human	The study investigated relationships between primary reinforcer responses and ratings of negative symptoms with the hypothesis that brain responses to temporal difference errors (dopamine neurons signal mismatches between expected and actual outcomes) in dopamine midbrain nuclei and target areas would be abnormal in schizophrenic patients. Researchers recruited a sample of 36 people (18 controls and 18 patients with a schizophrenia or schizoaffective diagnosis) where participants experienced a timed classical conditioning reward task, underwent an fMRI scan, and also participated in psychological

					assessments to determine levels of cognitive functioning, identify symptoms of depression, and quantify participants' ability to experience pleasure. Patients with schizophrenia showed reduced contrasts in brain activity evoked by positive and negative temporal difference error in multiple brain areas, patients showed reduced responses in numerous brain regions to delivery of the reinforcer on standard trials, and clinical ratings of avolition correlated with brain activity evoked by reinforcer delivery in the primary gustatory cortex and putamen. Limitations of this study are that the small sample size reduces generalizability and since it's published in 2009 more recent literature might provide more information. The strength of this study is that its simple but effective design increase replicability.
	Chung et al. 2016	Empirical paper	63	37% Female human	This study aimed to examine group differences between schizophrenia patients and controls in neural activation during cognitive control in reward and non-reward contexts. 36 SZ patients (M age =38) and 27 controls (M age = 35) completed a Stroop-like task during fMRI scanning, where some of the trials were preceded by a reward cue. SZ patients showed lower putamen activation to reward cues; and lower dlPFC activation during the cognitive control phase of the task was associated with negative symptoms. The study uses an interesting and theoretically-relevant task set up, but the small sample and ROI approach make replicability difficult.
	Katthagen et al. 2020	Empirical paper	43	33% Female human	The aim of this study was to investigate the relationship between elevated striatal dopamine synthesis capacity and striatal reward prediction error signaling in patients with schizophrenia. Striatal dopamine synthesis capacity was measured by FDOPA-PET and reward prediction error signal by fMRI scan in 20 unmedicated schizophrenia

					patients and 23 healthy controls as participants completed a reversal learning task. There was no significant difference in striatal dopamine synthesis capacity found between schizophrenia patients and healthy controls, and there was no correlation between dopamine synthesis capacity and striatal reward prediction error signaling in schizophrenia patients also this group performed worse on the reversal learning task compared to controls and demonstrated greater heterogeneity in the PET measures of dopamine synthesis. A limitation of this study is its small sample size, but its strength is its focus on unmedicated patients which eliminates antipsychotic drugs as potential confounding variables affecting dopamine measures. However, previous medication and illness history was not considered in the study's cross-sectional design, so patients with non-dopaminergic psychosis may have been included in the study.
	Murray et al. 2008	Empirical paper	25	28% Female human	Investigators examined the links between psychotic experience, reward learning and dysfunction of the dopaminergic midbrain and associated target regions to better understand how dopamine dysfunction generates psychotic symptoms. In this study 13 first episode psychosis patients (FEPs) with active psychotic symptoms and 12 healthy control participants performed a reward conditioning task while undergoing fMRI. Investigators found abnormal activation in FEPs compared to healthy controls associated with reward prediction error in the dopaminergic midbrain, striatum and limbic system. One of the first studies to associate negative symptoms with abnormal striatal activation, however, strength and reproducibility of findings are limited with less than 20 participants per group.

	Zhang et al. 2016	Systematic review and/or meta-analysis	53	N/A	This study sought to examine transdiagnostic neural substrates of three sub-domains of anhedonia (anticipatory, consummatory, emotional processing) in both SZ and MDD using ALE meta-analytic methods on fMRI studies. 53 studies (33 MDD and 24 SZ) were included in ALE analyses. Anticipatory anhedonia was associated with frontal-striatal regions (dACC, medial & middle frontal gyri); consummatory anhedonia was associated with decreased activation in the ventral basal ganglia. MDD and SZ showed significant overlap in neural activation associated with anticipatory and consummatory anhedonia but not emotion processing. The main strength of this study is the novel transdiagnostic approach to the neural mechanisms underlying sub-domains of anhedonia observed in multiple disorders. Limitations include those associated with ALE methods, and potential confounding effects of psychotropic drug effects on reward processing circuitry.
	Lee et al. 2019	Empirical paper	53	37% Female human	The study hypothesized that patients with schizophrenia would show reduced neural sensitivity (with a specific focus on the ventral striatum, the anterior cingulate cortex, and the ventromedial prefrontal cortex) to social rewards, but not to nonsocial rewards. Researchers had a sample of 53 individuals (27 controls and 26 patients with a schizophrenia diagnosis) undergo fMRI social and non-social reward tasks to determine neural sensitivity. Patients and controls showed comparable levels of neural sensitivity to nonsocial rewards, but patients showed blunted neural sensitivity to social rewards compared with controls and in the social reward condition, patients showed a significant correlation between optimal response and neural sensitivity to reward across all 3 brain regions. The study has a small sample size and all of the participants were utilizing antipsychotic medication which could have

					had an effect on the results of the study since the primary focus was on neural sensitivity.
	Segarra et al. 2016	Empirical paper	66	21% Female human	This study aimed to understand differences in brain activation, motivation, and anhedonia in depression and schizophrenia in comparison to those without illness. Using fMRI, authors measured activation of fronto-striatal regions during a motivation task in 24 patients with depression (Mean age 33, SD = 9.15), 21 patients with schizophrenia (Mean age = 32, SD = 7.4) and 21 healthy controls (Mean age 34, SD = 10.1). Authors found hypofunction in ventral striatal and orbitofrontal regions in depression and schizophrenia during unexpected reward receipt. A weakness of the study is that all groups were skewed to be more representative of males, especially the schizophrenia group.
	Waltz et al. 2013	Empirical paper	50	22% Female human	This study aimed to investigate what underlying neural pathways might explain the previously observed deficit on tasks of rapid reinforcement learning in patients with schizophrenia. A sample of 29 chronic schizophrenia patients and 21 matched normal controls completed an fMRI probabilistic reversal learning task to evaluate rapid reinforcement learning ability. Compared to healthy controls, schizophrenia patients showed similar magnitudes of punishment-evoked deactivation of the ventral striatum but reduced deactivation of the default mode network in the medial prefrontal cortex, correlating to negative symptoms of schizophrenia such as avolition and anhedonia. The schizophrenia patients included in this study were all outpatients being treated with antipsychotic medications, which may have influenced reward-related neural responses by interacting with dopamine activity.

	Juckel et al. 2006	Empirical paper	20	100% Male human	This study examined the association of negative symptoms with dysfunction of the brain reward system in schizophrenia. Researchers used fMRI to assess BOLD response in ventral striatum of 10 men with unmedicated schizophrenia (Mean age 26, SD = 7.8) compared to 10 healthy volunteers (Mean age 31, SD = 8.4) during presentation of reward-indicating and loss-indicating stimuli. Compared to healthy controls, the unmedicated schizophrenia group showed reduced ventral striatal activation during reward-induing cues, and decreased activation of the left ventral striatum was inversely correlated with the severity of negative symptoms. A major weakness of this study is the inclusion of solely male participants. Despite the small and limited sample size, this study demonstrated a relationship between reduced ventral striatum activation during reward anticipation and negative symptoms.
	Simon et al. 2010	Empirical paper	30	33% Female human	This study sought to delineate associations between specific depression and negative symptoms (apathy, anhedonia) and specific stages of reward processing (anticipation vs. receipt). 15 individuals with SZ and 15 nonpsychiatric controls completed the monetary incentive delay task during fMRI. Symptoms were assessed in SZ participants only. No group differences in VS activation during reward processing were observed. In SZ participants, lower VS activation during reward anticipation was associated with more severe apathy symptoms but not other negative symptoms; lower VS activation during reward receipt was associated with more severe anhedonia and depression symptoms but not other negative symptoms. A strength of this study is that it includes measurement of VS activation during both reward anticipation and receipt and links these to conceptually related symptoms for a more precise understanding of the neural mechanisms underlying apathy,

					anhedonia, and depression in SZ. However, results should be interpreted with caution due to the small sample size for fMRI analyses.
	Pornpattananangkul et al. 2019	Empirical paper	2566	48% Female human	Authors aimed to map anhedonia in children onto changes in intrinsic large-scale connectivity and task-evoked activation and in comparison to low mood, anxiety, and attention-deficit/hyperactivity disorder. Using data from the Adolescent Brain Cognitive Development study, authors examined functional MRI connectivity within large-scale networks, between networks, and between networks and subcortical regions during rest (mean Age 10, SD = 0.62, 48% girls). Functional MRI activation was examined during reward anticipation and working memory. The results demonstrated that hypoconnectivity at rest and hypoactivation during reward anticipation correlate with lack of intrinsic arousal connectivity during rest and diminishment of extrinsic reward-arousal activity during reward anticipation in children with anhedonia but not with low mood, anxiety, or ADHD. The strength of this study is the use of the ABCD data set, providing a large sample size allowing for well-powered findings.
	Morris et al. 2015	Empirical paper	36 and 5	100% Male human	This study examined associations between endogenous testosterone levels and ventral striatal activity during reward prediction error processing and whether this association is disrupted in individuals with SZ. Secondarily, the authors hypothesized this association is due to testosterone levels effecting human midbrain dopamine neurons via sex steroid receptors. 21 men with SZ and 15 men with no psychiatric diagnoses competed a reward prediction task during fMRI. Post-mortem midbrain tissue samples from 5 adult humans were used to examine the influence of testosterone on dopamine neurons. Ventral striatal (VS) activation during prediction error processing was associated with

					testosterone levels in nonpsychiatric controls but not in men with SZ. Postmortem tissue analysis indicated direct modulation of dopamine via sex steroids, consistent with evidence that testosterone modulates reward processing and consequently influences motivated behavior. A strength of this study is the examination of a potential mechanism for well-validated disruptions in VS activation during reward processing in SZ and its association with negative symptoms. Limitations include small sample sizes resulting in large confidence intervals for correlational analyses, and the possibility that antipsychotic medication contributes to disrupted VS activation in SZ due to dopamine antagonism.
Puberty and Frontal-Limbic Development (section 3.2)	Vijayakumar et al. 2018	Systematic review and/or meta-analysis	38 (brain structure) and 26 (brain function)	N/A	This systematic review evaluates (human) research on physical and hormonal indicators of pubertal maturation in relation to structural and functional brain development, focusing on reward, affective reactivity and cognitive control fMRI tasks. Most of the covered studies used Tanner staging or the Pubertal Development Scale to measure physical indicators of puberty, and examined testosterone and/or estradiol as hormonal indicators. The authors report inconsistency in findings and evaluate possible sources of this, including methodological variation. One of the most useful aspects of this review is the overview figures, pinning locations of findings across neuroimaging studies onto a single brain image. It also evaluates prominent theories on adolescent neurodevelopment and, similar to our review, finds that the evidence has only limited support for and sometimes contradicts e.g. the dual-systems and mismatch theories.
	Barendse and Pfeifer	Narrative review	N/A	N/A	This book chapter outlines how puberty, and its underlying hormonal and physical changes, might elicit a sensitive period for the development of the social brain. It evaluates the current literature on this topic and finds

	2021				that overarching conclusions are limited by methodological issues of prior research. The chapter provides a comprehensive overview of pubertal development, its links with structure and functioning of the social brain.
	Casey et al. 2011	Narrative review	N/A	N/A	This influential narrative review proposes the dual-systems model, which suggests that adolescent risk taking behavior is explained by a faster non-linear development of limbic and striatal regions, and slower, more linear development of the prefrontal cortex. They lay out evidence from neuroimaging studies, which at the time was a limited amount of mostly small-N studies. This theory has since garnered additional empirical support but also critique and empirical findings that contradict the theory (as we discuss in the current review).
	Pozzi et al. 2021	Systematic review and/or meta-analysis	48, 41, and 19	N/A	This fMRI meta-analysis evaluates neural activation during emotion reactivity (48 studies), and implicit (41 studies) and explicit (19 studies) emotion regulation in adolescence and young adulthood. Implicit regulation and emotion reactivity elicited activation in the amygdala and posterior temporal regions. Explicit emotion regulation was associated with activation in dorsolateral PFC, middle temporal gyrus and supramarginal gyrus. During implicit regulation, adolescents exhibited more consistent activation of the amygdala, fusiform gyrus, and thalamus but emerging adults displayed more consistent activation in the posterior superior temporal sulcus. The authors clearly outline the criteria for study inclusion, but their definition of implicit regulation is quite broad, leading to a mix of study designs being included under the meta-analysis of that construct. Also, the underlying empirical studies are mostly small-N studies with relatively lenient thresholds for activation.

	Pfeifer and Allen 2020	Narrative review	N/A	N/A	This narrative review explains how pubertal influences on brain development might trigger a cascading process of bidirectional influences between social experiences and brain functional or structural changes, which together shape risk for mental illness in adolescence. The authors provide a substantial amount of evidence for various pieces of the theoretical model, and provide testable hypotheses for future research. However, they discuss little research that illustrates the key idea of bidirectional or transactional influences.
Puberty and Frontal-Striatal Development (section 3.3)	Braams et al. 2015	Empirical paper	299	52% Female human	The current study used a longitudinal design to test the association between nucleus accumbens activity to pubertal development, rewards, and risk-taking behavior as well as whether changes in puberty and risk-taking behavior contributed to the presumed peak in nucleus accumbens activity. Researches used a longitudinal process over a two year period with a sample of 299 participants (48% male; Mean age =14.15 years; SD = 3.56) who engaged in fMRI reward tasks, provided saliva samples to test testerone levels, and utilized self-report measures to assess pubertal development, risk taking behavior, and risk/reward tendencies. Neural responses to rewards were shown peak during adolescence; testosterone levels were positively related to nucleus accumbens activity in both sexes. This study had a good sample size and longitudinal set up, making it one of the stronger studies on the topic.
	Shulman et al. 2015	Empirical paper	8270	49% Female human	The empirical study investigated sex differences in the developmental trajectories of self-reported impulse control and sensation seeking in adolescents and early adulthood. Authors examined impulse control and sensation-seeking by age and sex using longitudinal data from the National Longitudinal Study of Youth 1979 Child and Young Adult Survey (N = 8270). They found sex differences in both impulse control

					and sensation-seeking that increase with age, such that the window of heightened vulnerability to risk-taking during adolescence may be greater in magnitude and more protracted for males than for females. This study modeled sex-specific growth curves and sex differences in trajetory of the interested variables using a large, diverse and longitudinal sample. However, measures of sensation-seeking and impulse control are indirect measures of these cognitive processes.
	Silverman et al. 2015	Systematic review and/or meta-analysis	26	N/A	This meta-analysis evaluated fMRI studies with an activation likelihood estimation approach to determine the underlying core neural networks of adolescent reward processing. 26 fMRI studies on reward processing were included in the meta-analysis, and adolescent and adult reward-related brain activity were compared. Results showed similar subcortical brain networks in both adolescent and adult reward processing; adolescents exhibit greater likelihood of activation in the insula, putamen and amygdala, and adults showed greater activation of executive control regions in the frontal and parietal lobes. This meta-analysis' focus on studies that reported activation coordinates in whole-brain analyses allowed for effective determination of differences between adolescent and adult reward processing.This study was limited by its inclusion of studies with relatively small sample sizes and studies with wide variability in fMRI reward tasks.
	Goddings et al. 2014	Empirical paper	275	43% Female human	This study investigated pubertal influence on structural development of subcortical brain regions. Researchers used a longitudinal sample of 711 MRI scans from 275 individuals (117 F, M age 12.8 , SD = 3.3; 158 M, M age 13.8, SD = 3.4) to examine the relationship between age, Tanner pubertal stage, and volume in subcortical ROIs. Pubertal development and age both had independent and interactive influence on volume for the

					amygdala, hippocampus and putamen in both sexes and caudate in females. There was an interactive puberty-by-age effect on volume for the nucleus accumbens and globus pallidus in both sexes, and caudate in males. The authors utilized a longitudinal design, which reduces confounds due to variability in structural volume between individuals and allows researchers to examine differences in growth trajectories. The study could have been further strengthened by directly measuring pubertal hormones testosterone and estrogen.
	Wierenga et al. 2018	Empirical paper	271	54% Female human	The aim of this study was to figure out chronological age effects on development of subcortical brain volumes, test effects of pubertal measures, self-report pubertal stage and testosterone level, above and beyond chronological age, and test whether the influence of testosterone (on brain volumes) is puberty specific. It is a longitudinal study with 271 participants engaging in MRI scans, with physical pubertal maturation being assessed by the Pubertal Development Scale (PDS) for participants under 18 and testosterone levels were assessed in morning saliva samples. Significant age-related changes were observed for most volumes for both males and females, with a heterogeneous pattern of both volume increases and decreases with increasing age and puberty related increases in testosterone were associated with developmental changes in some subcortical brain volume. The strengths were the large sample size and the longitudinal design, including examination of non-linear associations.
	Koolschijn et al. 2014	Empirical paper	215	52% Female human	The aim of the study is to assess the association between pubertal stage, testosterone, LH, and estradiol with maturation in the anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), inferior frontal gyrus (IFG), and orbitofrontal cortex (OFC). Researchers collected the

					saliva and urine from a sample of 215 participants between the ages of 8-25 with normal intelligence to test their estradiol, testosterone, and LH levels, with the exclusion of girls utilizing hormonal contraceptives. Puberty Developmental Scale (PDS) was completed by all participants under the age of 18 to assess pubertal stage. Higher testosterone levels contributed to ACC and OFC gray matter volume maturation, higher estradiol levels contributed to smaller ACC gray matter volumes, there were no significant associations between DLPFC, IFG or subcortical structures and testosterone, estradiol or LH, and there was an association between pubertal stage and inferior frontal gyrus (IFG) gray matter volume. This study was effective in showing the relationship between gray matter maturation, pubertal development, and sex hormones.
	Neufang et al. 2009	Empirical paper	30	50% Female human	This study aimed to examine the associations between Tanner Stage, testosterone levels and grey matter volume in adolescents. 30 adolescents aged 8-15 (15 female and 15 male) completed a sMRI and underwent medical examination for Tanner stage and a blood draw for testosterone and estradiol. Amygdala volume was positively associated with Tanner Stage and testosterone levels in both sexes, whereas hippocampal volume was negatively associated with those pubertal indicators and striatal volumes were unrelated to Tanner stage or hormone levels. This is one of the first studies to look at hormone levels in association with brain structure, but it is limited by the small sample size and cross-sectional design.
	Chahal et al. 2021	Empirical paper	156	57% Female human	The aim of this study was to determine how sex differences in the maturation of white brain matter may relate to sex differences in reward and punishment-related behaviors, specifically investigating the morphometry of white matter fibers connecting the orbitofrontal cortex

					and nucleus accumbens. For a sample of 156 adolescents (89 females, 67 males), pubertal stage was determined using the Pubertal Development Scale, patients completed a Sensitivity to Punishment and Sensitivity to Reward Questionnaire, and diffusion-weighted MRI data were collected and underwent Fixel-based analysis. Results revealed sex differences in adolescent maturation of the fronto-accumbal tract: fronto-accumbal tract fiber density and cross-section was negatively associated with punishment sensitivity and positively associated with pubertal stage in males, and specifically fiber cross-section was related to reward sensitivity in both sexes. A weakness of this study is its cross-sectional design and also its usage of self-report measures of reward and punishment sensitivity that may be conducive to bias. Nevertheless, this study effectively demonstrates sex differences in white matter morphometry during puberty (in fiber density and fiber cross-section).
	Alarcón et al. 2017	Empirical paper	167	43% Female human	This cross-sectional analysis examine adolescent sex differences in the neural substrates of reward sensitivity during a risky decision-making task. Authors used fMRI in 96 boys (Mean age 14, SD = 1.3) and 71 girls (Mean age 14, SD = 1.3) to examine if girls compared to boys show heightened brain activation in the nucleus accumbens during reward receipt, mediated by testosterone and estradiol levels. Results demonstrated the nucleus accumbens was recruited more robustly in males during reward, and that sex hormones did not mediate this effect. The authors used adequate methodological design for sex hormone collection by utilizing venipuncture collection that controlled for time and menstural cycle phase. However, measurements were taken the week of the MRI and thus do not reflect acute mechanisms of action due to hormone levels on the day of MRI.

	Op de Macks et al. 2011	Empirical paper	50	66% Female human	The aim was to directly test the relationship between gonadal hormone concentrations and activity in the striatum in response to reward outcomes in adolescents across different stages of puberty. 50 adolescents 10-16 years old with 17 boys (M age = 13.5, SD = 2.3) and 33 girls (M age = 12.9, SD = 1.8) completed two self report measures of pubertal maturation, provided saliva samples to test for gonadal hormone levels, and participated in an MRI task where they could choose to take a (small or large) risk (to play) or not take a risk at all (skip or reset the trial). Individual differences in gonadal hormone levels at different stages of puberty are positively associated with individual differences in the neural response to monetary reward, suggesting that the drastic rise of gonadal hormone levels at puberty may contribute to increased reward sensitivity observed in adolescents but this finding was more prevalent in boys (for testosterone) than in girls (for testosterone and estradiol). A limitation was the small sample size and the lack of separation across puberty levels but a strength was that the results seemed to be in line with similar research results.
	Forbes et al. 2010	Empirical paper	77	51% Female human	This study examines puberty-specific changes in reward-related brain function in adolescents, of varied pubertal maturation, by stimulating striatal and medial prefrontal cortex (mPFC) reactivity and utilizing adults as a comparison. The sample included 19 adults as the comparative group and 77 adolescents within an 11-13 age range wherein 26 were pre/early pubertal adolescents and 51 were mid/late pubertal adolescents with the intent to maximize the pubertal variability within groups of girls (Mean age 11.49, SD =.60) and boys (Mean age 12.39, SD =.59) and to avoid confounding of puberty by age. Within adolescents, striatal reactivity during reward anticipation and reward outcome was positively correlated with subjective positive affect, striatal

					reactivity during reward outcome was negatively correlated with depressive symptoms, and medial prefrontal cortex reactivity during reward anticipation was modestly, positively correlated with depressive symptoms. This study's limitations include its cross-sectional design, reliance on single measures of mood and depressive symptoms, and the age range and size of the adult group; however I do think that there is a strength in varying the age of boys and girls to match pubertal variability across sexes.
	Ladouceur et al. 2019	Empirical paper	79	59% Female human	The goal of this study was to examine the extent to which striatal activation and cortico-striatal functional connectivity to cues predicting upcoming rewards would be positively associated with pubertal status and levels of DHEA, testosterone, and estradiol. Authors measured pubertal maturation with the Tanner stage and PDS, and had 32 boys (Mean age 12, SD = 0.95) and 47 girls (Mean age 11, SD = 0.83) perform a reward cue processing task during fMRI. In girls only, higher estradiol was associated with less striatum activation to reward cues and higher ventral striatum to putamen connectivity, whereas higher testosterone was associated with stronger ventral striatum to ACC and insula connectivity. A strength is the multiple samples taken to measure hormone levels. A limitation is the cross-sectional design and small groups.
	Poon et al. 2019	Empirical paper	67	46% Female human	This study examines adolescents' reward system functioning in relation to the level of pubertal maturation, specifically to pubertal hormone levels (testosterone and estradiol). In a sample of 67 adolescents aged 12-14 years, pubertal status was self-reported using the Pubertal Development Scale, testosterone and estradiol levels were determined from saliva samples, and participants completed a reward task while

					undergoing fMRI scans. This study found that testosterone levels were associated with activation in the lower bilateral dorso-lateral prefrontal cortex while estradiol was not associated with any of the reward-processing regions of interest, and that pubertal hormone levels were broadly related to connectivity between reward regions. A weakness of this study is its cross-sectional design, especially given its focus on how reward processing differs across pubertal stages. Results are consistent with prior research on reward processing in adolescents and provides additional support for the link between pubertal hormone levels and adolescent risky behaviors.
	Vijayakumar et al. 2019	Empirical paper	82	50% Female human	This study examines age and puberty-related changes in adolescent affective brain function. In a longitudinal sample of 82 adolescents (Mean age 10 at onset, SD = 0.31), authors measured salivary testosterone and affective reactivity of the amygdala to emotional faces in three fMRI sessions, each three years apart. Authors found nonlinear relations between changes in testosterone levels and limbic activation in response to emotional faces. The longitudinal design with three time-points, direct collection of testosterone, and modeling of non-linear associations were all strengths of this study in investigating age and puberty effects on brain development. This study could be strengthened with a collection of estrogen in addition to testosterone.
Hormonal mechanisms (section 4.1)	McCarthy et al. 2008	Narrative review	N/A	N/A	This narrative review describes how estradiol influences brain development, specifically apoptosis, synaptogenesis, morphometry of neurons and astrocytes, and their downstream effects on sex differentation. The influences of estradiol vary by brain region and can be protective or disruptive. The review mostly focuses on fetal brain development and on knowledge from animal research.

	Maninger et al. 2009	Narrative review	N/A	N/A	This comprehensive narrative review describes the effects of DHEA and DHEAS on the brain. It discusses general mechanisms of synthesis (in the adrenal and brain) and action (binding various receptors), as well as specific neurobiological effects such as neuroprotection, neurite growth, apoptosis, anti-glucocorticoid and anti-inflammatory action. Since this review was published in 2009, more recent literature is not covered, but more recent reviews are not as comprehensive or have another focus than the neurobiological actions.
	Schulz et al. 2009	Narrative review	N/A	N/A	This narrative review updates the organizational-activational hypothesis that exposure to steroid hormones early in development masculinizes and defeminizes neural circuits, programming behavioral responses to hormones in adulthood. It proposes that this not only occurs in fetal development, but also in adolescence. The evidence they provide is mainly derived from animal research on behaviors that are not always easy to translate to human behavior. Nonetheless, the paper is a good starting point for generating specific hypotheses on the influence of steroid hormones on neural circuits in adolescence.
	Stahl 2018	Narrative review	N/A	N/A	This narrative review discusses the role of three neurotransmitter systems in the etiology of psychosis: dopamine, serotonin and glutamate. The brief nature of the article doesn't allow for a lot of nuance, but it nonetheless provides a good overview of key theories with clear visuals.
Dopamine (section 4.2)	Markham et al. 2011	Narrative review	N/A	N/A	The aim of this narrative review was to summarize research on the effects of maternal immune activation, psychological stresses, and malnutrition on fetal brain development, and how that might be relevant to psychosis and depression. The authors conclude that these environmental factors have an impact on a wide range of brain

					developmental and behavioral outcomes, like disrupted PFC development and hyperdopaminergia, with stress exposure specifically having an influence on HPA axis programming. The review discusses a large body of literature, making it comprehensive but also difficult to judge the strength of the evidence for each element. It would have been beneficial to have a table or other overview summarizing which studies support and don't support each of the links between the perinatal factors and outcomes (i.e. for each arrow in their figure).
	Sinclair et al. 2014	Narrative review	N/A	N/A	This narrative review discusses the impact of sex hormones and stress hormones during adolescence on dopamine neurotransmission. They describe that sex hormones exert part of their effects on dopamine signaling through midbrain dopamine neurons, while stress hormones, at least in adulthood, act on dopamine receiving neurons in the PFC, striatum, and nucleus accumbens. The authors point out that the influences of sex and stress hormones on dopamine signaling might be relevant to schizophrenia, but don't comprehensively review the literature regarding sex hormones and schizophrenia.
	Gogos et al. 2010	Empirical paper	80 and 54	100% Female rats	This animal study examined whether estrogen treatment can reverse prepulse inhibition (PPI) disruption induced by 8-OH-DPAT or the dopamine D1/D2 receptor agonist apomorphine. It also compared these effects to the ability of serotonin or dopamine receptor antagonists in reversing the disruption by 8-OH-DPAT and apomorphine on PPI. 80 ovariectomized female rats were given a high or low or zero dose of estradiol and two weeks later tested for PPI. They were tested for PPI after administration of saline and 8-OH-DPAT or apomorphine. Another 54 ovariectomized rats were given a serotonin or dopamine receptor antagonist with or without 8-OH-DPAT or apomorphine and tested for

					PPI. Only the high dose of estradiol prevented the disruption of PPI by 8-OH-DPAT or apomorphine. The serotonin receptor antagonist WAY 100,635 and the dopamine D2 receptor antagonist haloperidol reversed the 8-OH-DPAT induced disruption in PPI. Haloperidol also reversed apomorphine induced PPI disruptions. A strength is the examination of both serotonin and dopamine antagonists. A limitation is that only females were included.
Glutamate (section 4.3)	Tibbo et al. 2004	Empirical paper	42	62% Female human	They hypothesized that high quantities of glutamate and glutamine could be an indicator of genetic risk for SZ in adolescence and be used to predict it using H-MRS. The study included 20 adolescents all of which had a parent with SZ (Mean Age=16.4) and 22 adolescents with low genetic risk as the HC (Mean Age= 16.7) and comparing their spectra within the right medial frontal lobe. They concluded that there were significantly higher quantities of glutamate and glutamine in the genetically at risk population compared to the low risk group. The paper is well written and the study seems thorough and well thought out. It is one of the few examining glutamate levels in an adolescent population.
	Poels et al. 2014	Narrative review	N/A	N/A	This narrative review summarizes the literature on the difference in glutamate levels and NMDA receptor density between people with schizophrenia and controls. Proton MRS studies suggest higher glutamate levels in the medial PFC in unmedicated SZ, and PET/SPECT studies support the theory of hypofunction of the NMDA receptor. The authors also outline prospects of future research. This review could have taken a more systematic approach to evaluate the proton MRS and PET/SPECT studies, including an outline of search strategy and table

					listing effect sizes of all studies, or a meta-analysis (like Egerton, et al., 2017 for GABA).
	Wei et al. 2014	Empirical paper	5 to 14 per experiment	50% Female rats	This study examines estrogen as a molecular mechanism underlying sex differences in glutamate-receptor mediated PFC response to stress in male and female rats. The investigators conducted a series of experiments in which male and female rats were exposed to repeated stress (RS) and then completed an explicit memory task designed to engage medial PFC regions. Authors examined 1) sex differences in the effect of RS on explicit memory and glutamatergic transmission in the mPFC, and 2) estrogen modulation of the effect of RS on explicit memory and glutamatergic transmission in the mPFC. Results indicate that 1) females do not show negative effects of RS on explicit memory or glutamatergic transmission but males do, and 2) that inhibition of estrogen in females and administration of estradiol in males reversed this effect. Collectively, these results indicate estrogen is protective against the effect of RS on PFC related functions and could explain sex differences in stress response/resilience. Strengths include: inclusion of both male and female rats, measurement of both cognitive function and glutamatergic function, systematic isolation of estrogen's modulation of the effect of RS in the PFC.
GABA (section 4.4)	Egerton et al. 2017	Systematic review and/or meta-analysis	16 and 7	N/A	This systematic review and meta-analysis summarized 1H-MRS studies (16 studies) and PET/SPECT studies (7 studies) that compared participants with schizophrenia-spectrum disorders or at high risk and controls on GABA or GABAA/benzodiazepine receptor availability. GABA did not differ between groups in the medial prefrontal cortex, parietal/occipital cortex or striatum; and there were no consistent group differences in GABAA/benzodiazepine receptor availability. The authors

					looked at several potential moderators to try to explain the non-significant results and high heterogeneity (without success). Likely because of the low number of empirical studies, they combined research on high risk populations and the whole range of SZ disorders, making it harder to draw specific conclusions.
	Piekarski et al. 2017	Empirical paper	unknown	100% Female mice	Researchers hypothesized that pubertal hormones (ovarian hormones in female mice) can regulate frontal lobe maturation by causing an increase in inhibitory neurotransmission. In mice subjects, researchers manipulated exposure to gonadal steroids to advance the age of puberty onset and then measured the strength of excitatory versus inhibitory neurotransmission to the cingulate cortex with slice electrophysiology. Mice behavior was also observed as mice were trained in a 4-choice discrimination and reversal task. Mice that were exposed to hormonal treatment and underwent early puberty exhibited greater inhibitory neurotransmission in the frontal cortex and altered behavior on cognitive tasks. This study effectively demonstrated the age-independent role of pubertal hormones in the neurodevelopment of the frontal cortex. A limitation of the study is its exclusion of male mice subjects.
	Lichenstein et al. 2016	Systematic review and/or meta-analysis	37	N/A	This mixed meta-analysis and narrative review evaluates research on the normal development of the ACC in adolescence and how its connectivity and function is altered in depression. The meta-analysis shows linear development of two ACC-linked tracts: the cingulum and anterior thalamic radiation. They conclude that ACC structural and functional connectivity is altered in adolescent depression. However, a closer look at their very helpful tables shows that the small samples, varied methods, and inconsistent findings of the empirical studies prevent a conclusion on how ACC connectivity is altered.

	Fornito et al. 2009	Narrative review	N/A	N/A	This narrative review summarizes structural neuroimaging and neuropathological studies comparing structure of the anterior cingulate cortex (ACC) between SZ groups (high-risk, first-episode, and schizophrenia) and controls. Authors conclude that SZ involves grey matter reductions in the ACC that may precede psychosis onset and progress with illness duration. With regard to the neuropathological findings, they report reductions in neuronal, synaptic, and dendritic density. This is a clearly structured article that bridges two separate bodies of literature, but could have used more discussion of the meaning of altered ACC structure for its functioning.
	Bloch et al. 2000	Empirical paper	16	100% Female human	This study investigated the role of changes in gonadal steroid levels in postpartum depression by stimulating two hormonal conditions related to pregnancy and parturition in euthymic women with and without a history of postpartum depression. In this clinical trial study, investigators collected self-report depression and mood ratings in 8 participants with a history of postpartum depression (Mean age 33.5, SD = 7.8) and 8 without a psychiatric history (mean age 33.5 , SD = 4.9) during three phases: baseline hormone levels, hypergonadal hormone levels to simulate pregnancy conditions, and hypogonadal hormone levels during withdrawal of hormones to stimulate the postpartum period. Results demonstrated a significant increase in depressive symptoms in women with a history of postpartum depression but not in the comparison group after hormone withdrawal, compared with baseline. This article illustrates a direct connection between hormone level and mood using a strong methodology that controlled for menstrual cycle variability and pre-existing mood disorders, though in the manipulative phases the study is not able to fully replicate the reproductive endocrine events during

					pregnancy, notably the length at which hypergonadal state occurs. A challenge with this study is also the small sample sizes for each group.
	Hantsoo et al. 2015	Narrative review	N/A	N/A	This article aims to give an overview of the etiology, characteristics and treatment of premenstrual dysphoric disorder. It discusses influences of progesterone on GABAa receptors and influences of estradiol of serotonin receptors and transporters as potential etiological mechanisms behind the mood symptoms. Although not directly relevant to SZ, this article illustrates a field that actively investigates hormonal influences on neurobiological processes. It shows that interindividual variation in the sensitivity of the brain to hormonal fluctuations might be an important factor, an idea that could be worth examining in the context of other mental illnesses too.
	Andréen et al. 2009	Narrative review	N/A	N/A	This narrative review focused on arguments indicating a link between GABA steroids (steroids produced mainly by the ovary during the menstrual cycle, by the adrenals, and by the testes in men) and negative mood effects. The review discussed the known negative mood effects (premenstrual dysphoric disorder (PMDD), the negative mood symptoms encountered during sequential progestagen addition to oestrogen treatment in postmenopausal women, and the side effects of oral contraceptives, and compared them to the positive (anxiolysis and sedation) effects and when GABA-A receptors mediate adverse effects a subset of individuals is sensitive to low doses or concentrations of GABA-A receptor modulators and thus respond with severe adverse emotional reactions when provoked (3-10%); and that about one third of individuals are moderately affected by the GABA-A receptor modulators at low dosage. The review gives a clear hypothesis regarding the mechanism behind the hormone to negative mood association, it covers a

					lot of human, experimental research, but some of the covered research is a bit outdated.
Serotonin (section 4.5)	Barth et al. 2015	Narrative review	N/A	N/A	This narrative review summarizes research on the complex influences of sex hormones on serotonin, dopamine, GABA and glutamate systems. They bridge findings from animal and human literature and discuss potential implications of the findings for the risk for mental illness in periods of life characterized by hormonal change. Summary tables could have helped to gain oversight of the findings discussed in this comprehensive review.
	Bethea et al. 2002	Empirical paper	20	100% Female primate	This study tested the hypothesis that raloxifene and arzoxifene (selective estrogen receptor modulators (SERMs)) would have similar effects as estradiol on serotonergic gene targets (tryptophan hydroxylase (TPH), serotonin reuptake transporter (SERT), and the 5HT1A autoreceptor) when tested in ovariectomized primates. Primates were ovariectomized and orally given estradiol, raloxifene or arzoxifene once per day by sipper bottles for 30 days. Estradiol, raloxifene, and arzoxifene led to an increase in TPH transcription, a decrease in SERT gene expression, and no effect on 5HT1A. This study used primates to figure out the effects of estradiol and the SERMs so it is unclear how well the results translate to a human population. This study is pretty innovative in its use of raloxifene and arzoxifene to mimic the results of estradiol without the potential risks of ovarian/uterine cancer.
	Gogos et al. 2006	Empirical paper	11	100% Female human	This study examined the effects of estrogen treatment on sensorimotor gating by investigating the modulation of prepulse inhibition (PPI) by the serotonin-I A (5-HTIA) receptor partial agonist, buspirone. Electromyogram activity was measured in 11 healthy women (Mean age

					26, SD = 2) following four acute treatment conditions: placebo, buspirone, estradiol, and combined buspirone and estradiol. Authors found that estrogen treatment, administered in appropriate experimental conditions prevented PPI deficits induced by 5-HTIA; at the 120 ms interstimulus interval buspirone caused a significant disruption of PPI and pretreatment with estrogen prevented this disruption. This study is one of the first to examine in a human sample how estrogen affects the 5-HT1A receptor modulation, previous studies were conducted in rat models. Generalizability to a female schizophrenia population is limited, due to a small sample size of healthy individuals and with findings in a specific ISI window that could be impacted by attentional influences.