CHUN, CHARLOTTE A., Ph.D. A Meta-Analysis of Context Integration Deficits Across the Schizotypy Spectrum. (2018).
Directed by Dr. Thomas R. Kwapil. 267 pp.

Schizophrenia is the most extreme manifestation of schizotypy, a continuum of symptoms and impairment that ranges from minimal impairment to full-blown psychosis. Schizophrenia has been described as a disorder of disrupted context integration (CI), the ability to assimilate internal and external information into coherent mental representations. CI in schizotypy is often measured with the AX-Continuous Performance Task (AX-CPT) and the Dot Pattern Expectancy (DPX) task. Research using these tasks indicated CI deficits along the schizotypy spectrum, but has primarily been conducted with schizophrenia patients. There have been mixed findings regarding outcomes such as error patterns and the association of CI deficits with schizotypy symptom dimensions. Further, conclusions were limited by generally small sample sizes, heterogeneous patient variables, and varied task parameters across studies. The current study used systematic review and meta-analysis to collect and synthesize data on AX-CPT and DPX performance across the schizotypy spectrum. CI impairment was present across the schizotypy spectrum. CI deficits in schizophrenia were substantial in magnitude and correlated with disorganized and negative symptom dimensions. Error patterns suggested a specific deficit in CI, which was larger than deficits attributed to broader cognitive impairment and general psychopathology. When examining subgroups, CI performance was comparable between chronic and first-episode schizophrenia patients. Groups at risk to develop schizophrenia demonstrated moderate CI impairment. The results were generally robust across task parameters and there was no evidence of reporting biases. In

sum, these findings lend additional support to theories suggesting that CI is a stable vulnerability factor for schizophrenia.

A META-ANALYSIS OF CONTEXT INTEGRATION DEFICITS ACROSS THE SCHIZOTYPY SPECTRUM

by

Charlotte A. Chun

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2018

Approved by

_____
Committee Chair

APPROVAL PAGE

This dissertation written by Charlotte A. Chun has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Thomas R. Kwapil

Committee Members _____
Michael J. Kane

_____
Suzanne M. Vrshek-Schallhorn

_____
Paul J. Silvia

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Thomas R. Kwapil for his enduring support, guidance, and humor during this project. I also thank my committee members, Drs. Michael J. Kane, Suzanne M. Vrshek-Schallhorn, and Paul J. Silvia for lending their expertise and providing input on this project.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Overview

The current study is a meta-analysis of context integration (CI) deficits in schizotypy-spectrum psychopathology. The study followed best practices and employed systematic review procedures to gather empirical data on CI in different groups along the schizotypy spectrum. Two CI tasks from the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia (CNTRACS; Gold et al., 2012) battery were included: the AX-Continuous Performance Task (AX-CPT; Servan-Schreiber, Cohen, & Steingard, 1996) and the Dot Pattern Expectancy (DPX; Jones, Sponheim, & MacDonald, 2010; MacDonald, Goghari et al., 2005) task. Meta-analysis was used to synthesize results and estimate the magnitude of CI impairment across the schizotypy spectrum. Variables thought to moderate or predict effect size were also examined, including factors such as symptom subtypes, patient characteristics, and task parameters. This is the first known, comprehensive meta-analysis of CI across the schizotypy spectrum.

Context Integration

> I only saw fragments: a few people, a kiosk, a house. To be quite correct, I cannot say that I see all of that, because the objects seemed altered from the usual. They

> did not stand together in an overall context, and I saw them as meaningless details…My impressions did not flow as they normally do (Matussek, 1987, pg. 92).

This account confronts us with the importance of context, which usually goes unnoticed but leaves a strange and fragmented world when disrupted. Context is a complicated construct that has been broadly used, but poorly defined, in the psychopathology literature. In this paper, I use context to mean both internal information—such as task schemas—and external information—such as perceptual features of the surrounding environment—that affect attention, interpretation, or behavior (Barch et al., 2004; Barch et al., 2001; Cohen & Servan-Schreiber, 1992). I conceptualize context integration (CI) as the adaptive, dynamic ability to assimilate internal and external information into a coherent mental representation.

Schizophrenia and Schizotypy

Schizophrenia is a severe mental illness that affects about 1% of the population (Diagnostic and Statistical Manual of Mental Disorders: DSM-5; American Psychiatric Association, 2013). It is part of a spectrum of psychotic and nonpsychotic illnesses, which in turn represent the most severe manifestations of a broader continuum of clinical and subclinical symptoms and impairment referred to as schizotypy (e.g., Kwapil & Barrantes-Vidal, 2015; Lenzenweger, 2010; Meehl, 1989). This paper will use the term schizotypy to refer to the broad construct ranging from subclinical to psychotic expression.

The schizotypy spectrum is heterogeneous and includes positive, negative, and disorganized symptom dimensions (Kwapil & Barrantes-Vidal, 2015; Liddle, 1987). Positive schizotypy involves excesses and distortions in perception and thought content. Perceptual abnormalities can occur in any sensory modality and range from mild illusions, or misperceptions of sensory stimuli, to hallucinations, or perceptions in the absence of sensory stimulation. Disruptions in thought content range from mild and transient magical, referential, and suspicious ideation all the way to full-blown delusions. Negative schizotypy includes diminished pleasure, social interest, thoughts, speech, affect, and motivation, and ranges from mild diminutions in these experiences to a complete lack of enjoyment, thought, expression, and engagement in the world (Kwapil & Barrantes-Vidal, 2015).

Disorganized schizotypy, which is thought to be especially relevant for disrupted context integration, involves impairment in the ability to regulate thoughts and actions, manifesting as odd or disrupted speech and peculiar behavior (Kwapil & Barrantes-Vidal, 2015). Disorganized speech is often hard to follow and involves disruptions such as tangentiality or inappropriate use of words. Disorganized behavior includes atypical movements and inappropriate affect (Arndt, Alliger, & Andreasen, 1991; Liddle, 1987). Disorganized symptoms range from mild cognitive disruption, circumstantial and tangential speech, and difficulty organizing and completing tasks to formal thought disorder, such as completely derailed thinking, incoherent speech, and grossly impaired behavior.

Theoretical Basis of CI Impairment in Schizotypy

Early phenomenologists described the development of schizophrenia in terms of impaired CI. These theorists were heavily influenced by Gestalt theory, which relies on principles of the tendency to perceive groups or complete forms from individual visual elements (e.g., Wertheimer, 1912; Wertheimer 1922). Incorporating ideas from Gestalt psychology, Matussek (1952) and Conrad (1958) proposed that schizophrenic impairment in integrating individual stimuli within the perceptual context could eventually lead to delusions. This occurs as the dimming of the perceptual field causes stimuli to stand out, which leads to attentional capture, feelings of uncertainty and anxiety, and ultimately, the formation of delusional thought as a means of restructuring the disorganized perceptual framework. As the delusion becomes stronger, experiences are interpreted as part of the newly formed context.

Modern theories have focused on a neural basis for impaired CI in schizophrenia. CI is proposed to occur at both the globalized cellular and the perceptual level (e.g., Phillips & Singer, 1997). Put simply, neurons have a contextual field that is sensitive to input from nearby cells. This causes nearby cells to synchronize firing, which is needed for normal perceptual integration (e.g., Phillips & Singer, 1997). The neural processing activates and influences top-down processing, which in turn can modulate contextual field responding (e.g., König, Chiang, & von Stein, 1997). It is thought that impaired coordination between bottom-up and top-down processing of contextual information may be associated with some symptoms of schizophrenia (e.g., Silverstein & Schenkel, 1997), which would indicate a problem with the neural mechanism of integration rather than

with vision. Accordingly, researchers have proposed that CI involves distributed network connectivity that is not limited to a particular area of the brain nor a singular cognitive function (e.g., Phillips & Silverstein, 2013; Ray et al., 2017). Phillips and Silverstein (2013) review information suggesting that widespread disruptions in coordination between nearby neurons and across longer-range brain regions have been found in schizophrenia, and that these processes are related to CI impairment.

It has been argued that impaired CI is not just an outcome of schizophrenia, but an underlying mechanism contributing to a number of cognitive, behavioral, and symptomatic manifestations of the illness (Barch & Braver, 2009; Cohen & Servan-Schreiber, 1992). It is proposed that a single CI process underlies attention, working memory, and inhibition aspects of executive control. In this way, context representations have been proposed as a top-down mechanism for focusing on task-relevant processes and for maintaining and updating these representations over time (Barch & Braver, 2009; Barch & Sheffield, 2017). For example, poor representation and maintenance of context may manifest as behavioral symptoms such as disorganized speech resulting from failure to interpret a phrase's meaning from the broader context of a sentence or conversation (Cohen & Servan-Schreiber, 1992).

Among other neurological abnormalities, schizophrenia involves disrupted dopamine functioning in the dorsolateral prefrontal cortex (DLPFC), which is associated with deficits in neurocognition and specifically, executive function (e.g., Cools, Brouwer, de Jong, & Slooff, 2000; Knable & Weinberger, 1997). CI deficits are thought to represent impaired function in the DLPFC and its connection with numerous neural

regions and neurotransmitter systems (see Barch & Sheffield, 2017 for review). In line with these predictions, CI deficits have been associated with diminished DLPFC activation and increased noise in the mesocortical dopamine system—that is, lower signal-to-noise ratio across the network—in empirical and computational studies (e.g., Barch et al., 2001; Braver, Barch, & Cohen, 1999; Lesh, Westphal, Niendam et al., 2013; MacDonald & Carter, 2003).

Phenomenological and neuroscientific theories of schizophrenia both posit a key role of disruptions in processing context at behavioral and neurological levels. However, there is ongoing debate about whether the underlying neural abnormalities are broadly distributed or more localized. Thus, questions remain about whether the mechanisms described above are actually interrelated, and about the timing of possible interactive processes. For example, do early neurodevelopmental abnormalities give rise to later impaired neural processing, which then result in abnormal perceptual experiences and manifestation of symptoms? In what ways is the process more nuanced, with disruptions at various levels of processing being interleaved or dependent upon environmental factors? As with many scientific fields, cross-cutting, interdisciplinary research is sorely needed to bridge together siloed theories from separate approaches.

Once CI deficits in schizophrenia were initially discovered, researchers began investigating two key questions: whether CI may be an endophenotype—a heritable, vulnerability marker (Gottesman & Gould, 2003)—of the disorder, and whether it is specific to schizophrenia versus related to psychopathology in general. To aid in the

detection of endophenotypes, Gottesman and colleagues have proposed six criteria throughout the years:

> 1) The endophenotype is associated with illness in the population. 2) The endophenotype is heritable. 3) The endophenotype is primarily state-independent (manifests in an individual whether or not illness is active) but may require a challenge to elicit the indicator...4) The endophenotype found in affected family members is found in non-affected family members at a higher rate than in the general population (familial association). 5) The endophenotype is more prevalent among the ill relatives of ill probands compared with the well relatives of the ill probands (i.e., co-segregation). 6) The endophenotype should be a trait that can be measured reliably, and ideally is more strongly associated with the disease of interest than with other psychiatric conditions (Chan & Gottesman, 2008, pg. 961-962).

Similarly, Nuechterlein and colleagues (1992) distinguished among three types of developmental processes: episode indicators, mediating vulnerability factors, and stable vulnerability indicators. Episode indicators are processes that are disrupted only during an illness episode; mediating vulnerability factors show continual impairment that is exacerbated during episodes; and stable vulnerability indicators involve consistent impairment, with severity unaffected by episodes (Nuechterlein et al., 1992).

Researchers have examined the diagnostic specificity of CI. Barch and colleagues (2003) proposed that CI is impaired in schizophrenia and in psychotic disorders more broadly. However, Barch et al. found a key distinction in timing between these psychiatric conditions: in non-schizophrenia psychotic disorders, impairment was not stable over time. Thus, CI may be an episode indicator in non-schizophrenia psychoses. In contrast, Barch and colleagues found that CI was consistently impaired in patients with schizophrenia across the span of a month, despite marked reductions in psychotic

symptoms during that time. Accordingly, they proposed that CI impairment is either a mediating vulnerability factor or stable vulnerability indicator for schizophrenia, supporting the suggestion that CI may represent an endophenotype for schizophrenia.

Overall, CI deficits are not unique to schizophrenia but patients with schizophrenia generally demonstrate impairment that is both more severe and more stable than other psychiatric groups, such as patients with other psychotic disorders or mood disorders (e.g., Barch et al., 2003; Cohen, Barch, Carter, & Servan-Schreiber, 1999; Reilly et al. 2017; Richard, Carter, Cohen, & Cho, 2013). In sum, the descriptive psychopathology and neuroscience literature both provide theoretical models for understanding the development and expression of schizophrenic psychopathology in terms of impaired CI.

Measurement of Context Integration

This review and meta-analysis specifically focused on measurement of CI using the AX-CPT and the DPX laboratory tasks. These tasks were selected because they map on to neural mechanisms implicated in CI (Braver, Barch, & Cohen, 1999). Furthermore, the tasks were designed to reveal a specific deficit in context processing; that is, greater impairment in CI compared to other areas of general cognition (Chapman & Chapman, 1973). This is accomplished through examination of differential error patterns, as described below.

AX-Continuous Performance Task. One of the most commonly used CI assessments is the AX-CPT, a modified continuous performance task (CPT) designed to assess the mental representation and maintenance of context. CPT tasks were originally

developed to assess sustained attention (Cohen, 2011). There are a variety of CPT tasks (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956) used in cognitive testing; for example, the single letter version, in which participants must respond to the letter X, the identical pairs version, in which participants must respond when a letter is shown twice in a row, or the standard AX version of the task, in which participants must respond when an A is followed by an X. In these CPT tasks, stimuli are presented rapidly and target trials occur infrequently so participants must maintain attention in order to respond correctly and detect the target amidst other stimuli. Standard versions are often used with patients with schizophrenia to assess vigilance (e.g., Nuechterlein, 1991).

Servan-Schreiber and colleagues (1996) modified the AX-CPT to include a higher proportion of target (AX) trials (80% in their original version) to keep the task goal relevant; because this new format creates a tendency to respond to X, they planned to assess BX errors in schizophrenia as an indication of difficulty with response inhibition. The AX-CPT and its nonverbal analog, the DPX (described below) were selected by the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia consortium as goal maintenance tasks for the CNTRACs battery, which includes cognitive measures recommended for cross-cutting research in schizophrenia (Barch et al., 2009). CNTRACs tasks were chosen by breakout groups of experts from various domains on the basis of each task demonstrating construct validity, mapping onto neural systems and animal models, demonstrating predicted outcomes in response to pharmacological manipulation in healthy subjects, showing evidence of impairment in schizophrenia, and exhibiting adequate psychometric properties.

The modified AX-CPT presents participants with a series of letters presented in succession; an item pair is made up of two sequentially presented letters, the cue followed by the probe. Researchers often manipulate the delay interval between the cue and probe, which may be different than the intertrial interval. They must make a positive (target) response when they see the letter A followed by the letter X (AX trials), and a negative (non-target) response to other letter pairs: AY, BX, and BY, with B representing any non-A cue and Y representing any non-X target. Although there are several versions of the task, the most common CI version involves 70% AX, 10% AY, 10% BX, and 10% BY trials.

Dot Pattern Expectancy Task. A nonverbal variant of the AX-CPT is the dot pattern expectancy task, which involves dot patterns instead of letters. This allows researchers to examine CI using novel spatial stimuli that may eliminate the potential confound of verbal strategies. Additionally, this variation typically involves a different proportion of trial types (69% AX, 12.5% AY, 12.5% BX, and 6% BY trials), providing a slightly greater number of the critical trials most strongly related to CI (BX trials) compared to the traditional AX-CPT format. The DPX is reportedly sensitive to CI impairment in both patients and their unaffected siblings (e.g., MacDonald, Goghari et al., 2005, Study 2), suggesting that it can detect deficits in less impaired individuals. Thus, it appears to be a promising tool for assessing CI deficits across the schizotypy spectrum.

The AX-CPT and DPX correlate between .63 and .80 (Strauss et al., 2014) and demonstrate comparable, adequate internal consistency and test-retest reliability overall;

however, the DPX has lower internal consistency for AY trials than the AX-CPT (Jones et al., 2010). A confirmatory factor analysis showed that BX trials from the AX-CPT and DPX tasks loaded on a context-processing factor, whereas AY trials from both tasks loaded on a preparatory factor. AX trials from both tasks loaded simultaneously on the context processing and preparatory factors, which were uncorrelated overall. Finally, BY errors serve as a check that participants understood task instructions and attended to the task – in that they neither are the correct target, nor follow the correct cue (MacDonald, Goghari et al., 2005, Study 1).

The cognitive neuroscience theories described above are particularly useful because they provide specific, testable predictions that relate AX-CPT/DPX performance to neurocognitive dysfunction in schizophrenic psychopathology. With these tasks, researchers are generally more interested in error patterns than in overall performance. Theoretically, AY trials should be the most difficult for individuals with intact context processing because they keep the task goal in mind and prepare to respond once they see the letter A. BX trials should be the most difficult for people with impaired context processing because the X triggers a response when the goal is not properly maintained (a target pair never begins with B) (Barch & Braver, 2009). This differentiation in cognitive abilities by trial allows for the examination of differential deficits in CI within a single task. For example, some studies have shown a double dissociation pattern, in which individuals on the schizotypy spectrum who presumably have impaired CI made more errors than controls on trials where poor CI is a disadvantage (BX trials) but fewer errors

on trials where poor CI is advantageous (AY trials) (Barch et al., 2001; MacDonald, Pogue-Geile, Johnson, & Carter, 2003).

Performance on CPT tasks is typically evaluated using signal detection theory to quantify participants' ability to distinguish a signal amidst background noise (e.g., Swets, 1961). For example in the standard AX-CPT, correctly detected AX targets are considered *hits*, failures to detect targets are considered *misses*, and inappropriate positive responses to non-target trials are considered *false alarms.* A sensitivity or detectability index, d-prime (d'), is calculated to quantify a person's ability to detect hits and ignore false alarms (Atkinson, 1963). The index of d' is the primary measure of vigilance in CPT studies in the schizophrenia literature (e.g., Nuechterlein, Green, Calkins et al., 2015; Wohlberg & Kornetsky, 1973).

Response patterns from the CNTRACs tasks are often used to compute *d'context*, a modified d' index designed to measure context sensitivity: d'context = $z(AX_{hits})$ - $z(BX_{false\ alarms})$. This value compares proportions of responses between AX hits and BX errors (rather than all false alarms) using a *z*-transformation, with lower d'context (more BX false alarms and AX misses) reflecting worse CI ability. d'context is thought to yield a better estimate of sensitivity to context than independent examination of errors (Cohen et al., 1999; Servan-Schreiber et al., 1996) and has become a gold standard for measuring context-processing ability in schizophrenia.

If differential error patterns and d'context reflect initial representation of context, then manipulation of the cue-probe delay is thought to reveal processes related to the maintenance of context representations (Barch & Braver, 2009). When CI abilities are

intact, BX accuracy should increase across longer delay periods (that is, at least across those typically reported in these tasks) because participants have more time to prepare a non-target response following the B-cue. In contrast, longer cue-probe delays should diminish AY accuracy in healthy participants because they have more time to prepare (an incorrect) positive response to the A-cue. The opposite pattern is predicted for people with poor CI, who have difficulty keeping task instructions in mind and maintaining the context of how a given cue should prepare them to respond. In schizotypy-spectrum groups, therefore, longer cue-probe delays should result in lower BX accuracy because the appropriate mental context ("B means I should prepare a negative response") is not maintained well across time, making them more likely to commit a false alarm when presented with an X. On the other hand, AY accuracy should theoretically increase with longer delay time in schizotypy-spectrum groups because deterioration of the appropriate mental context ("A means I should prepare a positive response") is actually beneficial in AY trials, which require a negative response (Barch & Braver, 2009).

Empirical Review of CI Impairment in Schizotypy Spectrum Psychopathology

To examine disrupted context processing as a potential endophenotypic marker of the illness, it must be assessed premorbidly. Simply demonstrating that patients with schizophrenia have difficulty integrating context cannot distinguish whether CI deficits reflect an underlying risk factor or are an outcome of the disorder due to experience of symptoms, stigma, or medication effects and simply part of the generalized performance deficit reported in schizophrenia (Chapman & Chapman, 1978). However, if CI impairment manifests early in the disorder and is present in people without clinically

significant levels of symptoms and functional impairment, this—among other criteria—lends credence to the theory of CI as an endophenotype. Thus, it is important to assess CI in different types of samples along the schizotypy spectrum. Studies using the AX-CPT and DPX have been conducted with participants who have schizophrenia and schizoaffective disorder, schizotypal personality disorder (SPD; a non-psychotic personality disorder on the schizotypy spectrum) (American Psychiatric Association, 2013), those at ultra-high risk to develop schizophrenia, first-degree relatives, and psychometrically identified schizotypy. However, the majority of these studies were conducted with patients with schizophrenia.

Published patient studies report CI deficits on CNTRACs tasks in terms of error patterns, reaction time, and d'context scores when comparing schizophrenia groups to healthy and psychiatric control groups (e.g., Barch et al., 2003; Cohen et al., 1999; Reilly et al., 2017). For example, patients with schizophrenia show diminished d'context—indicative of poorer CI ability—compared to healthy controls on both the AX-CPT (e.g., Barch et al., 2001) and the DPX (e.g., Henderson et al., 2012).

Regarding specific error patterns in schizophrenia, increased BX errors are reported in patient studies; however, the literature is variable on findings for AX, AY, and BY accuracy. Although findings are mixed in patients with schizophrenia, the most common pattern appears to be more AX and BX errors than AY and BY errors, relative to controls. Additionally, even when patients were found to perform worse than controls across all trials, a specific deficit was still seen in which patients perform worse on BX trials relative to AY trials (Henderson et al., 2012). Some studies even show a crossover

effect, in which patients outperform controls on AY trials despite worse BX performance (Barch et al., 2001; MacDonald et al., 2003). In sum, there is variability in error patterns across studies, but published studies seem to support CI deficits in patients with schizophrenia.

Only two identified studies have compared individuals with SPD to controls on the AX-CPT and DPX. They found CI deficits in SPD compared to healthy subjects and non-Cluster A personality disorders (Barch et al., 2004; McClure, Barch, Flory, Harvey, & Siever, 2008). Two other studies have examined CI in people with SPD pre- and post-pharmaceutical treatment (McClure et al., 2007; McClure et al., 2010). All four SPD studies showed greater BX relative to AY errors within (pre-treatment) SPD. Taken altogether, this provides initial evidence of a specific deficit in CI in non-psychotic individuals with schizotypy-spectrum disorders.

Only one known study has examined CI in ultra high-risk (prodromal) participants who have no history of psychosis (Niendam et al., 2014). These at-risk individuals showed 1) attenuated psychotic symptoms, 2) brief, limited intermittent psychotic symptoms, or 3) a substantial decline in functioning over the past year plus either SPD or a first-degree relative with psychotic disorder. In this study, the ultra high-risk participants showed more BX errors and lower d'context than healthy controls, again demonstrating preliminary support for CI impairment in earlier stages of schizotypic psychopathology.

Unaffected relatives of patients with schizophrenia are at greater risk for schizophrenia-spectrum psychopathology than the general population, presumably due to

shared genetic vulnerability (e.g., Tandon, Keshavan, & Nasrallah, 2008; MacDonald, Goghari et al., 2005). Unaffected relatives often show impairment intermediate to patients and control participants across a number of domains but there have been mixed findings in the literature (e.g., Delawalla, Csernansky, & Barch, 2008; Lopez-Garcia, Espinoza, Santos, Marin, & Sanchez-Pedreño, 2013; MacDonald, Goghari et al., 2005, Study 2; MacDonald et al., 2003; Richard et al., 2013). Thus, there is inconclusive evidence regarding the magnitude of possible CI impairment in unaffected relatives of patients with schizophrenia.

Only one group study (Sloat 2007, *unpublished dissertation*) was identified that used the AX-CPT to examine CI in a subclinical group scoring high on schizotypy questionnaires. Using two-tailed *t*-tests, no group differences were found for BX errors or d'context scores compared to healthy controls (our calculations from group summary data). However, extreme group designs can be problematic in psychometric identification studies because of the use of arbitrary cutoff points and the likelihood of underestimating effect sizes due to reduced statistical power (e.g., Preacher, Rucker, MacCallum, & Nicewander, 2005). Therefore, since continuous ratings of schizotypy have been shown to predict psychopathology better than extreme group scores (Kwapil, Gross, Silvia, & Barrantes-Vidal, 2013), we believe an examination of CI in subclinical schizotypy using continuous scores is indicated for future research.

In sum, CI deficits are frequently found across the schizotypy continuum in clinical samples and some at-risk samples. Studies assessing unaffected relatives and psychometrically identified schizotypy were few and reported mixed findings.

Nonetheless, there is preliminary evidence that CI impairment may begin early in the course of schizotypy-spectrum psychopathology. Additionally, these findings do not appear to simply reflect generalized performance deficits and suggest a specific impairment in context processing.

Other Factors That May Influence CI Impairment in the Schizotypy Spectrum

Diagnostic status alone may not necessarily explain poor CI ability in people with schizotypy-spectrum disorders. There are a number of other factors that may affect cognitive performance—and specifically, CI—in both clinical and subclinical manifestations of schizotypy, including medication status, illness duration and symptom dimensions.

Medication Status. Psychotropic medications can influence performance on cognitive tasks as certain commonly prescribed medications have been known to aid or impair cognition in schizophrenia (see Barch, 2005 for review). Barch and colleagues (2001) previously suggested that CI impairment is not likely caused by antipsychotic medication, since CI deficits are seen in medication-naïve patients with schizophrenia. Indeed, the current literature points to CI impairment compared to controls in both medicated (e.g., Chung, Mathews, & Barch, 2011; Fornito, Yoon, Zalesky, Bullmore, & Carter, 2011; MacDonald & Carter, 2003; Yoon et al., 2012) and unmedicated patients (e.g., Barch et al., 2003; Lesh et al., 2015; Niendam et al., 2014; Richard et al., 2013); however, it is still important to assess and rule out medication effects whenever possible.

Illness Duration. It has been suggested that length of illness may play a role in CI ability in schizophrenia (Lesh et al., 2013). However, the AX-CPT/DPX literature

suggests that both patients with first-episode (e.g., Braver, Barch, Cohen, 1999, Study 4; Lesh et al., 2013) and chronic (e.g., Cohen et al., 1999; Perlstein, Dixit, Carter, Noll, & Cohen, 2003; Stratta, Daneluzzo, Bustini, Prosperini, & Rossi, 2000) schizophrenia show substantial CI deficits compared to controls. Further, Thoma and Daum (2008) did not find any effect of illness duration on CI deficits within a group of chronic patients. Only one published study has directly compared recent-onset and chronic patient groups: Servan-Schreiber et al. (1996) found that unmedicated, multi-episode patients had lower d'context scores than unmedicated, first-episode patients, though they did not report the magnitude of this effect. Therefore, there is little direct evidence of how illness duration impacts CI in schizophrenia.

However, studies comparing patients with first-episode schizophrenia and chronic schizophrenia often are confounded with medication status (as first-episode patients are more likely than chronic patients to be medication naïve). Studies have examined interactions between medication status and illness duration. Lesh and colleagues (2015) found that both medicated and unmedicated first-episode patients were impaired relative to healthy subjects but that, not surprisingly, deficits were significantly more pronounced in the unmedicated group. However, few studies have conducted such nuanced examinations of task performance. Thus, medication status and illness duration are both important factors to consider when assessing CI abilities in schizotypy-spectrum disorders.

Symptom Dimensions. Because schizotypy and schizophrenia are heterogeneous, it is important to consider the effect of their symptom dimensions. Different schizotypic

symptom dimensions are proposed to have distinct etiological pathways (e.g., Andreasen & Carpenter, 1993; Myin-Germeys, Krabbendam, Jolles, Delespaul, & van Os, 2002) and differential patterns of cognitive impairment (e.g., Bora, Yucel, & Pantelis, 2009; Keefe et al., 2006; Liddle, 1987). Assessing schizotypy as a singular construct can lead to inconsistent, uninterpretable, and invalid findings (Kwapil & Chun, 2015); thus it is important to examine CI deficits in relation to specific symptom dimensions.

Disorganized Symptoms. Disorganized symptoms have the strongest theoretical and empirical link with CI deficits. Disorganized symptoms have been associated with decreased cortical activity in the dorsolateral prefrontal cortex (Goghari, Sponheim, & MacDonald, 2010; Yoon et al., 2008) and with impairment in certain aspects of executive function (Nieuwenstein, Aleman, & de Haan, 2001) broadly. Cohen and colleagues (1999) proposed an association between CI (as measured by the AX-CPT) and disorganization, especially formal thought disorder (Cohen et al., 1999). Of the schizotypic symptom dimensions, the greatest support has been found for the association of CI deficits and disorganized symptoms, with almost every study showing significant associations in various groups along the schizotypy spectrum (e.g., Ceccherini-Nelli, Turpin-Crowther, & Crow, 2007; Jones et al., 2010, *Study 3*; McClure et al., 2008; Richard et al., 2013; Sloat, 2007, *unpublished dissertation*).

Negative Symptoms. Cohen and Servan-Schreiber (1992) were the first to propose a theoretical link between negative symptoms and CI deficits. Negative symptoms are associated with diminished frontal and prefrontal cortex dopamine-linked activity (hypofrontality) and executive functioning deficits (e.g., Andreasen, Flaum, Swayze,

Tyrrell, & Arndt, 1990; Bora et al., 2009; Servan-Schreiber et al., 1996). Further, recent work in patients with schizophrenia has shown hypoactivation of the fronto-parietal network, which is thought to be important for cognitive control, in relation to task-based negative symptoms (poor reward learning) (Culbreth, Gold, Cools, & Barch, 2016). Thus, there are theoretical and functional bases for associations between negative symptoms and poor executive functioning; logically, this theory would extend to CI deficits but there have been less explicit connections between negative symptoms and CI in the theoretical literature compared to disorganized symptoms. Accordingly, associations between CI and negative symptoms have been demonstrated in the empirical literature but less reliably than in the disorganized symptom dimension. Patient studies have shown mixed results, with both significant negative (e.g., Barch et al., 2003; Javitt, Rabinowicz, Silipo, & Dias, 2007; Owoso et al., 2013; Niendam et al., 2014) and null (e.g., Ceccherini-Nelli et al., 2007; Cohen et al., 1999; MacDonald & Carter, 2003; Stratta et al., 2000) correlations between negative symptoms and CI ability. Additionally, the association between negative symptoms and CI deficits has not been supported across the entire schizotypy spectrum: null findings were found in clinical high-risk (Niendam et al., 2014), SPD (McClure et al., 2008), and psychometrically identified schizotypy (Sloat, 2007, *unpublished dissertation*) groups.

Positive Symptoms. Early phenomenologists described a possible connection between CI and delusions (a prominent positive symptom of schizophrenia) (Conrad, 1958; Matussek, 1952). Similarly, Kapur (2003) noted that the development of psychotic symptoms is associated with a sense of "aberrant salience" that arises when exaggerated

significance is assigned to inappropriate stimuli. He proposed that this process is mediated by excessive, context-independent dopamine release. Stratta and colleagues (2000) have further suggested that hallucinations might represent a failure in source monitoring associated—to some extent—with poor CI. Thus, there is some theoretical basis for the association between CI and positive symptoms. However, these theories have not received strong empirical support: most studies using the AX-CPT and DPX find no associations between positive symptoms and CI in various schizotypy groups (e.g., Gold et al., 2012; Javitt et al., 2007; McClure et al., 2008; Niendam et al., 2014; Owoso et al., 2013). Furthermore, one study found that associations with positive symptoms were no longer significant after accounting for disorganized symptoms (Becker, 2012, *unpublished dissertation*), suggesting that co-occurring symptoms may explain findings in the few studies that do show a relationship between the positive dimension and CI deficits.

Summary and Limitations of Symptom Association Studies. Disorganized symptoms show consistent associations with CI impairment across the schizotypy spectrum, negative symptoms show associations with CI impairment that are less consistent than the findings for disorganization and may not be present in all at-risk groups, and positive symptoms do not show strong associations with CI across the schizotypy spectrum. Bear in mind that the use of inappropriate measures or items in symptom dimension conceptualizations may limit the validity of findings; examples include social anxiety conceptualized as a disorganized symptom, increased excitability as a negative symptom, and conceptual disorganization as a positive symptom.

Inconsistent operationalization and measurement of symptom dimensions makes it difficult to draw valid conclusions from an already inconsistent literature. Conceptualization of symptom dimensions driven by theory rather than by measurement tools would allow for a better understanding of how these symptoms relate to cognitive factors such as CI.

Summary and Limitations of the Literature

In summary, the published literature suggests that a specific deficit in CI ability is present in patients with schizophrenia and some subclinical schizotypy groups. Studies examining symptom correlations indicated that CI may be associated with both negative and disorganized symptoms, but is most clearly related to the disorganization dimension. However, most researchers have treated schizotypy spectrum groups as unidimensional and have not specifically designed their studies to assess symptom associations. Because schizotypy is heterogeneous, differential patterns of impairment may exist within symptom dimensions and mask overall effects. Furthermore, studies that did employ a multidimensional perspective often measured symptom dimensions in ways that were inconsistent with the theoretical literature and were inconsistent across studies. Examining symptom correlations across studies using theory-driven symptom conceptualizations would provide a more valid assessment of associations between symptoms and CI ability.

There was considerable heterogeneity among studies regarding factors such as patient characteristics, control group characteristics, task parameters, and analyses conducted. This heterogeneity can make it difficult to directly compare studies in a

qualitative manner. Meta-analysis allows for a quantitative examination of the factors that account for variance in CI ability, which would provide a more powerful test of whether variables such as participant characteristics and task parameters moderate CI outcomes in different schizotypy spectrum groups. It is also likely that many of the reviewed studies may have been underpowered due to small sample sizes. In this case, a meta-analysis would provide greater power to examine CI impairment in schizotypy.

Selective reporting presents a potential issue in this literature: researchers often conduct planned contrasts on specific variables, but do not report outcomes for each trial type, which could affect the results of a meta-analysis due to missing data. There is also a possibility that file drawer effects, in which null and negative results are less likely to be published, have inflated the appearance of positive results (Rosenthal, 1979). Qualitative review of the literature is particularly susceptible to file drawer effects and fails to estimate the magnitude of CI deficits in the patient population (e.g., Cooper, 2010). Using systematic review to collect unpublished data can diminish the impact of publication bias. Examination of reported results across studies can also give an idea of how selective reporting affects results.

In summary, the foundational literature on CI deficits in schizophrenia is extensive, but we know much less about CI in subclinical schizotypy, particularly from a multidimensional perspective. The use of small samples and inclusion of heterogeneous study variables have limited the conclusions that can be drawn. Many of these shortcomings could be improved through use of systematic review and meta-analysis.

Despite the large number of studies examining CI in schizophrenia and schizotypy, there are no known, comprehensive meta-analyses of AX-CPT and DPX performance. Other meta-analyses have included various CPT tasks to assess factors such as cognitive remediation (e.g., Grynszpan et al., 2011) or executive function-related brain abnormalities (e.g., Fusar-Poli et al., 2007; Goghari, 2011; Minzenberg, Laird, Thelen, Carter, & Glahn, 2009) but this is the first meta-analysis to specifically assess AX-CPT and DPX for this purpose. Thus, this will be the first known meta-analysis to assess CI deficits across the schizotypy spectrum using the CNTRACs tasks.

Meta-Analysis

Meta-analysis is a statistical methodology that allows researchers to synthesize research in a systematic manner, estimate effect sizes, and increase overall power through the accumulation of data. Meta-analysis offers many advantages over traditional narrative approaches for summarizing research, which may: a) lack a standardized approach, b) employ imprecise search strategies, c) miss relevant research, d) gather information improperly, and e) be unable to show the overall magnitude of effect across studies (Cooper, 2010).

With meta-analysis, design flaws from the primary research studies being analyzed tend to "wash out," similar to how random sampling in experimental research can prevent confounds from participant characteristics. Meta-analysis can also address the file drawer effect, in which null results or findings that challenge accepted research are less likely to be reported, by obtaining data from unpublished studies and applying

corrective techniques. Thus, there are meta-analytic techniques that can protect validity of the synthesized results even when drawn from an imperfect literature base.

Goals

The goals of this project were to conduct an exhaustive review of the published and unpublished literature and to conduct a meta-analysis to: 1) examine overall CI impairment in the schizotypy-spectrum, 2) examine CI impairment in specific clinical and at-risk groups along the schizotypy spectrum, 3) examine associations between CI impairment and symptom dimensions in the existing literature and in a subset of studies with theory-driven symptom conceptualizations, and 4) examine the impact of patient characteristics and task variables on between-group outcomes. Specific hypotheses are detailed in the Method section below. Although hypotheses are not always specified in meta-analysis, it is recommended that meta-analytic studies running subgroup analyses make *a priori* predictions to demonstrate credibility (Sun, Briel, Walter & Guyatt, 2010).

CHAPTER II

METHOD

Project planning and *a priori* decisions for this project were pre-registered at Open Science Framework (https://osf.io/qhguz/) before analyses were conducted. The meta-analysis was conducted in keeping with the evidence-based guidelines proposed by the Meta-Analysis Reporting Standards (MARS; APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). The MARS guidelines were the outcome of an American Psychological Association task force that evaluated various standards for reporting meta-analyses from other scientific domains and made them more applicable to the social sciences. These steps are considered the gold standard for conducting meta-analyses in psychology. A list of general methodological decisions made in this study based upon best practices is provided in Appendix B.

Literature Search

As recommended by Cooper (2010), broad search terms were used (see Table 1) to search the databases of PsycInfo, PubMed, Google Scholar, and Scopus. In addition, searches were conducted in the following leading journals that publish empirical studies of schizotypy and schizophrenia: *Journal of Abnormal Psychology, Archives of General Psychiatry, Biological Psychology, Biological Psychiatry, Schizophrenia Research, Schizophrenia Bulletin, JAMA Psychiatry, Neuropsychology, American Journal of*

Psychiatry, Psychological Medicine, Journal of Clinical Neuropsychology, Psychiatry Research, and Schizophrenia Research: Cognition. Additionally, references were examined from key papers and 227 known researchers in the field were contacted for unpublished data. The researchers that manage the CNTRACS battery provided a contact list for individuals that have requested the use of the AX-CPT and DPX; therefore, this information was used as the basis of the contact list, adding in other key authors in the field. The primary author, who is an advanced graduate student in clinical psychology, screened studies found using the literature search using the criteria presented in Table 1.

Gathering Information from Studies

Information from the studies that met inclusion criteria were coded using a systematic spreadsheet, adapted from Cooper (2010): see Appendix C. The following general categories were coded: article characteristics, task characteristics, setting characteristics, and participant characteristics. Information was gathered by two coders: the primary author and a trained senior undergraduate research assistant. Discrepancies were discussed and resolved between coders. In rare cases when the resolution was unclear, Thomas Kwapil, Ph.D. (clinical psychologist), who has expertise in schizotypy spectrum research, was consulted. The two coders had an 82% initial agreement rate for 38 samples that were dual-coded.

Analyses

General Analysis Strategy. Analyses were planned *a priori*—that is, before analyses were begun—and pre-registered at Open Science Framework. Analyses were run using Comprehensive Meta-Analysis (CMA) (Borenstein, Hedges, Higgins, &

Rothstein, 2015), Version 3, a software program designed to conduct meta-analyses. There are two main models used in meta-analysis: fixed-effects and random-effects. Briefly, a fixed-effect model assumes that variance in the average effect size is due to sampling error, whereas a random-effects model assumes it is due to systematic error (Cooper, 2010). Based on consultation with Paul Silvia, Ph.D. (social psychologist), who has expertise in meta-analysis, a random-effects model was selected for the proposed meta-analysis. Random-effects models are more commonly used in psychology and there is reason to believe that error may vary systematically across the studies reviewed in this project.

Between-group effect sizes on task performance were calculated as or converted to Hedges *g* and within-group effect sizes on symptom-task associations were calculated as or converted from bivariate Pearson correlations (*r*) to Fisher's *z* scores. Effect sizes from each study were weighted by inverse variance and summarized using CMA software. 95% confidence intervals were computed to estimate the variance for each average effect size.

Heterogeneity among effects was analyzed using the *Q*-statistic, tau^2, and I^2. The *Q*-statistic is based on a chi-squared distribution and represents the ratio of observed variation among studies to sampling error within studies. Tau^2, which estimates the variance of true effects, and I^2, which estimates the ratio of observed variance representing true differences among studies, are both derived from *Q* (Borenstein, 2009). For subgroups, analyses were calculated with separate estimates of tau^2 for each subgroup. This statistical method is most appropriate for studies and subgroups in which

effect sizes are assumed to vary naturally. Throughout data analysis, best practices according to Borenstein (2009) and Higgins and Green (2011) were followed.

Although a large number of analyses were conducted, analyses were limited to those that were planned *a priori*. Post-hoc alpha adjustment procedures (e.g., Bonferroni correction) were not used so alpha was set at 0.05. As noted by O'Keefe (2003), post-hoc alpha adjustment procedures have three serious limitations: 1) they reduce statistical power, 2) the principles justifying their use are not consistently applied, and 3) the consistent application of the principle leads to undesirable research practices. Nevertheless, results are interpreted with caution, giving more credence to findings in the predicted direction and those showing consistent patterns across the literature. Due to the issues with null hypothesis significance testing, hypotheses generally focus more on the size of predicted effects than on significance, using suggested effect size interpretations as general guidelines (Cohen, 1988; Cohen, 1992; Sawilowsky, 2009). For Cohen's mean difference, *d* of 1.20 is considered very large, *d* of 0.80 to 1.19 is considered large, *d* of 0.50 to 0.79 is considered medium, *d* of 0.20 to 0.49 is considered small, and *d* of less than 0.20 is considered very small/negligible. For Pearson's bivariate correlations, *r* of |0.50| and above is considered large, *r* of |0.30| to |0.49| is considered medium, *r* of |0.10| to |0.29| is considered small, and *r* of less than |0.10| is considered very small/negligible.

Multiple comparisons from within the same group of participants were handled by halving the sample size for that group based on the number of comparisons, in order to include both sets of comparisons without "double counting" that group. This method is recommended by meta-analysis experts for dealing with multiple groups from one study

(Higgins & Green, 2011). For example, for studies in which a schizophrenia group was compared to two different psychiatric comparison groups, the effect size for the schizophrenia group versus the first psychiatric comparison group was entered with half of the schizophrenia group sample size, and the same for the schizophrenia group versus the second psychiatric comparison group.

All outcomes were assessed for primary comparisons between the schizotypy-spectrum group and healthy controls for short and long delay conditions: errors and reaction time for all four trial types, and d'context. To limit the number of analyses for primary psychiatric comparison, subgroup, moderation, and meta-regression analyses, only key outcomes were examined: typically AY errors, BX errors, and d'context combined across delay conditions, which are considered the most critical variables for assessing CI. For symptom correlations, only BX errors and d'context (again, combined across delay conditions) associations were assessed since these are the best indicators of CI impairment. Although these outcomes all tap CI ability, it was decided *a priori* to assess them both (rather than selecting one or averaging outcomes) for theoretical reasons. Intact CI could theoretically lead to greater AY errors in healthy controls when the A cues preparation for response, whereas disrupted CI may lead to greater BX errors in those at risk for schizophrenia when the X probe triggers a response if cue context is not held in mind. Thus, the results for these two outcomes were examined in parallel. d'context was selected because it is thought to yield a better estimate of sensitivity to context than independent error rates (Servan-Schreiber et al., 1996). Therefore, even though d'context and BX errors are not independent (BX errors are used to calculate

d'context), they are not redundant and both contribute helpful information based on theoretical considerations. Results for d'context and BX errors were run separately so as not to violate the assumption of independent data in meta-analysis.

Note that candidate subgroups and moderating factors were decided before data was collected (although some of the originally planned analyses could not be conducted due to limited number of studies). Candidate analyses were also planned *a priori* but were modified based on power estimations according to availability of the data collected; thus the final subset of pre-planned analyses was selected after studies were coded but before analysis was begun.

There is no consensus on appropriate cutoffs for planning subgroup analyses; however, in line with recommendations from Borenstein (2009) and Fu et al. (2011), planned subgroup analyses were only carried out when there were at least 6 studies per subgroup or characteristic being assessed. Although the participant sample sizes in this literature are often small, key effect sizes for this literature tend to be moderate to large. Furthermore, many of the planned analyses far surpassed this sample cutoff; therefore, I estimated that most of the analyses described below would have adequate power. Note, however, that actual power in a random-effects model will also be determined by the variance across studies (Borenstein, 2009); therefore, I kept in mind that when heterogeneity among studies is large, the study may be underpowered to detect subtle effects.

Medication status of patients with schizophrenia was initially proposed as a moderating factor but there were not enough studies examining unmedicated patients to

provide adequate power for these analyses; additionally, medication status and number of illness episodes were confounded (i.e., unmedicated patients were typically experiencing their first episode of schizophrenia). I also hoped to examine results across short and long delay conditions separately for key outcomes in subgroup, symptom correlation, moderation, and meta-regression analyses but there were not enough studies to allow for this so results were collapsed across delay for most analyses (see below for exceptions).

Specific Analyses and Detailed Hypotheses

Primary Analyses, Between-Group

Healthy Control Comparisons. All schizotypy-spectrum groups (all schizophrenia-spectrum diagnoses, unaffected relatives, and psychometrically-identified schizotypy groups) together were compared against healthy controls for AX, AY, BX, and BY errors and reaction time, and d'context for short (<3500ms) and long (>3500ms) length of cue-probe intervals.

The schizotypy-spectrum group was expected to demonstrate worse performance than healthy controls on AX[1] and BX trials at the level of a medium effect size, and somewhat better performance on AY trials at the level of a small effect size. No substantial effect (i.e., effect size in the very small range) was hypothesized for BY

[1] An initial hypothesis of no difference in AX errors was posted to Open Science Framework. However, this hypothesis was made in error as it misrepresents CI theory, which would predict deficits in both BX and AX errors since both trial types tap CI ability (MacDonald, Goghari et al., 2005).

errors, consistent with theory of a specific deficit in CI. It was expected that the schizotypy-spectrum group would demonstrate slower performance (longer reaction times) across all trial types. The schizotypy-spectrum group was hypothesized to have lower d'context scores compared to healthy controls, at the level of a medium effect size.

Psychiatric Comparisons. All patients with schizophrenia (chronic and first-episode) were compared on AY errors, BX errors, and d'context against all psychiatric comparison groups. Patients with schizophrenia were expected to demonstrate worse performance on BX trials and d'context scores compared to psychiatric comparison patients, at the level of a small to medium effect. No substantial effect was predicted for AY errors.

Subgroup Analyses, Between-Group

Diagnostic Groups. The magnitude of the difference between patients with schizophrenia and healthy controls on AY errors, BX errors and d' prime context was compared to the magnitude of the difference between the at-risk group and healthy controls on the same measures.

Patients with schizophrenia were expected to demonstrate poorer performance than healthy controls on BX trials and d'context scores, at the level of a medium to large effect size, and somewhat better performance on AY trials compared to healthy controls at the level of a small to medium effect. It was hypothesized that the at-risk group would show somewhat worse performance on BX trials and d'context scores—though these effects were predicted to be small—and no difference in AY errors compared to healthy controls. Thus, it was predicted that effect sizes for indirect comparison analyses

examining BX errors and d'context would be larger in patients with schizophrenia versus healthy controls compared to the at-risk group versus healthy controls.

Diagnostic Groups by Illness Length. The magnitude of the difference between patients with chronic schizophrenia and healthy controls on AY errors, BX errors, and d'context was compared to the magnitude of the differences between first-episode patients and healthy controls, and between the at-risk group and healthy controls on the same measures.

It was predicted that patients with schizophrenia (both chronic and first-episode) and the at-risk group would show the same pattern of results described above (for diagnostic groups). In terms of effect sizes for d'context and BX errors for indirect comparisons, it was expected that the difference between patients with chronic schizophrenia and healthy controls would be greater than the difference between first-episode patients and healthy controls, which in turn would be larger than the difference between the at-risk group and healthy controls.

Symptom-Task Correlations, Within-Group. Meta-analyses were run on within-group correlations of positive, negative, and disorganized symptoms with d'context and BX errors in patients with schizophrenia (chronic and first-episode). Symptom-task correlations were conducted: 1) using all correlations from authors' original symptom dimension conceptualizations (or from our conceptualizations when no conceptualization was provided and/or raw data was available) and 2) using only correlations from symptom dimension conceptualizations that were decided *a priori* to appropriately measure the symptom constructs (conceptualizations were made in consultation with Dr.

Kwapil; see Appendix D for details). These subset analyses were planned because there appeared to be inappropriate symptom conceptualizations in the initial pool of studies examined for inclusion, as described earlier.

Symptom-Task Correlation Hypotheses. It was predicted that disorganized and negative symptoms would be associated with worse performance for both outcome measures, that is, positively associated with BX errors (lower accuracy) and negatively associated with d'context (worse CI). No substantial effect was predicted for the association of positive symptoms with BX errors or d'context. It was predicted that disorganization symptoms would be more strongly associated with both BX errors and d'context than either positive or negative symptoms. Additionally, it was expected that associations would be stronger in analyses using the subset of appropriate symptom conceptualizations: a small effect was predicted for negative symptom associations with task performance and a medium effect for disorganization symptom associations.

Moderation Analyses, Between-Group. Moderation analyses were conducted by grouping data according to a moderating variable and calculating the Q-statistic to determine whether effect sizes from the different groups are homogenous—that is, whether the grouping variable explains a significant amount of variance in the average effect size—according to best practices in the field (Borenstein, 2009; Higgins & Green, 2011).

Moderation analyses were conducted to examine whether between-group effect sizes for all schizotypy-spectrum groups compared to healthy controls for AY errors, BX errors, and d'context were affected by cue duration (dichotomized at 500 ms) and cue-

probe interval (dichotomized at 3500ms). For the cue-probe interval analyses (categorically in this section, as well as continuously in the section below), short and long delay outcomes were not averaged; studies that examined both long and short cue-probe interval delays were entered with two sets of outcomes at half the sample size each. Because there were not enough studies for each diagnostic group using the DPX, I only assessed AX-CPT versus DPX task as a moderator of effect size for the difference between patients with chronic schizophrenia and healthy controls (again, for AY errors, BX errors, and d'context) to avoid confounding task effects with diagnostic effects. These analyses were planned to help inform task parameters for future research.

Moderation Analysis Hypotheses. For cue duration analyses, it was predicted that effect sizes for schizotypy-spectrum groups versus healthy controls would be slightly larger for the short cue condition (cues displayed for 500ms or shorter) compared to the long cue condition (cue displayed for 1000ms), but that these effects would be small at best. For cue-probe interval analyses, it was predicted that effect sizes for schizotypy-spectrum groups versus healthy controls would be larger for the long delay condition (3500ms or longer) than the short delay condition (less than 3500ms), at the level of a small to medium effect. For task type analyses, it was predicted that effect sizes for patients with schizophrenia versus healthy controls would be slightly larger for the DPX compared to the AX-CPT, but that these effects would be small at best.

Meta-Regression, Between-Group. Meta-regression analyses were conducted to assess the extent to which key task performance outcomes were predicted by a pre-planned set of continuous variables (Borenstein, 2009). Between-group effect sizes for all

schizotypy-spectrum groups versus healthy controls for AY errors, BX errors, and d'context were regressed on the following continuous variables: cue-probe interval duration (short and long delay outcomes not averaged; see above) and total number of trials. These analyses were planned to assess whether there were clear cut-points in the data relevant to recommendations for future research (e.g., "Large effects were observed in studies that ran at least X number of trials."). Between-group effect sizes for all patients with schizophrenia versus healthy controls for AY errors, BX errors, and d'context were regressed on patients' length of illness.

Meta-Regression Hypotheses. Cue-probe interval duration, total number of trials, and length of illness in patients with schizophrenia were expected to positively predict effect sizes for the schizotypy-spectrum group versus healthy controls. That is, effects were hypothesized to be larger for longer delays, more trials, and more chronic patients.

Interpreting the Evidence

There are a number of ways to interpret results in light of potential file-drawer effects and *p*-hacking. *P*-hacking occurs when researchers obtain significant results by various post-hoc methods; for example, by stopping data collection as soon as significant results are reached, altering statistical analyses selected, or only reporting analyses that are statistically significant (Simonsohn, Nelson, & Simmons, 2014). Funnel plots, the trim-and-fill method, and Simonsohn's *p*-curve (Simonsohn, 2017) were used to assess for asymmetry and skew in the distribution of findings, as rough estimations of possible publication bias and selective reporting to aid in the interpretation of the meta-analytic findings.

Funnel Plots and the Trim-and-Fill Method. Using Comprehensive Meta Analysis software, funnel plots were created by plotting effect size (using a random effects model) on standard error, which is the recommended choice of axis to display a study's precision. The graphs are inverted so studies that are more precise (i.e., have the largest sample sizes) are shown at the top of the graph. A vertical line is drawn through the summary effect estimate and diagonal lines are drawn representing the expected 95% confidence intervals around the summary effect estimate: summary effect estimate + 1.96*standard error and summary effect estimate – 1.96*standard error. When between-study heterogeneity is not significant, 95% of study effects fall within the bounds of these confidence interval lines (Sterne & Egger, 2001). When a plot is asymmetrical and more than 5% of studies fall outside the confidence interval lines, it may be indicative of reporting bias, heterogeneity among studies, or chance. Reporting biases that can lead to asymmetrical funnel plots include publication bias, selective reporting of outcomes, or *p*-hacking (Sterne et al., 2011). However, asymmetrical funnel plots do not necessarily indicate inappropriate research practices: they can also reflect true variability among effects across samples, in which sample size has an influence on effect size (Sterne et al., 2011).

Using Comprehensive Meta Analysis software, Duval & Tweedie's (2000) trim-and-fill method was applied to the funnel plots. This method uses symmetry assumptions to estimate the number of missing studies, impute these missing values, and calculate an adjusted summary effect estimate incorporating the imputed values. This method can reveal the extent to which possible biases influence findings (if the funnel plot

asymmetry is indeed due to reporting bias) and can provide a more conservative estimate of the overall effect.

Funnel plots were examined and the trim-and-fill method was applied for key outcomes: AY errors, BX errors, and d'context for short and long delay conditions in primary analyses of schizotypy-spectrum groups versus healthy control groups, and for combined delay conditions in primary analyses of schizophrenia patient groups versus psychiatric comparison groups. Note that the choice of which outcomes to examine was not pre-registered but paralleled previous decisions to reduce the number of results presented by only examining key outcomes. Funnel plots and trim-and-fill techniques were not used to assess outcomes for subgroup, moderation, or meta-regression analyses since they are redundant; that is, the same outcomes are described in these types of analyses, just organized in a different manner.

The P-Curve. Simonsohn's (2017) *p*-curve application is a tool designed to detect *p*-hacking. It involves plotting out reported *p*-values from a set of pre-defined studies in the literature and interpreting the evidential value: an effect has evidential value if is not likely caused by *p*-hacking alone. Evidential value is determined by examining the shape of the *p*-curve, relying on the assumption that it is more likely to obtain small statistically significant *p*-values than large ones (Simonsohn, Nelson, & Simmons, 2014; Simonsohn, Simmons, & Nelson, 2014). This technique is an important complement to Duval and Tweedie's trim-and-fill method because trim-and-fill theory relies on the assumption that reporting biases arise from small effect sizes. In reality, statistical significance is often what drives reporting and publication in psychology

research (Simonsohn, Simmons et al., 2014); therefore, it is important to assess the distribution of *p*-values in addition to effect sizes.

When examining *p*-values that are less than 0.05, if the curve is skewed to the right—that is, if there are more low *p*-values (e.g., 0.001) than high *p*-values (e.g., 0.045)—evidential value is demonstrated. When the *p*-curve is skewed to the left—that is, if the number of *p*-values just below 0.05 are greater than the number of *p*-values much smaller than 0.05—it suggests that "intensive *p*-hacking" is likely present and the effect lacks evidential value (Simonsohn, Simmons et al., 2014; Simonsohn, Nelson et al., 2014). Specifically, a combination test probing both the half *p*-curve (*p*-values<0.025) and the full *p*-curve (*p*-values<0.05) is used. Evidential value is considered to be present if a) the half *p*-curve is skewed to the right at $p<0.05$, or b) both the half and full *p*-curve are skewed to the right at $p<0.10$ (Simonsohn, Simmons, & Nelson, 2015).

Analytical decisions were made before *p*-curve analysis began and were pre-registered. The reported *p*-values were plotted using Simonsohn's online *p*-curve application, version 4.052 (Simonsohn, 2017). The *p*-values were plotted onto separate *p*-curves for positive, negative, and disorganized symptom correlations with BX errors and with d'context. Authors' original symptom dimension conceptualizations were used (or our conceptualizations when no conceptualization was provided and/or raw data was available). Because the *p*-curve method is designed to assess for selective reporting, I only plotted *p*-values that were reported directly in the text (i.e., not those that corresponded to any values I calculated or received via correspondence with authors).

The *p*-values for symptom correlations were plotted by inputting *r* statistics into the online application. It is not appropriate to plot dependent effects on the same *p*-curve (Simonsohn, Nelson et al., 2014); therefore, decisions were made to handle multiple outcomes reported by a single study. Short and long delay conditions were plotted separately: one set of *p*-curves was created with *p*-values from short delay conditions and combined or unknown delay conditions, and a second set of *p*-curves was made with *p*-values from long delay conditions and combined or unknown delay conditions. Thus, these curves should be considered independently. See Appendix E for description of other issues with multiple outcomes and justification of methodological decisions.

It was decided *a priori* not to create *p*-curves for error patterns or d'context because the data were not well-suited for this purpose: most studies used ANOVA tests to assess global between-group effects and/or group by delay interactions, and then ran follow-up tests for simple effects. Only *p*-values obtained from directly testing simple effects should be included in a *p*-curve because using simple effects that were part of an interaction will bias the *p*-curve towards demonstrating evidential value even when there is none (Simonsohn, Nelson et al., 2014). Thus, it would not be appropriate to include studies that tested interactions; however, including those that did not would limit analyses to a small collection of studies that are perhaps unrepresentative of the broader literature. With these considerations in mind, *p*-curves were only created for symptom correlation effects within patients with schizophrenia and not for any between-group effects.

The Fail-Safe N. The fail-safe N estimates how many studies with null findings it would take to yield the calculated average effect size insignificant (Rosenthal, 1979).

However, the fail-safe N relies on null hypothesis significance testing, ignoring the effect sizes of hypothetical, unpublished studies (Higgins & Green, 2011) and is insensitive to *p*-hacking (Simonsohn, Nelson et al., 2014). Thus, it is not recommended for use in meta-analysis by the leading experts in the field and was not reported in the current study.

CHAPTER III

RESULTS

Outcomes of the Literature Search

A total of 47 independent samples were selected for final inclusion, displayed in Appendix F; 41 studies were included in meta-analyses and 6 studies were only included in summary point estimates According to guidelines from Preferred Reporting Items for Systematic reviews and Meta-Analyses statement (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009), a diagram of search results is presented as a flowchart in Figure 1, including reasons for study exclusion. 39 independent samples were included using patients with schizophrenia, 8 samples with unaffected relatives of patients with schizophrenia, 4 samples with individuals with SPD, 1 sample with those at ultra high risk to develop schizophrenia, and 1 sample with psychometric identification of schizotypy. Psychiatric comparison groups included non-schizophrenia psychotic patients: 5 samples of bipolar disorder with psychotic features and 4 samples of non-psychotic depressed patients.

Appendix B provides a detailed description of the *a priori* methodological decisions that were made for obtaining, reviewing, coding, and excluding studies; Appendix G provides additional information and methodological decisions pertaining to specific studies. Great care was taken to follow these pre-planned procedures and data were recorded and coded using the systematic and exhaustive coding sheet in Appendix

C. Note that both published and unpublished studies often lacked complete information including sample characteristics, results of statistical analyses, and descriptions of the task parameters. Repeated efforts were made to contact investigators during 7 weeks to obtain missing information. However, at times determinations about specific studies had to be made. For the sake of transparency, these are summarized in Appendix G.

As described in detail in Appendix B, the use of overlapping samples and reporting on data in multiple published articles was a substantial issue. Furthermore, authors often did not provide adequate descriptions of their samples or respond to direct communication; thus, unambiguous identification of sample overlap was not always possible. When unclear, a conservative approach was taken and the study was excluded. Overall, 20 samples were excluded due to confirmed or likely re-use of participants across articles. Note that inclusion of multiple studies reporting the same data runs the risk of artificially inflating conclusions about effects.

Demographic Information

Summary demographic information for all included studies is provided in Table 2. Schizotypy-spectrum groups included patients with chronic schizophrenia, patients with first-episode schizophrenia (medicated and unmedicated), individuals with SPD, those at ultra high risk, unaffected relatives of patients with schizophrenia, and those with psychometrically identified schizotypy. There was a large majority of patients with chronic schizophrenia in the schizotypy-spectrum group. Healthy control groups included those with no current (and often no past) mental illness. Psychiatric comparison groups included non-schizophrenia psychotic patients, non-psychotic depressed patients, bipolar

patients (with and without psychotic features), non-Cluster A personality disorders, and a psychometrically-identified "vulnerable to depression" group.

Point Estimates of Task Performance

Point estimates representing average mean and standard deviation weighted by sample size were computed in both short and long delay conditions for errors and reaction times across trial type and d'context. Graphs of these point estimates are presented in Figures 2 to 7. For visibility, error bars were not included on error and reaction time graphs, but precise weighted means and standard deviations can be found in Tables H1 to H3 in Appendix H.

Point estimates were computed for the following groups: first-episode unmedicated schizophrenia patients, first-episode medicated schizophrenia patients, chronic schizophrenia patients, individuals at-risk, healthy controls, and psychiatric controls. Note that these basic point estimates do not take into account between-study variance and are weighted by sample size rather than by inverse variance; thus observations of possible group differences may not be synonymous with meta-analysis findings. Nonetheless, these estimates combine data from a larger set of participants since several studies did not include control participants and thus could not contribute between-group effect sizes for the meta-analyses.

There were not enough studies reporting medication status across different schizophrenia groups to examine medication status as a moderator. The point estimates displayed in the figures show first-episode groups separately by medication status and thus offer a complementary view of the data; however, note that some estimates include

few samples or only a single study. In sum, these point estimates are observational and meant only to aid the reader in visualizing the magnitude of unstandardized deficits. Therefore, information on group differences will only be interpreted from meta-analysis results presented below.

Primary Analyses

Healthy Control Comparisons. A combined schizotypy-spectrum group (all patients with schizophrenia and at-risk schizotypy groups) was compared to the healthy control group across all outcomes. Results are presented in Tables 3 and 4; forest plots of d'context effects for short and long delay conditions are presented in Figure 8. Due to the large number of primary comparisons presented here, confidence intervals around effect estimates are not described in the text for this section but the reader is encouraged to consider confidence intervals from the tables.

Contrary to hypotheses, the schizotypy-spectrum group showed worse performance than healthy controls on almost every outcome. They demonstrated significantly more errors across all trial types for short delay intervals, and across AX, BX, and BY trials for long delay intervals; there was no significant between-group difference for AY errors in the long delay condition. In line with hypotheses, the schizotypy-spectrum group showed significantly longer reaction times across all trial types and lower d'context for both short and long delay intervals. Very small effects were found for AY and BY errors (long delay). Small effect sizes were found for AY and BY errors (short delay), BX reaction time (long delay), and AY reaction time (short and long delay). As predicted, medium effects were found for BX reaction times and d'context

(short delay), AX and BX errors, and for AX and BY reaction times (short and long delay). A large effect was found for d'context in the long delay condition. Significant heterogeneity across studies was found for all outcomes, except short delay BX errors.

Confidence intervals around effects were indirectly compared to make observational inferences about effect size differences among trial types. For short delay trials, the confidence intervals around AX errors slightly overlapped with those for AY and BY errors; the confidence intervals for BX errors showed no overlap with those for AY and BY errors. For long delay trials, neither AX errors nor BX errors' confidence intervals overlapped with those for AY and BY errors, suggesting that the effects for CI-critical trials were larger than for trials that tap preparatory or general ability.

Confidence intervals were also used to make post-hoc inferences about the influence of cue-probe delay on task performance. See results sections IV and V below for planned assessments of cue-probe interval as a categorical moderator and a continuous predictor of key outcomes; the current indirect observations provide complementary information across a broader range of outcomes. There were no significant differences in magnitude of between-group effect sizes between short and long delay conditions for any error, reaction time, or d'context outcomes.

In summary, the schizotypy-spectrum group generally performed worse than the healthy control group. Significant between-study variance in these effects was observed. As predicted, effects for AX and BX errors were generally larger than those for AY and BY errors, suggesting a specific deficit in context processing. Post-hoc observations suggested that the magnitude of between-group effects may not differ between delay

conditions but these comparisons should be interpreted with caution and in light of subsequent analyses.

Psychiatric Comparisons. A combined schizophrenia group was compared with a combined psychiatric comparison group. Results are presented in Table 5; a forest plot of d'context effects is presented in Figure 9. As predicted, the schizophrenia group showed significantly lower d'context than psychiatric comparison groups, at the level of a medium effect size, with confidence intervals ranging from a small to large effect and significant heterogeneity across studies. Patients with schizophrenia made significantly more BX errors than other psychiatric patients at the level of a small effect size, with confidence intervals ranging from a very small to small effect. Finally, there was no overall between-group difference in AY errors: a very small effect was observed with confidence intervals in the very small effect range. There was no significant heterogeneity across studies for BX and AY errors. In sum, the schizophrenia group demonstrated worse CI ability compared to other psychiatric groups and the magnitude of these effects conformed to hypotheses.

Subgroup Analyses

Diagnostic Groups. As planned, results were re-run dividing the schizotypy-spectrum group into subgroups of individuals with schizophrenia and those at risk. Results are presented in Tables 6 to 8; a forest plot of d'context effects is presented in Figure 10. As hypothesized, the schizophrenia group showed significantly lower d'context than healthy controls at the level of a large effect size, with confidence intervals ranging from a medium to large effect, and more BX errors at the level of a medium

effect size, with confidence intervals in the medium effect range. Contrary to predictions, patients with schizophrenia made more AY errors at the level of a small effect size, with confidence intervals ranging from a very small to small effect. Significant heterogeneity across studies was observed for all effects in the schizophrenia group.

The pattern and magnitude of CI deficits in the at-risk group conformed to hypotheses: compared to healthy controls, the at-risk group demonstrated significantly lower d'context at the level of a small effect size, with confidence intervals ranging from a small to medium effect, and more BX errors at the level of a small effect size, with confidence intervals ranging from a very small to medium effect. There were no between-group differences for AY errors: a very small effect was observed, with confidence intervals ranging from a very small to small effect. There was no significant heterogeneity across studies for any effects in the at-risk group.

Indirect comparisons showed that when each group was compared to healthy controls, the schizophrenia group had stronger effects than the at-risk group for AY errors, BX errors, and d'context, as predicted. Indirect comparisons are observational and should be interpreted cautiously. Additionally, a significant amount of unexplained variance remained in the model after accounting for diagnostic status. In sum, as expected both patients with schizophrenia and at-risk individuals showed CI deficits, and indirect comparisons suggested that these effects are more pronounced in patients.

Diagnostic Groups by Illness Length. As planned, results were re-run dividing the schizotypy-spectrum group into subgroups of patients with chronic schizophrenia, first-episode patients, and individuals at risk to develop the disorder. Results are

presented in Tables 9 to 11. Results for at-risk groups versus healthy control groups were identical to those presented in the previous section above (Section 1, Diagnostic groups). Findings for chronic and first-episode patients generally conformed to hypotheses. When compared to healthy controls, both patient groups showed significantly lower d'context at the level of a large effect size, with confidence intervals ranging from a medium to large effect; and more BX errors at the level of a medium effect size, with confidence intervals ranging from a medium to large effect. Contrary to predictions, however, the chronic group made more AY errors than controls at the level of a small effect, with confidence intervals ranging from a very small to medium effect. The first-episode group did not make significantly more AY errors than controls: a very small effect was observed, with confidence intervals ranging from a very small to small effect. There was significant heterogeneity across studies for all effects in the chronic group and for BX errors in the first-episode group.

The overall strength of effects differed significantly across groups for AY errors, BX errors, and d'context; however, there was still unexplained variance left in the model after accounting for diagnostic grouping by illness length. Using indirect comparisons with healthy controls, the chronic patient group showed stronger effects than the at-risk group for AY errors and d'context. The first-episode patient group showed a stronger effect than at-risk group for d'context. Although the overall group effect for BX errors was significant, confidence intervals around the at-risk group estimate overlapped slightly with confidence intervals for both the chronic and first-episode groups.

Unexpectedly, there was no difference in strength of effects between the chronic and first-episode patients groups for any outcomes.

In sum, hypothesized CI deficits were found in those at-risk for developing schizophrenia and in patients with both first-episode and chronic schizophrenia. Indirect comparison indicated that, contrary to predictions, the magnitude of deficits in first-episode and chronic patients was similar relative to healthy controls.

Symptom-Task Correlations

Authors' Conceptualizations and Our Conceptualizations. Within-group correlations between symptom dimensions and key task outcomes were examined for patients with schizophrenia. Contrary to expectations, results did not appreciably differ when analyses were re-run with a subset of studies corresponding to our narrower, *a priori* conceptualizations of symptom dimensions: effects were slightly—but not significantly—larger using our symptom conceptualizations. Therefore, results are only described in the text for the full set of studies, corresponding to the authors' original conceptualizations (or to our conceptualization when none was provided or when raw data was available). See Tables 12 to 14 for results from authors' conceptualizations and Tables 15 to 17 for results from our conceptualizations.

Forest plots of d'context correlations with symptom dimensions are presented in Figures 11 to 13. Based on authors' conceptualizations, positive symptoms did not show significant associations with d'context or BX errors: as expected, very small effects were found, with confidence intervals in the small to very small effect range. As hypothesized, negative symptoms were significantly correlated with d'context and BX errors at the level

of a small effect size, with confidence intervals ranging from a very small to small effect. No significant heterogeneity across studies was observed for positive or negative symptom correlations. Disorganization symptoms showed significant associations with d'context and BX errors at the level of a small effect size, with confidence intervals ranging from a very small to medium effect. Note that these findings were in the predicted direction but the associations were somewhat smaller than expected. All significant symptom effects indicated worse performance in the schizophrenia group: disorganization and negative symptoms were negatively correlated with d'context scores and positively correlated with BX errors. Significant heterogeneity was present across studies for disorganized symptom correlations.

In sum, disorganized and negative symptoms showed small associations with CI deficits in patients with schizophrenia; the positive symptom dimension showed negligible associations consistent with hypotheses. However, contrary to hypotheses, the magnitude of effects did not significantly differ among the three symptom dimensions.

Moderation Analyses

Cue Duration. Length of cue presentation was assessed as a categorical moderator of between-group outcomes for a combined schizotypy-spectrum group versus healthy controls. Results are presented in Tables 18 to 20. Long cues (those lasting for 1,000ms) showed significantly stronger effects than short cues (those lasting 500ms or less) for AY errors. Note that the schizotypy-spectrum group made more errors than the healthy control group so cue duration moderated the effect in the opposite direction from

what was predicted. Cue duration did not moderate effects for BX errors or d'context, consistent with predictions that effects would be small at best.

Cue-Probe Interval. Length of cue-probe delay was assessed as a categorical moderator of outcomes between schizotypy-spectrum and healthy control groups. Results are presented in Tables 21 to 23. Cue-probe interval moderated outcomes for AY accuracy: short cue-probe intervals (less than 3500ms) showed significantly stronger effects than long cue-probe intervals (3500ms or longer) for AY errors. Again, these AY errors were in the opposite direction from what was hypothesized. Significant variance in AY errors across studies still remained with cue-probe duration in the model. Contrary to predictions, cue-probe interval did not moderate BX or d'context outcomes.

Task Type. Use of AX-CPT versus DPX task was assessed as a categorical moderator of outcomes between patients with chronic schizophrenia and healthy control groups; as noted earlier, analyses were limited to these groups because there were not enough studies using the DPX with other schizotypy-spectrum groups. Results are presented in Tables 24 to 26. Task type did not moderate the magnitude of effect between patients with chronic schizophrenia and healthy controls for AY errors, BX errors, or d'context, consistent with predictions that effects would be small at best.

In summary, long cues and short cue-probe delays were associated with stronger effects for AY errors (with the schizotypy spectrum group making more errors than controls). Note, however, that these variables are moderating effects of negligible to small magnitude *against* the predicted direction. None of the variables moderated BX

error or d'context effects, suggesting that schizotypy-spectrum CI deficits are robust across the examined task parameters.

Meta-Regression

Total Number of Trials. Total number of trials run in the experiment was assessed as a continuous predictor of outcomes between schizotypy-spectrum and healthy control groups. All results are presented in Tables 27 to 29. A graph of number of trials regressed on AY errors is presented in Figure 14 and a graph of number of trials regressed on d'context is presented in Figure 15. Number of trials explained significant variance in the magnitude of effect for AY errors and d'context; however, significant variance still remained with this predictor in the model. Number of trials was negatively associated with these effects, meaning that as the number of trials increased, studies tended to find smaller positive effects for AY errors and stronger negative effects for d'context (each moving more in the predicted direction). Contrary to hypotheses, number of trials did not predict the magnitude of the between-group effect for BX errors.

As seen by examining the regression lines on the graph in Figure 14, the AY error effect approaches zero around 300 trials. In Figure 15, the regression line for d'context predicts medium effects for studies using 250 trials or less and large d'context effects for studies using more than 250 trials.

Cue-Probe Interval. Cue-probe delay was assessed as a continuous predictor of outcomes between schizotypy-spectrum and healthy control groups. Results are presented in Tables 30 to 32. Contrary to hypotheses, duration of the cue-probe interval did not predict effect magnitudes for AY errors, BX errors, or d'context.

Length of Illness. Length of illness in patients with schizophrenia (both chronic and first-episode) was assessed as a continuous predictor of outcomes between patients and healthy control groups. Results are presented in Tables 33 to 35. Patients' length of illness did not predict effect size for AY errors, BX errors, or d'context, contrary to predictions.

In sum, total number of trials predicted outcomes for AY errors and d'context, with more trials associated with effects more strongly in the direction of CI deficits. Cue-probe duration and length of illness were not significant predictors of CI outcomes.

Interpreting the Evidence

Funnel Plots and the Trim-and-Fill Method

Primary Analyses: Healthy Control Comparisons. Funnel plots were examined for AY errors, BX errors, and d'context scores in short and long delay conditions for effects in all schizotypy-spectrum groups versus healthy controls. The trim-and-fill method was applied using a random effects model for all analyses. Trim-and-fill results are presented in Table 36, including estimated confidence intervals around adjusted effect sizes. Three to nine values were imputed per outcome. None of the adjusted Hedges' g estimates differed significantly from (unadjusted) observed estimates, based on overlapping confidence intervals around the estimates. A funnel plot of d'context effects for short and long delay conditions including imputed values is presented in Figure 16.

Primary Analyses: Psychiatric Comparisons. Funnel plots were examined and the trim-and-fill method was applied for AY and BX errors, and d'context scores (combined across delay conditions) for effects in schizophrenia groups versus psychiatric

comparison groups. Results are presented in Table 37. Zero to three values were imputed per outcome and none of the adjusted estimates differed significantly from observed estimates. A funnel plot of d'context effects including imputed values is presented in Figure 17.

Symptom Correlations. Funnel plots were examined and the trim-and-fill method was applied for disorganized, negative, and positive symptom correlations with BX errors and d'context scores (combined across delay conditions) within patients with schizophrenia. Results are presented in Table 38. Zero to three values were imputed per outcome and none of the adjusted estimates differed significantly from observed estimates. A funnel plot of disorganization symptom correlations with d'context including imputed values is presented in Figure 18.

Simonsohn's P-Curve. Separate *p*-curves were created for symptom correlations of positive, negative, and disorganized symptoms with BX errors and d'context at short and long delay. Results are presented in Tables 39 and 40. None of the generated *p*-curves were significant for *p*-hacking. *P*-curves for disorganized and negative symptom correlations with d'context demonstrated evidential value for both short and long delay conditions, meaning that there is evidence these effects are not solely due to *p*-hacking. All other p-curves either could not be generated because all *p*-values were greater than 0.05 or results were inconclusive, meaning that evidential value could neither be established nor ruled out. Because there were few significant *p*-values per condition, power was low for these inconclusive analyses, ranging from 5% to 53%. Overall, there was no evidence that symptom correlation findings were solely due to *p*-hacking.

CHAPTER IV

DISCUSSION

Summary of Findings

Impaired CI ability is proposed to be a specific deficit in schizophrenia that presents before the onset of psychosis and may be implicated in the development of certain schizotypic symptoms. Numerous studies have used the CNTRACs tasks to assess CI in patients with schizophrenia, but fewer have examined CI in non-psychotic manifestations of schizotypy. The literature to date is limited by the use of small samples (which is not unusual in studies of patients with severe mental illness), varying task parameters, and heterogeneous patient characteristics that complicate interpretation of findings (especially from individual studies). Meta-analysis is particularly suitable for mitigating many of these limiting factors. Thus, the magnitude of CI impairment in schizotypy-spectrum groups and the impact of variables thought to moderate performance outcomes were examined using meta-analysis. The current study was the first comprehensive meta-analysis to examine CI impairment on AX-CPT and DPX tasks in schizotypy-spectrum psychopathology. Furthermore, the study followed best practices for meta-analytic review described by experts in the field (Borenstein, 2009; Higgins & Green, 2011).

A final total of 47 independent samples was included for point estimates of task outcomes, between-group comparisons, and within-group associations of people with

schizophrenia and schizoaffective disorder (chronic, first-episode, medicated, and unmedicated), SPD, ultra high-risk to develop schizophrenia, unaffected relatives of patients with schizophrenia, and schizotypy identified using psychometric methods. Meta-analysis results supported predicted levels of CI impairment across the schizotypy-spectrum. As hypothesized, deficits in patients with schizophrenia were substantial and, unexpectedly, appeared stable across illness duration. As expected, the at-risk groups showed milder CI disruption compared to healthy controls than patients with schizophrenia. Consistent with hypotheses, disorganized and negative symptoms were negatively correlated with CI ability, whereas positive symptoms were unrelated. Larger between-group effects were found when more trials were given but, contrary to predictions, no other task parameters appreciably predicted or moderated outcomes.

Results did not differ when the trim-and-fill method was used to adjust effect size estimates. Further, *p*-curve analyses showed no evidence that symptom correlation findings were due solely to *p*-hacking. Although reporting biases can never be fully ruled out, they do not seem to have grossly impacted findings from this literature, which supports the validity of the current meta-analysis results.

Interpretation of Results

When CI abilities were examined across the entire schizotypy-spectrum, the schizotypy group performed worse than healthy controls on most task outcomes. It is important to demonstrate differential patterns of impairment in psychopathology—and specifically, in schizophrenia—to distinguish a specific cognitive impairment from generalized performance deficits due to issues such as inattention and amotivation

(Chapman & Chapman, 1973; MacDonald & Carter, 2002). Accordingly, meta-analysis results revealed substantial deficits on AX and BX trials in the schizotypy-spectrum group that surpassed errors made on BY trials. Since AX and BX trials are thought to tap CI ability (MacDonald, Goghari et al., 2005) and BY trials are considered a rough estimator of general ability to perform the cognitive task (and hence, generalized impairment), these differential effect sizes are consistent with theory of a specific cognitive deficit in CI.

Negligible to small effects were found for AY accuracy, which theoretically reflects maintenance of the task goal. The schizotypy-spectrum group made slightly more AY errors than the healthy control group. It was hypothesized that the schizotypy-spectrum group would make *fewer* AY errors, thus, these findings were against the predicted direction. AY errors were proposed to reflect intact context processing, since good maintenance of task goals would theoretically result in the A cue leading participants to prepare response (Barch & Braver, 2009). The proposed crossover effect—in which patients with schizophrenia would perform worse than controls on BX trials but better on AY trials—is theoretically compelling but has not received strong or consistent empirical support.

Only two published studies have found patients with schizophrenia to out-perform controls on AY trials (Barch et al., 2001; MacDonald et al., 2003), and the current meta-analysis showed worse performance across the board in schizotypy-spectrum groups. One group of researchers has suggested that greater AY errors in schizophrenia may reflect difficulty with response inhibition (MacDonald, Goghari, et al., 2005), which could

explain why they did not globally out-perform healthy controls. Thus, it seems more likely that the BX-AY discrepancy constitutes a relative difference within the schizotypy-spectrum psychopathology—as has often been demonstrated in the literature (e.g., Barch et al., 2004; Henderson et al., 2012)—rather than an absolute difference between groups. The fact that individuals on the schizophrenia spectrum perform worse on BX trials than on AY trials is indicative of a specific deficit: that is, CI is more disrupted than response inhibition (or other cognitive abilities tapped by the task).

Differentiating CI Deficits in the Schizotypy Spectrum. CI disruption is proposed to demonstrate diagnostic specificity, that is, to be impaired in schizophrenia beyond deficits that may be due to psychosis in general or to psychopathology more broadly (e.g., Cohen et al., 1999; MacDonald, Carter et al., 2005). To examine this hypothesis, we compared schizophrenia groups with a variety of other psychiatric patients, including those with non-schizophrenia psychosis. Disruptions in CI were more pronounced in patients with schizophrenia than in other psychiatric groups, indicating that poor CI may be specifically disrupted in schizophrenia beyond general psychiatric impairment.

It follows logically that CI deficits would be stronger in more severe manifestations of schizotypy-spectrum psychopathology; however only a handful of studies have examined this hypothesis directly. Relative to patients with schizophrenia, unaffected family members have shown attenuated deficits in some studies (Delawalla et al., 2006; MacDonald et al., 2003; Richard et al., 2013) and comparable levels of impairment in other studies (MacDonald, Goghari et al., 2005; Poppe et al., 2015).

Niendam and colleagues (2014) found no difference between ultra high-risk participants and first-episode patients. However, no published studies have compared individuals with SPD or psychometrically identified schizotypy to patients with schizophrenia. Thus, to test the hypothesis that patients will be more impaired than at-risk groups across a broader range of studies, we examined CI deficits among different schizotypy-spectrum subgroups according to diagnostic status and illness length using indirect comparisons. Indirect comparisons should be interpreted with caution, but suggested that patients with schizophrenia have more severe CI deficits than at-risk individuals. In terms of unstandardized errors, results translated to patients and at-risk groups making approximately 4% more BX errors than healthy controls (with a larger standardized effect in patients by indirect comparison due to a narrower confidence interval). It is important to note that individuals who have never experienced psychosis, and most of whom will never develop schizophrenia, are still showing significant, albeit attenuated, cognitive impairment. This supports theories that CI may be a precursor to schizophrenia, or—as some have proposed (Barch & Braver, 2009)—even a mechanism influencing its development.

To further analyze CI deficits in schizophrenia, we ran additional subgroup analyses with chronic patients, first-episode patients (those whose illness had lasted 1.5 years or less), and again with the same at-risk group. Results were similar for chronic and first-episode groups: both showed large CI deficits compared to healthy controls. Although indirect comparisons should be interpreted cautiously, these results were supported by meta-regression analyses showing that continuous measures of illness

duration did not predict CI outcomes. This is partially consistent with findings from Richard and colleagues' (2013) longitudinal study that showed that CI performance in first-episode patients was comparable between baseline and 1-year follow up for short delay conditions, but improved for long delay conditions. Current meta-analysis results (collapsed across delay conditions) suggest that CI deficits may be stable throughout the illness, but additional longitudinal designs examining changes from initial episodes across the course of the illness are required to clarify these relationships. Given that CI appears to be present premorbidly and at initial episodes, persists throughout the illness, and is not episode-limited, it may be a stable vulnerability indicator for schizophrenia (Nuechterlein et al., 1992), consistent with previous researchers' theories (Barch et al., 2003; Richard et al., 2013).

Current evidence of impaired context processing in schizotypy-spectrum psychopathology does not contradict any of Chan and Gottesman's (2008) requirements for the demonstration of an endophenotypic marker. Further, based on research to date, CI appears to fulfill 5 of the 6 criteria for an endophenotype: CI deficits are associated with illness, state-independent, found in non-affected family members, can be measured reliably, and shows diagnostic specificity. Although the current meta-analysis did not address heritability, previous research has shown that DPX performance is influenced by Val158Met COMT polymorphism, a gene associated with risk for schizophrenia (Lopez-Garcia et al., 2015; MacDonald, Carter, Flory, Ferrell, & Manuck, 2007). However, it remains unclear whether CI is more prevalent among affected versus unaffected relatives

of schizophrenia probands. Investigating this final criterion is an important next step if CI is to be established as an endophenotype.

In sum, poor CI is a strong candidate to be considered an endophenotypic marker of schizophrenia, but further research is required. If CI were shown to be an endophenotype, it would allow researchers to quickly and non-intrusively identify those purportedly at-risk to develop schizophrenia. This could aid in the identification of individuals for prophylactic intervention and would open the door for research investigating risk and protective variables that may be implicated in cognitive ability and development of schizophrenia-spectrum disorders.

CI and Schizotypy Symptom Dimensions. The schizotypy spectrum is multidimensional, with different symptom dimensions showing differential patterns of cognitive impairment; thus, outcomes were examined separately for each symptom dimension. Within the schizophrenia group, disorganized and negative symptoms showed inverse correlations with CI, whereas positive symptoms were unassociated. Correlations between CI ability and the disorganized symptoms were smaller than expected. This may, in part, reflect that assessment of CI across symptom dimensions has not been a primary goal of most studies. Much of the literature on the topic includes correlations that were run in mixed symptom patient groups with varying levels of each symptom type, using conceptualizations driven largely by measures rather than by theory. We propose that future studies should be directly designed to assess symptom correlations with CI, beginning with a strong theoretical conceptualization of each symptom dimension, valid operationalization and measurement of these symptoms, recruitment of schizotypy-

spectrum groups that have strong representation of each symptom dimension, and examination of these symptom-CI relationships in an adequately powered sample.

Despite the shortcomings of the literature on multidimensional expression of CI in schizotypic psychopathology, the current findings of disorganized and negative symptom associations with CI indicate that these relationships are robust. These differential symptom correlations lend support to research describing distinct cognitive versus emotional pathways for schizophrenia (e.g., Myin-Germeys et al., 2002) and are consistent with neuroscience theories and reviews connecting neurocognitive impairment with disorganized and negative symptoms (Barch et al., 2001; Cohen et al., 1999; Goghari et al., 2010; Yoon et al., 2008).

CI associations in the literature have been particularly prominent for disorganized symptoms. For example, MacDonald, Carter, and colleagues (2005) showed that within patients with schizophrenia, diminished prefrontal activity following B-cues was associated with greater BX errors, as well as greater disorganized symptoms (showing stronger associations than with negative or positive symptoms). The authors concluded that cognitive disorganization in schizophrenia is related to impairment in the top-down ability to represent and maintain appropriate context of the B-cue, which should cue participants not to respond to any letter that follows, even the letter X (MacDonald, Carter et al., 2005). Furthermore, Yoon and colleagues (2008) found that disorganized symptoms were related to decreased activity in the DLPFC and diminished connectivity to a broader neural network implicated in cognitive control. The authors theorized that altered DLPFC function and connectivity may lead to poor top-down control underlying

impaired task performance and behavioral symptoms of disorganization (Yoon, et al., 2008).

Early phenomenologists (e.g., Conrad, 1958; Matussek, 1952) proposed connections between positive symptoms and the dis-integration of an individual from their environmental context. These purported associations described in the phenomenological literature were not supported by the results of the present meta-analysis: positive symptoms showed negligible correlations with CI. This seems to indicate that the cognitive-perceptual disruptions associated with positive symptoms of schizophrenia may not involve the same processes as CI assessed with cognitive tasks such as the AX-CPT and DPX.

There is further debate on the extent to which neural mechanisms of CI are broadly distributed versus localized in the specific regions of the brain, such as the DLPFC (and its interactions with other regions). Overall, there are gaps in our understanding of the integration of "context" across different levels of processing. Whether impairments found on neurocognitive tasks are synonymous with behavioral manifestations of poorly integrated context remains an empirical question. Longitudinal studies integrating CI measurements across multiple domains should help reveal confluence or discontinuity of the processes proposed to underlie neural, cognitive, perceptual, and behavioral outcomes, and could help scientists in different fields of research refine the language they use when describing these phenomena.

Factors Moderating CI Deficits. Effects of different task parameters were examined using moderation and meta-regression analyses. When comparing schizotypy-

spectrum participants with healthy controls, categorical measures of task type and cue duration did not moderate BX accuracy or d'context effects, nor did categorical or continuous measures of cue-probe interval duration. This is surprising given theories that CI ability should deteriorate more for schizotypy-spectrum groups than for controls across longer delay periods (Barch & Braver, 2009). This may suggest that any performance deficits due to poor maintenance of context do not reliably manifest above and beyond those due to initial poor representation of context, at least across the delay periods used in these tasks (the maximum delay period in the included studies was 9500ms). Authors aiming to investigate the possible distinction between representation and maintenance of context may want to consider using a cue-probe interval of longer than one second in the long delay condition because delays shorter than one second did not affect outcomes across studies; it is still possible that context maintenance difficulties are present and would only be revealed using longer delay periods.

Because the use of novel dot patterns makes cue maintenance more challenging in the DPX compared to the AX-CPT (Barch et al., 2009), it is logical that the DPX may be more sensitive to subtle deficits in at-risk individuals by avoiding possible ceiling effects. Since use of DPX versus AX-CPT was not distributed evenly across studies with different diagnostic groups, task type could only be assessed as a moderator in schizophrenia groups versus healthy control groups. Task type did not moderate outcomes; however, at-risk groups were not included in these comparisons. Therefore these results do not address whether the DPX may find larger effects in at-risk groups

than the AX-CPT. Further study of the DPX with at-risk groups is needed to examine whether it may be a more appropriate task to measure CI in less impaired individuals.

Overall, CI deficits were generally robust across task type and various timing parameters, but total number of trials did predict the size of between-group differences with more trials predicting stronger d'context effects. Based on regression line predictions, it is therefore recommended that researchers use at least 250 trials to capture large effects.

Limitations and Recommendations for Future Research

The current study was limited according to gaps in the literature. The presented meta-analysis ignored the possibly confounding factor of medication status. There were not enough studies examining medicated, first-episode patients (k=3 studies) to provide adequate power for subgroup analysis. Most patients with chronic schizophrenia included in this literature were medicated and in non-acute phases of the illness. More research is needed that examines the impact of medication and distinguishes chronicity from medication status. Only a few studies have directly approached these issues (Barch et al., 2003; Lesh et al., 2015; Woodward, 2016, *personal communication*). Using a cross-sectional design, Lesh and colleagues (2015) found that unmedicated first-episode patients had lower d'context scores than medicated first-episode patients, at the level of a medium effect size (Hedges' g=-0.67). In preliminary cross-sectional results from Neil Woodward, Ph.D., at the time of personal communication, medicated first-episode patients (n=15) had lower d'context scores than chronic medicated patients (n=35) at the level of a small effect size (Hedges' g=-0.19). Finally, Barch and colleagues' (2003)

longitudinal design showed that first-episode patients had comparable CI performance when they were medication-naïve at baseline and following antipsychotic treatment 4 weeks later. Overall, it is still unclear how medication status may impact CI ability or whether it might interact with illness length to affect cognitive outcomes. Research directly comparing medicated and unmedicated chronic schizophrenia patients with medicated and medication-naïve first-episode patients could help clarify these relationships, including research differentiating among various types of unmedicated patients: never medicated, currently medication non-compliant, individuals who do not need or want medication, etc.

Cognitive impairment often predicts functional outcome in schizophrenia, thus, it would be beneficial to understand how CI impairments may influence functioning in social, community, occupational, as well as related domains such as quality of life. These relationships could not be assessed in the current meta-analysis because few studies looked at associations with functional outcome. Those that did reported some significant correlations of CI with global and community functioning and with performance-based skills in patients with schizophrenia (Gold et al., 2012; Richard et al., 2013; Sheffield et al., 2014; Stratta et al., 2000; Todd et al., 2014), with effect sizes ranging from negligible to large. Especially given the possible applications to intervention, assessing functional outcome in conjunction with CI is an area in need of further study.

As described earlier, treating schizotypy-spectrum psychopathology as unidimensional is problematic because it can mask true effects that are differentially expressed across symptom dimensions. Use of unidimensional models, inconsistent or

invalid symptom measurements, and unequal representation of symptom dimensions in a sample could limit the validity of findings and may account for some of the significant heterogeneity among studies. Therefore, theory-driven, multidimensional assessments of CI deficits in schizotypy are recommended.

Some of the most prolific researchers in this field frequently re-used participants across multiple published studies. This is problematic (and contrary to the American Psychological Association's piecemeal publication practices), especially when sample sizes are often very small; for example, there are multiple instances in which data has previously been reported for over half of the published sample. Further, the overlap in samples is not always clearly described in the main text—or sometimes at all. Ethical issues in publishing aside, this practice is problematic because it inflates the perceived reliability of findings from individual research groups in qualitative reviews and hinders synthesis of data in quantitative reviews. If researchers insist on re-using participant data across publications, it is recommended that they state this clearly in the main text of the article and provide subset analyses in the group of new participants so readers can distinguish what data is novel.

Finally, while interpreting results from the current meta-analysis, it is important to keep in mind that significant heterogeneity among studies was present in many of the analyses. The diagnostic and parametric variables used to predict CI outcomes did not fully account for variability in task performance, suggesting that there is still much to learn about what impacts CI ability. I put forth several hypotheses about variables that may have contributed to heterogeneity, such as unexamined medication effects and

poorly measured variation in symptom expression, but it cannot be ruled out that other confounding factors (e.g., educational background, intelligence quotient, socioeconomic status, or motivation) are partially accounting for results.

Summary and Conclusions

CI impairment appears to be a component of schizotypy-spectrum psychopathology that is present across the schizotypy spectrum and in premorbid, active, and residual phases of schizophrenia, as opposed to simply being a disease marker or a consequence or sequelae of the many catastrophic effects of schizophrenia. CI is associated with disorganized and negative symptoms of schizophrenia and appears stable throughout the course of illness. These results support theories that CI may be a stable vulnerability factor for schizophrenia. Further study is still warranted to confirm if CI is indeed an endophenotype, to understand the neural mechanisms underlying CI, to clarify CI's role in the expression of schizotypy spectrum symptoms and impairment, and to determine the extent to which CI impairment can be remediated by interventions.

To summarize, recommendations for future research include focus on longitudinal studies; inclusion of at-risk samples, particularly SPD, prodromal individuals or those at ultra high-risk to develop schizophrenia, and psychometrically identified schizotypy; research designed to differentiate effects of antipsychotic medication and duration of illness in or number of illness episodes; research specifically designed to measure differential associations of symptom dimensions with CI; and examination of the relationship between CI and functional outcome. Researchers wishing to assess goal maintenance with the AX-CPT or DPX may find it beneficial to use long delay periods of

ten seconds or longer. Administration of at least 250 trials is suggested in order to detect large effects.

The implication of CI in the development and expression of schizotypy-spectrum psychopathology has strong roots in phenomenological, neural, and cognitive neuroscience theories. Foundational research in this area has established a strong basis for CI impairment that is linked to cognitive and neurological outcomes. Important next steps include achieving a more nuanced understanding of the ways specific factors may interact to influence CI ability in schizotypy and integrating this research across other fields of study for a more holistic understanding of the mechanisms at play.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist, 63,* 839-851.

Andreasen, N. C., & Carpenter, W. R. (1993). *Diagnosis and Classification of Schizophrenia. Schizophrenia Bulletin, 19,* 199-214.

Andreasen, N. C, Flaum, M., Swayze, V W, Tyrrell, G., & Arndt, S. (1990). Positive and negative symptoms in schizophrenia. *Archives of General Psychiatry, 47,* 615-621.

Arndt, S., Alliger, R. J., & Andreasen, N. C. (1991). The distinction of positive and negative symptoms: The failure of a two-dimensional model. *British Journal of Psychiatry, 158,* 317-322.

Atkinson, R. C. (1963). A variable sensitivity theory of signal detection. *Psychological Review, 70,* 91-106.

Barch, D. M. (2005). The cognitive neuroscience of schizophrenia. *Annual Review of Clinical Psychology, 1,* 321-353.

Barch, D.M., Berman, M.G., Engle, R., Hurdelbrink Jones, J., Jonides, J., MacDonald, A., … Sponheim, S.R. (2009) CNTRICS final task selection: Working memory. *Schizophrenia Bulletin, 35*, 136-152.

Barch, D. M., & Braver, T. S. (2009). Context, goals, and behavioral regulation in schizophrenia: Psychological and neural mechanisms. In K. Salzinger & M. R. Serper (Eds.), *Behavioral mechanisms and psychopathology: Advancing the explanation of its nature, cause, and treatment*. (pp. 41–76). Washington, DC, US: American Psychological Association.

Barch, D. M., Carter, C. S., Braver, T. S., Sabb, F. W., MacDonald, A., Noll, D. C., & Cohen, J. D. (2001). Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Archives of General Psychiatry*, *58*, 280–288.

Barch, D. M., Carter, C. S., MacDonald, A. W., Braver, T. S., & Cohen, J. D. (2003). Context-processing deficits in schizophrenia: diagnostic specificity, 4-week course, and relationships to clinical symptoms. *Journal of Abnormal Psychology, 112*, 132–143.

Barch, D. M., Mitropoulou, V., Harvey, P. D., New, A. S., Silverman, J. M., & Siever, L.J. (2004). Context-processing deficits in schizotypal personality disorder. *Journal of Abnormal Psychology, 113, 4*, 556-568.

Barch, D. M., & Sheffield, J. M. (2017). Cognitive Control in Schizophrenia Psychological and Neural Mechanisms. In Egner, T. (Ed.) *The Wiley handbook of cognitive control* (pp. 556-580). Chichester, West Sussex, UK: John Wiley & Sons.

Becker, T. M. (2012). Understanding the relationship between goal maintenance and disorganized speech. *Unpublished Dissertation.*

Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive functioning in schizophrenia, schizoaffective disorder and affective psychoses: meta-analytic study. *The British Journal of Psychiatry: the Journal of Mental Science*, *195*, 475-482.

Borenstein, M. (2009). Introduction to meta-analysis. Chichester, U.K: John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., Rothstein, H. (2015). Comprehensive Meta-Analysis (Version 3) [Computer software]. Englewood, NJ: Biostat. Available from http://www.meta-analysis.com/index.php

Braver, T. S., Barch, D. M., & Cohen, J. D. (1999). Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. *Biological Psychiatry*, *46*, 312-328.

Ceccherini-Nelli, A., Turpin-Crowther, K., & Crow, T. J. (2007). Schneider's first rank symptoms and continuous performance disturbance as indices of dysconnectivity of left- and right-hemispheric components of language in schizophrenia. *Schizophrenia Research, 90,* 203-213.

Chan, R.C., Gottesman, I.I. (2008). Neurological soft signs as candidate endophenotypes for schizophrenia: a shooting star or a Northern star? *Neuroscience and Biobehavioral Reviews, 32,* 957–71

Chapman, L., & Chapman, J. (1973). Problems in the measurement of cognitive deficits. *Psychological Bulletin, 79,* 380–385.

Chapman, L. J., & Chapman, J. P. (1978). The measurement of differential deficit. *Journal of Psychiatric Research, 14,* 303-311.

Chung, Y. S., Mathews, J. R., & Barch, D. M. (2011). The effect of context processing on different aspects of social cognition in schizophrenia. *Schizophrenia Bulletin, 37,* 1048–1056.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: L. Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin: 112,* 155-159.

Cohen, R. (2011). Continuous performance tests. In J.S. Kreutzer, J. Deluca, & B. Kaplan (Eds). Encyclopedia of Clinical Neuropsychology (pp. 699-701). New York: Springer.

Cohen, J. D., Barch, D. M., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology, 108,* 120–133.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99,* 45-77.

Conrad, K. (1958). *Die beginnende Schizophrenie.* Stuttgart, Germany: Thieme Verlag.

Cools, R., Brouwer, W. H., De Jong, R., & Slooff, C. (2000). Flexibility, inhibition and planning: Frontal dysfunctioning in schizophrenia. *Brain & Cognition, 43,* 108–112.

Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Los Angeles: Sage.

Culbreth, A. J., Gold, J. M., Cools, R., & Barch, D. M. (2016). Impaired Activation in Cognitive Control Regions Predicts Reversal Learning in Schizophrenia. *Schizophrenia Bulletin, 42,* 484-493.

Delawalla, Z., Barch, D. M., Eastep, J. L. F., Thomason, E. S., Hanewinkel, M. J., Thompson, P. A., & Csernansky, J. G. (2006). Factors mediating cognitive deficits and psychopathology among siblings of individuals with schizophrenia. *Schizophrenia Bulletin, 32*, 525–537.

Delawalla, Z., Csernansky, J. G., & Barch, D. M. (2008). Prefrontal cortex function in nonpsychotic siblings of individuals with schizophrenia. *Biological Psychiatry, 63*, 490-497.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463.

Ellison-Wright, I., Glahn, D., Laird, A., Thelen, S., & Bullmore, E. (2008). The Anatomy of First-Episode and Chronic Schizophrenia: An Anatomical Likelihood Estimation Meta-Analysis. *The American Journal of Psychiatry, 165,* 1015-1023.

Fornito, A., Yoon, J., Zalesky, A., Bullmore, E. T., & Carter, C. S. (2011). General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance. *Biological Psychiatry, 70*, 64–72.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T., Griffith, L., ... Trikalinos, T. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology, 64,* 1187-1197.

Fusar-Poli, P., Perez, J., Broome, M., Borgwardt, S., Placentino, A., Caverzasi, E., … others. (2007). Neurofunctional correlates of vulnerability to psychosis: a systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews, 31,* 465–484.

Goghari, V. M. (2011). Executive functioning-related brain abnormalities associated with the genetic liability for schizophrenia: an activation likelihood estimation meta-analysis. *Psychological Medicine, 41,* 1239–1252.

Goghari, V. M., Sponheim, S. R., & MacDonald, A. W. (2010). The functional neuroanatomy of symptom dimensions in schizophrenia: A qualitative and quantitative review of a persistent question. *Neuroscience and Biobehavioral Reviews, 34,* 468-508.

Gold, J. M., Barch, D. M., Carter, C. S., Dakin, S., Luck, S. J., MacDonald, A. W. ., Ragland, J. D., ... Strauss, M. (2012). Clinical, functional, and intertask correlations of measures developed by the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia Consortium. *Schizophrenia Bulletin, 38,* 144-152.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *The American Journal of Psychiatry, 160,* 636-645.

Grynszpan, O., Perbal, S., Pelissolo, A., Fossati, P., Jouvent, R., Dubal, S., & Perez-Diaz, F. (2011). Efficacy and specificity of computer-assisted cognitive remediation in schizophrenia: a meta-analytical study. *Psychological Medicine, 41,* 163–173.

Henderson, D., Poppe, A. B., Barch, D. M., Carter, C. S., Gold, J. M., Ragland, J. D., … MacDonald, A. W. (2012). Optimization of a goal maintenance task for use in clinical applications. *Schizophrenia Bulletin*, *38*, 104–113.

Higgins, J. P. T. & Green, S. (Eds.) (2011). *Cochrane Handbook for Systematic Reviews of Interventions 5.1.0*. The Cochrane Collaboration 2011. Accessible at: http://www.cochrane-handbook.org

Javitt, D. C., Rabinowicz, E., Silipo, G., & Dias, E. C. (2007). Encoding vs. retention: Differential effects of cue manipulation on working memory performance in schizophrenia. *Schizophrenia Research, 91,* 159-168.

Jones, J. A., Sponheim, S. R., & MacDonald, A. W. (2010). The dot pattern expectancy task: reliability and replication of deficits in schizophrenia. *Psychological Assessment, 22,* 131-41.

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry, 16,* 13–23.

Keefe, R. S. E., Bilder, R. M., Harvey, P. D., Davis, S. M., Palmer, B. W., Gold, J. M., Meltzer, H. Y., ... Perkins, D. O. (2006). Baseline Neurocognitive Deficits in the CATIE Schizophrenia Trial. *Neuropsychopharmacology, 31,* 2033-2046.

Knable, M. B., & Weinberger, D. R. (1997). Dopamine, the prefrontal cortex and schizophrenia. *Journal of Psychopharmacology, 11,* 123-131.

König, P., Chiang, C., & von, S. A. (1997). Internal context and top-down processing. *Behavioral and Brain Sciences, 20,* 691-692.

Kwapil, T.R. & Barrantes-Vidal, N. (2015). Schizotypy: Looking back and moving forward. *Schizophrenia Bulletin (suppl 2):* S366-S373.

Kwapil, T.R, & Chun, C.A. (2015). The psychometric assessment of schizotypy. In O. Mason & G. Claridge, Eds., *Schizotypy: New Dimensions*. London: Routledge.

Kwapil, T. R., Gross, G. M., Silvia, P. J., & Barrantes-Vidal, N. (2013). Prediction of psychopathology and functional impairment by positive and negative schizotypy in the Chapmans' ten-year longitudinal study. *Journal of Abnormal Psychology, 122*, 807–815.

Lenzenweger, M. F. (2010). Schizotypy and schizophrenia: The view from experimental psychopathology. New York, NY: Guilford Press.

Lesh, T. A., Tanase, C., Geib, B. R., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., ... Carter, C. S. (2015). A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *Jama Psychiatry, 72,* 226-234.

Lesh, T. A., Westphal, A. J., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., … Carter, C. S. (2013). Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia. *NeuroImage: Clinical, 2*, 590–599.

Liddle, P. F. (1987). Schizophrenic syndromes, cognitive performance and neurological dysfunction. *Psychological Medicine, 17*, 49-57.

Lopez-Garcia, P., Espinoza, L. Y., Santos, P. M., Marin, J., & Sanchez-Pedreño, F. O. (2013). Impact of COMT genotype on cognition in schizophrenia spectrum patients and their relatives. *Psychiatry Research, 208*, 118-124.

MacDonald, A. W. & Carter, C. S. (2002). Cognitive experimental approaches to investigating impaired cognition in schizophrenia: A paradigm shift. *Journal of Clinical and Experimental Neuropsychology, 24*, 873-882.

MacDonald, A. W., & Carter, C. S. (2003). Event-related FMRI study of context processing in dorsolateral prefrontal cortex of patients with schizophrenia. *Journal of Abnormal Psychology, 112*, 689–697.

MacDonald, A. W., Carter, C. S., Flory, J. D., Ferrell, R. E., & Manuck, S. B. (2007). COMT val158Met and executive control: a test of the benefit of specific deficits to translational research. *Journal of Abnormal Psychology, 116,* 306–312.

MacDonald, A. W., Carter, C. S., Kerns, J. G., Ursu, S., Barch, D. M., Holmes, A. J., … Cohen, J. D. (2005). Specificity of prefrontal dysfunction and context processing deficits to schizophrenia in never-medicated patients with first-episode psychosis. *American Journal of Psychiatry, 162*, 475–484.

MacDonald, A. I., Goghari, V. M., Hicks, B. M., Flory, J. D., Carter, C. S., & Manuck, S. B. (2005). A Convergent-Divergent Approach to Context Processing, General Intellectual Functioning, and the Genetic Liability to Schizophrenia. *Neuropsychology, 19*, 814-821

MacDonald, A., Pogue-Geile, M. F., Johnson, M. K., & Carter, C. S. (2003). A specific deficit in context processing in the unaffected siblings of patients with schizophrenia. *Archives of General Psychiatry*, *60*, 1-15.

Macmillan, N. & Kaplan, H. (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*, 185-199.

Matussek, P. (1952). Untersuchungen über die Wahnwahrnehmung (Studies in Delusional Perception). *Archiv für Psychiatrie und Nervenkrankheiten*, *184*, 279-319.

Matussek, P. (1987). Studies in delusional perception. *Brain, Behavior, and Evolution, 59*, 21-32.

McClure, M. M., Barch, D. M., Romero, M. J., Minzenberg, M. J., Triebwasser, J., Harvey, P. D., & Siever, L. J. (2007). The Effects of Guanfacine on Context Processing Abnormalities in Schizotypal Personality Disorder. *Biological Psychiatry, 61*, 1157–1160.

McClure, M. M., Barch, D. M., Flory, J. D., Harvey, P. D., & Siever, L. J. (2008). Context processing in schizotypal personality disorder: evidence of specificity of impairment to the schizophrenia spectrum. *Journal of Abnormal Psychology, 117,* 342-54.

McClure, M. M., Harvey, P. D., Goodman, M., Triebwasser, J., New, A., Koenigsberg, H. W., … Siever, L. J. (2010). Pergolide treatment of cognitive deficits associated with schizotypal personality disorder: continued evidence of the importance of the dopamine system in the schizophrenia spectrum. *Neuropsychopharmacology, 35,* 1356–1362.

Meehl, P. E. (1989). Schizotaxia revisited. *Archives of General Psychiatry, 46,* 935-944.

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, *66*, 811–822.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Journal of Clinical Epidemiology, 62,* 1006-1012.

Myin-Germeys, I., Krabbendam, L., Jolles, J., Delespaul, P. A., & van Os, J. (2002). Are cognitive impairments associated with sensitivity to stress in schizophrenia? An experience sampling study. *The American Journal of Psychiatry, 159*, 443–449.

Niendam, T. A., Lesh, T. A., Yoon, J., Westphal, A. J., Hutchison, N., Ragland, D. J., … Carter, C. S. (2014). Impaired context processing as a potential marker of psychosis risk state. *Psychiatry Research - Neuroimaging*, *221*, 13–20.

Nieuwenstein, M. R., Aleman, A., & de Haan, E. H. F. (2001). Relationship between symptom dimensions and neurocognitive functioning in schizophrenia: a meta-analysis of WCST and CPT studies. *Journal of Psychiatric Research, 35,* 119-125.

Nuechterlein, K. H. (1991). Vigilance in schizophrenia and related disorders. In S. R. Steinhauer, J. H. Gruzelier, & J. Zubin (Eds.), *Handbook of schizophrenia. Vol. 5: Neuropsychology, psychophysiology, and information processing* (pp. 397–433). Amsterdam: Elsevier.

Nuechterlein, K. H., Dawson, M. E., Gitlan, M., Ventura, J., Goldstein, M. J., Snyder, K. S., … Mintz, J. (1992). Developmental processes in schizophrenic disorders: Longitudinal studies of vulnerability and stress. *Schizophrenia Bulletin, 18,* 387–425.

Nuechterlein, K. H., Green, M. F., Calkins, M. E., Greenwood, T. A., Gur, R. E., Gur, R. C., & ... Braff, D. L. (2015). Attention/vigilance in schizophrenia: Performance results from a large multi-site study of the Consortium on the Genetics of Schizophrenia (COGS). *Schizophrenia Research, 163,* 38-46.

O'Keefe, D. (2003). Colloquy: Should familywise alpha be adjusted? Against familywise alpha adjustment. *Human Communication Research, 29,* 431–447.

Owoso, A, Carter, C. S., Gold, J. M., Macdonald, A W., Ragland, J. D., Silverstein, S. M., … Barch, D. M. (2013). Cognition in schizophrenia and schizo-affective disorder: impairments that are more similar than different. *Psychological Medicine, 43,* 2535-2545.

Perlstein, W. M., Dixit, N. K., Carter, C. S., Noll, D. C., & Cohen, J. D. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry, 53,* 25-38.

Phillips, W. A., & Silverstein, S. M. (2013). The coherent organization of mental life depends on mechanisms for context-sensitive gain-control that are impaired in schizophrenia. *Frontiers in Psychology, 4, 307,* 1-15.

Phillips, W. A., & Singer, W. (1997). In search of common foundations for cortical computation. *The Behavioral and Brain Sciences, 20*, 657-83.

Poppe, A. B., Carter, C. S., Minzenberg, M. J., & MacDonald, A. W. (2015). Task-based functional connectivity as an indicator of genetic liability to schizophrenia. *Schizophrenia Research, 162,* 118-123.

Preacher, K. J., Rucker, D. D., MacCallum, R. C., & Nicewander, W. A. (2005). Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychological Methods*, *10,* 178–192.

Ray, K. L., Lesh, T. A., Howell, A. M., Salo, T. P., Ragland, J. D., MacDonald, A. W., Gold, J. M., ... Carter, C. S. (2017). Functional network changes and cognitive control in schizophrenia. *Neuroimage: Clinical, 15,* 161-170.

Reilly, J. L., Hill, S. K., Gold, J. M., Keefe, R. S., Clementz, B. A., Gershon, E., Keshavan, M. S., ... Sweeney, J. A. (2017). Impaired Context Processing is Attributable to Global Neuropsychological Impairment in Schizophrenia and Psychotic Bipolar Disorder. *Schizophrenia Bulletin, 43,* 397-406.

Richard, A. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2013). Persistence, diagnostic specificity and genetic liability for context-processing deficits in schizophrenia. *Schizophrenia Research, 147,* 75-80.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results, *Psychological Bulletin, 86,* 638-641.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology 20,* 343-350.

Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8,* 467-474.

Servan-Schreiber, D., Cohen, J. D., & Steingard, S. (1996). Schizophrenic deficits in the processing of context. A test of a theoretical model. *Archives of General Psychiatry, 53,* 1105-1112.

Sheffield, J. M., Gold, J. M., Strauss, M. E., Carter, C. S., MacDonald, A. W., Ragland, J. D., … Barch, D. M. (2014). Common and specific cognitive deficits in schizophrenia: Relationships to function. *Cognitive, Affective & Behavioral Neuroscience, 14,* 161–174.

Silverstein, S. M., & Schenkel, L. S. (1997). Schizophrenia as a model of context-deficient cortical computation. *Behavioral and Brain Sciences, 20,* 696-697.

Simonsohn, U. (2017). P-curve app (Version 4.052). [Online computer software]. Retrieved from http://www.p-curve.com/app4/.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143,* 534-547.

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2014). P-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. *Perspectives on Psychological Science, 9,* 666-681.

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller (2015). *Journal of Experimental Psychology, 144,* 1146.

Sloat, V. C. (2007). Context processing in psychometrically defined schizotypes (Doctoral dissertation). Retrieved from ProQuest Information and Learning Company. (UMI Number: 3285814).

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis. *Journal of Clinical Epidemiology, 54,* 1046-1055.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., ... Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343.*

Stratta, P., Daneluzzo, E., Bustini, M., Prosperini, P., & Rossi, A. (2000). Processing of context information in schizophrenia: relation to clinical symptoms and WCST performance. *Schizophrenia Research, 44,* 57–67.

Strauss, M. E., McLouth, C. J., Barch, D. M., Carter, C. S., Gold, J. M., Luck, S. J., … Silverstein, S. M. (2014). Temporal stability and moderating effects of age and sex on CNTRACS task performance. *Schizophrenia Bulletin, 40,* 835–844.

Sun, X., Briel, M., Walter, S. D., & Guyatt, G. H. (2010). Updating criteria to evaluate the credibility of subgroup analyses. *British Medical Journal, 7751,* 850-854.

Swets, J. A. (1961). Detection theory and psycho- physics: A review. *Psychometrika, 26,* 49-63.

Tandon, R., Keshavan, M., Nasrallah, H. (2008). Schizophrenia, "Just the Facts" What we know in 2008. 2. Epidemiology and etiology. *Schizophrenia Research, 102,* 1-18.

Thoma, P., & Daum, I. (2008). Working memory and multi-tasking in paranoid schizophrenia with and without comorbid substance use disorder. *Addiction, 103,* 774–786.

Todd, J., Whitson, L., Smith, E., Michie, P. T., Schall, U., & Ward, P. B. (2014). What's intact and what's not within the mismatch negativity system in schizophrenia. *Psychophysiology, 51,* 337–347.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegungen. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 61,* 161-265.

Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt: I. Prinzipielle Bermerkungen. *Psychologische Forschung, 1,* 47-58.

Wohlberg, G.W., Kornetsky, C. (1973). Sustained attention in remitted schizophrenics. *Archives of General Psychiatry, 28,* 533-537.

Yoon, J. H., Minzenberg, M. J., Ursu, S., Walters, R., Wendelken, C., Ragland, J. D., & Carter, C. S. (2008). Association of Dorsolateral Prefrontal Cortex Dysfunction With Disrupted Coordinated Brain Activity in Schizophrenia: Relationship With Impaired Cognition, Behavioral Disorganization, and Global Function. *American Journal of Psychiatry, 165*, 1006–1014.

Yoon, J. H., Nguyen, D. V., McVay, L. M., Deramo, P., Minzenberg, M. J., Ragland, J. D., Niendam, T., ... Carter, C. S. (2012). Automated classification of fMRI during cognitive control identifies more severely disorganized subjects with schizophrenia. *Schizophrenia Research, 135*, 28-33.

APPENDIX A

TABLES AND FIGURES

Table 1

Literature Search and Screening Criteria

Literature Search Terms	Title/Abstract Screening Criteria	Full-Text Screening Criteria
• AX-CPT and schizo*; dot pattern expectancy task and schizo*; DPX and schizo*; CNTRACS and schizo*; context* integration and schizo*; context* processing and schizo*; context* and ultra high risk • AX-CPT and prodromal; dot pattern expectancy task and prodromal; DPX and prodromal; CNTRACS	• If the title was completely irrelevant, I did not read abstract • If the title mentioned any neural, cognitive, or behavioral topic and schizotypy or another psychiatric disorder in which researchers use patients with schizophrenia as a psychiatric comparison group (e.g., mood disorder, other	• For each article, I electronically searched for "CPT," "DPX," "CNTRACS," and "supplem*." I used additional search functions for other article types, as described below. • I first skimmed the article to determine whether it was empirical. If the article did not appear to be an empirical article or did not have the structure of a

and prodromal; context* integration and prodromal; context* processing and prodromal

- AX-CPT and high risk; dot pattern expectancy task and high risk; DPX and high risk; CNTRACS and high risk; context* integration and high risk; context* processing and high risk
- AX-CPT and psychosis; dot pattern expectancy task and psychosis; DPX and psychosis; CNTRACS and psychosis; context* integration and psychosis; context* processing and psychosis

psychotic disorder, personality disorders), I read the abstract.

- If the abstract mentioned using the CNTRACS tasks or broadly mentioned neuropsychological or cognitive testing without listing which measures, I flagged it for full-text screening.
- If the title and/or abstract was written in English or French, I screened it as-is. If the title and or/abstract was written in another language, I used Google translation to screen for appropriateness.

traditional article, I skimmed the entire full-text to make sure the article did not contain new empirical data (e.g., a review article with a brief empirical section). If the article was not empirical, I also used the search function to search for "method*."

- If the article was empirical, I read the methods section to determine whether they used CNTRACS tasks (AX-CPT or DPX with 10+ total trials and 60+% AX trials). I also skimmed the structure of the article to look for multiple experiments within the article.
- If the article did use the AX-CPT and/or

DPX, I then read the methods section to determine whether they used a schizotypy measure or diagnosis. If the article used one of the CNTRACS tasks but did not include a measure of schizotypy, I also searched for "schizo*."

- I read supplemental materials when available, to look for relevant empirical data

Table 2

Summary of Demographic Information Across all Included Studies.

	First-Episode, Unmedicated	First-Episode, Medicated	Chronic Schizophrenia	At-Risk Group	Healthy Controls	Psychiatric Controls
Age: M (SD)	**23.51 (1.60)** [k=6, n=176]	**20.37 (0.59)** [k=4, n=112]	**34.83 (5.00)** [k=30, n=2096]	**33.63 (9.77)** [k=14, n=442]	**30.42 (6.65)** [k=41, n=2674]	**33.37 (6.47)** [k=12, n=447]
Education (years): M (SD)	**12.68 (0.43)** [k=6, n=176]	**12.47 (0.35)** [k=4, n=112]	**11.83 (1.69)** [k=24, n=1553]	**13.61 (1.17)** [k=10, n=347]	**12.93 (2.66)** [k=34, n=2201]	**13.76 (1.30)** [k=7, n=156]
Length of illness (years): M (SD)	**0.58** [k=1, n=22]	**0.59 (0.21)** [k=2, n=38]	**8.06 (4.36)** [k=12, n=864]	--	--	--
% Female	**29.16%** [k=6, n=176]	**26.84%** [k=4, n=112]	**35.81%** [k=30, n=2096]	**49.94%** [k=14, n=442]	**48.96%** [k=40, n=2654]	**53.97%** [k=12, n=447]
% White	--	**0%** [k=1, n=15]	**52.00%** [k=13, n=676]	**54.06%** [k=7, n=227]	**60.67%** [k=15, n=1181]	**69.86%** [k=3, n=259]

% Black	--	**80.00%** [k=1, n=15]	**43.67%** [k=8, n=798]	**43.42%** [k=3, n=108]	**27.98%** [k=7, n=527]	**21.15%** [k=4, n=284]
% Latino	--	**0%** [k=1, n=15]	**28.00%** [k=5, n=243]	**34.92%** [k=4, n=126]	**53.77%** [k=4, n=160]	**2.70%** [k=2, n=37]
% Asian	--	--	**50.75%** [k=4, n=1180]	**8.33%** [k=3, n=108]	**69.70%** [k=3, n=1066]	**3.3%** [k=2, n=240]
% Other Race/Ethnicity	--	**20.00%** [k=1, n=15]	**7.22%** [k=3, n=611]	**2.60%** [k=2, n=84]	**7.10%** [k=3, n=385]	**4.23%** [k=3, n=259]
% Medicated	**0%** [k=6, n=176]	**78.22%** [k=3, n=61]	**98.05%** [k=18, n=1126]	--	--	--

k=number of studies providing information. n=number of participants providing information. Note that for categorical variables, k and n do not correspond to the total percentage in that category for the entire group, but rather to the percentage

only within the subset of studies providing data for that category. See Results section in text for a description of group composition.

Table 3

Primary Analyses: Random Effects Model of Schizotypy-Spectrum Groups versus Healthy Control Groups for Short Delay.

	Hedges' *g*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
Short AX_{Err}	**0.59**	0.41	0.77	$1.22E^{-10}$	51.59	20	$1.31E^{-4}$	61.23	0.095
Short AY_{Err}	0.33	0.19	0.46	$1.94E^{-6}$	33.59	21	0.040	37.48	0.036
Short BX_{Err}	**0.61**	0.50	0.71	0	22.22	21	0.39	5.49	0.0035
Short BY_{Err}	0.32	0.14	0.50	$6.48E^{-4}$	51.73	19	$7.26E^{-5}$	63.27	0.099
Short AX_{RT}	**0.50**	0.30	0.69	$4.24E^{-7}$	36.42	16	0.0025	56.07	0.083
Short AY_{RT}	0.48	0.29	0.66	$2.53E^{-7}$	41.18	17	$8.79E^{-4}$	58.72	0.080
Short BX_{RT}	**0.62**	0.44	0.81	$6.71E^{-11}$	43.12	17	$4.61E^{-4}$	60.57	0.088
Short BY_{RT}	**0.68**	0.50	0.85	$5.15E^{-14}$	30.25	16	0.017	47.11	0.060
Short d'context	***-0.84***	-1.06	-0.63	$1.64E^{-14}$	96.00	23	$6.84E^{-11}$	76.04	0.20

Err=errors. RT=reaction time. Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 4

Primary Analyses: Random Effects Model of Schizotypy-Spectrum Groups versus Healthy Control Groups for Long Delay.

	Hedges' *g*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
Long AX_{Err}	**0.55**	0.42	0.68	0	38.35	24	0.032	37.43	0.033
Long AY_{Err}	0.00	-0.17	0.18	0.97	82.07	25	$5.38E^{-8}$	69.54	0.12
Long BX_{Err}	**0.59**	0.47	0.72	0	44.53	27	0.018	39.36	0.037
Long BY_{Err}	0.19	-0.01	0.39	0.06	100.39	24	$2.57E^{-11}$	76.09	0.17
Long AX_{RT}	**0.66**	0.46	0.85	$2.29E^{-11}$	62.61	21	$5.11E^{-6}$	66.46	0.12
Long AY_{RT}	0.47	0.31	0.64	$1.77E^{-8}$	47.08	21	$9.15E^{-4}$	55.40	0.072
Long BX_{RT}	0.49	0.33	0.65	$3.67E^{-9}$	45.00	21	0.0017	53.34	0.066
Long BY_{RT}	**0.58**	0.41	0.74	$5.67E^{-12}$	41.21	20	0.0035	51.46	0.062
Long d'context	**-0.76**	-0.91	-0.61	0	67.11	27	$2.87E^{-5}$	59.77	0.08

Err=errors. RT=reaction time. Hedges' *g*: medium effects in bold.

Table 5

Primary Analyses: Random Effects Model of Schizophrenia Groups versus Psychiatric Comparison Groups

Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
AY_{Err} Total	0.02	-0.12	0.15	0.81	1.95	5	0.86	0.00	0.00
Richard et al., 2013	-0.16	-0.54	0.21	0.40					
Woodward, 2016	-0.15	-0.85	0.54	0.66					
Barch et al., 2003	-0.10	-0.55	0.35	0.67					
Holmes et al., 2005	-0.05	-0.97	0.87	0.91					
Woodward, 2016	-0.01	-0.81	0.79	0.98					
Reilly et al., 2017	0.08	-0.09	0.25	0.34					
BX_{Err} Total	0.33	0.19	0.46	$1.65E^{-6}$	3.30	7	0.86	0.00	0.00
Woodward, 2016	0.00	-0.69	0.69	1.00					
Richard et al., 2013	0.26	-0.12	0.63	0.18					
Woodward, 2016	0.29	-0.51	1.10	0.47[1]					
Ceccherini-Nelli et al., 2007	0.33	-0.51	1.17	0.44					

Reilly et al., 2017	0.33	0.16	0.50	$1.04E^{-4}$					
Barch et al., 2003	0.34	-0.12	0.79	0.14					
Holmes et al., 2005	**0.56**	-0.37	1.50	0.24					
Ceccherini-Nelli et al., 2007	***0.94***	0.10	1.78	0.028					
$d'_{context}$ Total	**-0.73**	-1.07	-0.40	$1.73E^{-5}$	36.38	9	3.39E-5	75.26	0.19
Cohen et al., 1999	***-1.87***	-2.42	-1.31	$4.17E^{-11}$					
Holmes et al., 2005	***-1.55***	-2.61	-0.50	0.0039					
Ceccherini-Nelli et al., 2007	***-1.29***	-2.17	-0.41	0.0039					
Thoma & Daum, 2008	***-0.90***	-1.50	-0.30	0.0035					
Ceccherini-Nelli et al., 2007	**-0.69**	-1.55	0.17	0.11					
Woodward, 2016	-0.41	-1.10	0.29	0.25					
Reilly et al., 2017	-0.37	-0.54	-0.20	$1.55E^{-5}$					
Barch et al., 2003	-0.36	-0.81	0.09	0.12					
Richard et al., 2013	-0.35	-0.73	0.03	0.072					
Woodward, 2016	-0.13	-0.93	0.67	0.75					

Err=errors. Total meta-analytic effects for each outcome in gray. Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 6

Subgroup Analyses: Random Effects Model of Diagnostic Groups for AY Errors.

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Schz Total		0.29	0.16	0.42	$1.79E^{-5}$
Schz vs. HC	Todd et al., 2014	-0.46	-0.89	-0.033	0.035
Schz vs. HC	Barch et al., 2001	-0.34	-1.09	0.41	0.37
Schz vs. HC	Chung et al., 2011	-0.19	-0.67	0.30	0.45
Schz vs. HC	Yoon et al., 2014	-0.07	-0.81	0.66	0.85
Schz vs. HC	Stratta et al., 2000	-0.05	-0.66	0.56	0.87
Schz vs. HC	MacDonald, 2002	0.00	-0.89	0.89	1.00
Schz vs. HC	MacDonald, 2002	0.00	-0.93	0.93	1.00
Schz vs. HC	Richard et al., 2013	0.03	-0.35	0.42	0.87
Schz vs. HC	Barch et al., 2008	0.03	-0.38	0.44	0.87
Schz vs. HC	Barch et al., 2003	0.11	-0.25	0.47	0.54
Schz vs. HC	Lopez-Garcia et al., 2016	0.12	-0.65	0.90	0.75

Schz vs. HC	Becker, 2012	0.13	-0.33	0.59	0.58
Schz vs. HC	Lesh et al., 2015	0.17	-0.62	0.95	0.68
Schz vs. HC	Jones et al., 2010	0.17	-0.23	0.57	0.41
Schz vs. HC	Yoon et al., 2012	0.17	-0.21	0.56	0.38
Schz vs. HC	Lesh et al., 2015	0.21	-0.39	0.81	0.49
Schz vs. HC	Sheffield et al., 2015	0.30	-0.10	0.69	0.14
Schz vs. HC	Woodward, 2016	0.37	-0.30	1.04	0.28
Schz vs. HC	Reilly et al., 2017	0.43	0.28	0.58	$1.66E^{-8}$
Schz vs. HC	Woodward, 2016	0.49	-0.06	1.04	0.081
Schz vs. HC	Lopez-Garcia et al., 2015	**0.53**	0.06	1.00	0.027
Schz vs. HC	Sheffield et al., 2014	**0.53**	0.17	0.90	0.0045
Schz vs. HC	Braver et al., 1999	**0.55**	-0.14	1.24	0.119
Schz vs. HC	Edwards et al., 2010	**0.61**	-0.06	1.27	0.077
Schz vs. HC	Zhang et al., 2015	**0.70**	0.57	0.84	0
Schz vs. HC	Gold et al., 2012; Henderson et al., 2012	**0.71**	0.47	0.95	$1.13E^{-8}$

Schz vs. HC	Sheffield et al., 2014	**0.71**	0.34	1.08	$1.81E^{-4}$
Schz vs. HC	MacDonald & Carter, 2003	**0.79**	0.11	1.47	0.024
Schz vs. HC	Holmes et al., 2005	***0.81***	-0.17	1.78	0.10
At-risk Total		-0.04	-0.24	0.15	0.66
At-risk vs. HC	MacDonald, 2002	-0.39	-1.25	0.48	0.38
At-risk vs. HC	Barch et al., 2004	-0.30	-0.80	0.20	0.24
At-risk vs. HC	MacDonald, 2002	-0.22	-1.18	0.73	0.65
At-risk vs. HC	Lopez-Garcia et al., 2015	-0.14	-0.66	0.37	0.59
At-risk vs. HC	McClure et al., 2008	-0.08	-0.47	0.30	0.67
At-risk vs. HC	Sloat, 2007	0.08	-0.42	0.58	0.75
At-risk vs. HC	Lopez-Garcia et al., 2016	0.20	-0.56	0.97	0.60
At-risk vs. HC	Richard et al., 2013	0.25	-0.21	0.72	0.28

Schz=schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 6, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within schizophrenia groups?	Schizophrenia	75.72	28	$2.84E^{-6}$	63.02	0.065
-Is there significant variance in effects within at-risk groups?	At-Risk	4.14	7	0.76	0.00	0.00
-Does the grouping variable explain significant variance in the model?	Total within	79.86	35	$2.30E^{-5}$		
	Mixed effects					
-Does the effect differ between subgroups?	Total between	7.82	1	0.0052		

Table 7

Subgroup Analyses: Random Effects Model of Diagnostic Groups for BX Errors.

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Schz Total		**0.71**	0.61	0.80	0
Schz vs. HC	Lopez-Garcia et al., 2016	-0.04	-0.81	0.73	0.92
Schz vs. HC	Holmes et al., 2005	0.29	-0.65	1.23	0.54
Schz vs. HC	Sheffield et al., 2014	0.46	0.09	0.83	0.014
Schz vs. HC	Woodward, 2016	0.47	-0.20	1.14	0.17
Schz vs. HC	MacDonald & Carter, 2003	**0.50**	-0.17	1.17	0.14
Schz vs. HC	Reilly et al., 2017	**0.50**	0.35	0.66	$4.92E^{-11}$
Schz vs. HC	Barch et al., 2008	**0.52**	0.10	0.94	0.015
Schz vs. HC	Jones et al., 2010	**0.54**	0.14	0.95	0.0088
Schz vs. HC	Ceccherini-Nelli et al., 2007	**0.56**	-0.12	1.24	0.11
Schz vs. HC	Woodward, 2016	**0.56**	0.0066	1.11	0.047
Schz vs. HC	Chung et al., 2011	**0.58**	0.092	1.07	0.020

Schz vs. HC	Lesh et al., 2015	**0.58**	-0.024	1.19	0.060
Schz vs. HC	Richard et al., 2013	**0.59**	0.19	0.99	0.0035
Schz vs. HC	Edwards et al., 2010	**0.63**	-0.041	1.30	0.066
Schz vs. HC	Barch et al., 2001	**0.64**	-0.13	1.40	0.10
Schz vs. HC	Yoon et al., 2010	**0.65**	0.26	1.05	0.0012
Schz vs. HC	Sheffield et al., 2014	**0.68**	0.31	1.06	$3.02E^{-4}$
Schz vs. HC	Todd et al., 2014	**0.70**	0.27	1.14	0.0015
Schz vs. HC	Sheffield et al., 2015	**0.71**	0.31	1.11	$5.40E^{-4}$
Schz vs. HC	Gold et al., 2012; Henderson et al., 2012	**0.74**	0.50	0.99	$2.21E^{-9}$
Schz vs. HC	MacDonald, 2002	**0.75**	-0.17	1.68	0.11
Schz vs. HC	Lesh et al., 2015	**0.76**	-0.055	1.57	0.07
Schz vs. HC	Zhang et al., 2015	**0.79**	0.66	0.93	0
Schz vs. HC	Perlstein et al., 2003	***0.85***	0.13	1.57	0.020
Schz vs. HC	MacDonald, 2002	***0.88***	-0.10	1.86	0.079
Schz vs. HC	Barch et al., 2003	***0.90***	0.52	1.28	$2.88E^{-6}$

Schz vs. HC	Braver et al., 1999	***0.93***	0.22	1.65	0.01
Schz vs. HC	Fornito et al., 2011	***1.04***	0.45	1.64	$6.07E^{-4}$
Schz vs. HC	Yoon et al., 2014	***1.10***	0.31	1.89	0.0066
Schz vs. HC	Stratta et al., 2000	***1.14***	0.49	1.80	$6.53E^{-4}$
Schz vs. HC	Becker, 2012	***1.21***	0.71	1.70	$2.04E^{-6}$
Schz vs. HC	Lopez-Garcia, 2015	***1.84***	1.29	2.39	$5.60E^{-11}$
At-risk Total		0.36	0.15	0.58	0.0011
At-risk vs. HC	Richard et al., 2013	-0.08	-0.54	0.38	0.74
At-risk vs. HC	Sloat, 2007	-0.06	-0.56	0.44	0.82
At-risk vs. HC	Lopez-Garcia, 2016	0.20	-0.56	0.97	0.60
At-risk vs. HC	Paavola, 2013	0.35	-0.04	0.73	0.078
At-risk vs. HC	MacDonald, 2002	0.47	-0.40	1.33	0.29
At-risk vs. HC	Barch et al., 2004	**0.53**	0.02	1.04	0.043
At-risk vs. HC	MacDonald, 2002	**0.53**	-0.44	1.50	0.29
At-risk vs. HC	McClure et al., 2008	**0.58**	0.19	0.98	0.0038

At-risk vs. HC	Lopez-Garcia, 2015	***0.87***	0.34	1.41	0.0015

Schz=schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics, very large effects in underlined bold italics.

Table 7, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within schizophrenia groups?	Schizophrenia	43.80	31	0.063	29.22	0.02
-Is there significant variance in effects within at-risk groups?	At-Risk	11.64	8	0.17	31.29	0.033
-Does the grouping variable explain significant variance in the model?	Total within	55.44	39	0.042		
	Mixed effects					
-Does the effect differ between subgroups?	Total between	8.14	1	0.0043	8.14	1

Table 8

Subgroup Analyses: Random Effects Model of Diagnostic Groups for d'context

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Schz Total		***-0.94***	-1.08	-0.80	0
Schz vs. HC	Sheffield et al., 2014	***-1.83***	-2.26	-1.40	0
Schz vs. HC	Cohen et al., 1999	***-1.47***	-1.96	-0.97	$5.25E^{-9}$
Schz vs. HC	Todd et al., 2014	***-1.39***	-1.86	-0.92	$6.00E^{-9}$
Schz vs. HC	MacDonald, 2002	***-1.37***	-2.41	-0.32	0.010
Schz vs. HC	Holmes et al., 2005	***-1.36***	-2.40	-0.31	0.011
Schz vs. HC	Lesh et al., 2015	***-1.29***	-2.15	-0.44	0.0029
Schz vs. HC	Ceccherini-Nelli et al., 2007	***-1.27***	-2.01	-0.54	$6.67E^{-4}$
Schz vs. HC	MacDonald, 2002	***-1.25***	-2.23	-0.28	0.012
Schz vs. HC	Braver et al., 1999	***-1.20***	-1.94	-0.47	0.0014
Schz vs. HC	Merrill et al., 2017	***-1.20***	-1.77	-0.62	$4.33E^{-5}$
Schz vs. HC	Barch et al., 2003	***-1.18***	-1.57	-0.79	$2.96E^{-9}$

Schz vs. HC	Dias et al., 2013	***-1.08***	-1.83	-0.33	0.0050
Schz vs. HC	Gold et al., 2012; Henderson et al, 2012	***-1.03***	-1.28	-0.78	$1.11E^{-15}$
Schz vs. HC	Woodward, 2016	***-1.03***	-1.73	-0.32	0.0043
Schz vs. HC	Chung et al., 2011	***-1.02***	-1.53	-0.51	$8.98E^{-5}$
Schz vs. HC	Poppe et al., 2015	***-0.98***	-1.60	-0.36	0.0020
Schz vs. HC	Sheffield et al., 2014	***-0.83***	-1.20	-0.45	$1.60E^{-5}$
Schz vs. HC	Barch et al., 2001	***-0.80***	-1.58	-0.026	0.043
Schz vs. HC	Barch et al., 2008	**-0.71**	-1.13	-0.29	$9.81E^{-4}$
Schz vs. HC	Richard et al., 2013	**-0.69**	-1.09	-0.29	$6.71E^{-4}$
Schz vs. HC	Poppe et al., 2016	**-0.66**	-1.06	-0.27	$9.95E^{-4}$
Schz vs. HC	Reilly et al., 2017	**-0.66**	-0.82	-0.51	0
Schz vs. HC	Delawalla et al., 2006	**-0.65**	-1.15	-0.16	0.0098
Schz vs. HC	Woodward, 2016	**-0.59**	-1.14	-0.039	0.036
Schz vs. HC	Lesh et al., 2015	**-0.58**	-1.19	0.027	0.061
Schz vs. HC	Jones et al., 2010	**-0.51**	-0.91	-0.10	0.014

Schz vs. HC	Thoma & Daum, 2008	-0.45	-1.05	0.14	0.13
Schz vs. HC	MacDonald & Carter, 2003	-0.20	-0.86	0.46	0.55
At-risk Total		-0.40	-0.56	-0.23	$3.70E^{-6}$
At-risk vs. HC	McClure et al., 2008	**-0.72**	-1.12	-0.32	$4.29E^{-4}$
At-risk vs. HC	MacDonald, 2002	**-0.70**	-1.69	0.29	0.16
At-risk vs. HC	Poppe et al., 2015	-0.49	-1.01	0.034	0.067
At-risk vs. HC	Barch et al., 2004	-0.40	-0.91	0.10	0.12
At-risk vs. HC	Richard et al., 2013	-0.37	-0.84	0.094	0.12
At-risk vs. HC	MacDonald, 2002	-0.34	-1.21	0.52	0.44
At-risk vs. HC	Delawalla et al., 2008	-0.29	-0.71	0.12	0.16
At-risk vs. HC	Sloat, 2007	-0.28	-0.79	0.22	0.27
At-risk vs. HC	Delawalla et al., 2006	-0.09	-0.55	0.37	0.70

Schz=schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics, very large effects in underlined bold italics.

Table 8, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within schizophrenia groups?	Schizophrenia	61.62	27	$1.61E^{-4}$	56.18	0.068
-Is there significant variance in effects within at-risk groups?	At-Risk	5.12	8	0.74	0	0
-Does the grouping variable explain significant variance in the model?	Total within	66.74	35	0.00097		
	Mixed effects					
-Does the effect differ between subgroups?	Total between	23.45	1	$1.30E^{-6}$		

Table 9

Subgroup Analyses: Random Effects Model of Diagnostic Groups by Illness Length for AY Errors.

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Chronic Total		0.34	0.19	0.50	0.00
Chronic vs. HC	Todd et al., 2014	-0.46	-0.89	-0.033	0.035
Chronic vs. HC	Chung et al., 2011	-0.19	-0.67	0.30	0.45
Chronic vs. HC	Stratta et al., 2000	-0.05	-0.66	0.56	0.87
Chronic vs. HC	MacDonald, 2002	0.00	-0.89	0.89	1.00
Chronic vs. HC	MacDonald, 2002	0.00	-0.93	0.93	1.00
Chronic vs. HC	Barch et al., 2008	0.03	-0.38	0.44	0.87
Chronic vs. HC	Lopez-Garcia et al., 2016	0.12	-0.65	0.90	0.75
Chronic vs. HC	Becker, 2012	0.13	-0.33	0.59	0.58
Chronic vs. HC	Jones et al., 2010	0.17	-0.23	0.57	0.41
Chronic vs. HC	Sheffield et al., 2015	0.30	-0.10	0.69	0.14
Chronic vs. HC	Reilly et al., 2017	0.43	0.28	0.58	$1.66E^{-8}$

Chronic vs. HC	Woodward, 2016	0.49	-0.061	1.04	0.081
Chronic vs. HC	Lopez-Garcia et al., 2015	**0.53**	0.061	1.00	0.027
Chronic vs. HC	Sheffield et al., 2014	**0.53**	0.17	0.90	0.0045
Chronic vs. HC	Edwards et al., 2010	**0.61**	-0.065	1.27	0.077
Chronic vs. HC	Zhang et al., 2015	**0.70**	0.57	0.84	0.00
Chronic vs. HC	Gold et al., 2012; Henderson et al., 2012	**0.71**	0.47	0.95	$1.13E^{-8}$
Chronic vs. HC	Sheffield et al., 2014	**0.71**	0.34	1.08	$1.81E^{-4}$
Chronic vs. HC	MacDonald & Carter, 2003	**0.79**	0.11	1.47	0.024
Chronic vs. HC	Holmes et al., 2005	***0.81***	-0.17	1.78	0.10
FE Total		0.13	-0.04	0.30	0.14
FE vs. HC	Barch et al., 2001	-0.34	-1.09	0.41	0.37
FE vs. HC	Yoon et al., 2014	-0.07	-0.81	0.66	0.85
FE vs. HC	Richard et al., 2013	0.03	-0.35	0.42	0.87
FE vs. HC	Barch et al., 2003	0.11	-0.25	0.47	0.54
FE vs. HC	Lesh et al., 2015	0.17	-0.62	0.95	0.68

FE vs. HC	Yoon et al., 2012	0.17	-0.21	0.56	0.38
FE vs. HC	Lesh et al., 2015	0.21	-0.39	0.81	0.49
FE vs. HC	Woodward, 2016	0.37	-0.30	1.04	0.28
FE vs. HC	Braver et al., 1999	**0.55**	-0.14	1.24	0.12
At-risk Total		-0.04	-0.24	0.15	0.66
At-risk vs. HC	MacDonald, 2002	-0.39	-1.25	0.48	0.38
At-risk vs. HC	Barch et al., 2004	-0.30	-0.80	0.20	0.24
At-risk vs. HC	MacDonald, 2002	-0.22	-1.18	0.73	0.65
At-risk vs. HC	Lopez-Garcia et al., 2015	-0.14	-0.66	0.37	0.59
At-risk vs. HC	McClure et al., 2008	-0.08	-0.47	0.30	0.67
At-risk vs. HC	Sloat, 2007	0.08	-0.42	0.58	0.75
At-risk vs. HC	Lopez-Garcia et al., 2016	0.20	-0.56	0.97	0.60
At-risk vs. HC	Richard et al., 2013	0.25	-0.21	0.72	0.28

FE= first-episode schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 9, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within chronic groups?	Chronic	58.73	19	$6.13E^{-6}$	67.65	0.067
-Is there significant variance in effects within first-episode groups?	First-Episode	4.09	8	0.85	0	0
-Is there significant variance in effects within at-risk groups?	At-Risk	4.14	7	0.76	0	0
-Does the grouping variable explain significant variance in the model?	Total within	66.96	34	$6.30E^{-4}$		
	Mixed effects					
-Does the effect differ between subgroups?	Total between	9.79	2	0.0075		

Table 10

Subgroup Analyses: Random Effects Model of Diagnostic Groups by Illness Length for BX Errors.

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Chronic Total		**0.70**	0.58	0.82	0
Chronic vs. HC	Lopez-Garcia et al., 2016	-0.04	-0.81	0.73	0.92
Chronic vs. HC	Holmes et al., 2005	0.29	-0.65	1.23	0.54
Chronic vs. HC	Sheffield et al., 2014	0.46	0.09	0.83	0.014
Chronic vs. HC	MacDonald & Carter, 2003	**0.50**	-0.17	1.17	0.14
Chronic vs. HC	Reilly et al., 2017	**0.50**	0.35	0.66	$4.92E^{-11}$
Chronic vs. HC	Barch et al., 2008	**0.52**	0.10	0.94	0.015
Chronic vs. HC	Jones et al., 2010	**0.54**	0.14	0.95	0.0088
Chronic vs. HC	Ceccherini-Nelli et al., 2007	**0.56**	-0.12	1.24	0.11
Chronic vs. HC	Woodward, 2016	**0.56**	0.007	1.11	0.047
Chronic vs. HC	Chung et al., 2011	**0.58**	0.092	1.07	0.020
Chronic vs. HC	Edwards et al., 2010	**0.63**	-0.041	1.30	0.066

Chronic vs. HC	Sheffield et al., 2014	**0.68**	0.31	1.06	$3.02E^{-4}$
Chronic vs. HC	Todd et al., 2014	**0.70**	0.27	1.14	0.0015
Chronic vs. HC	Sheffield et al., 2015	**0.71**	0.31	1.11	$5.40E^{-4}$
Chronic vs. HC	Gold et al., 2012; Henderson et al., 2012	**0.74**	0.50	0.99	$2.21E^{-9}$
Chronic vs. HC	MacDonald, 2002	**0.75**	-0.17	1.68	0.11
Chronic vs. HC	Zhang et al., 2015	**0.79**	0.66	0.93	0
Chronic vs. HC	Perlstein et al., 2003	***0.85***	0.13	1.57	0.020
Chronic vs. HC	MacDonald, 2002	***0.88***	-0.10	1.86	0.079
Chronic vs. HC	Stratta et al., 2000	***1.14***	0.49	1.80	$6.53E^{-4}$
Chronic vs. HC	Becker, 2012	***1.21***	0.71	1.70	$2.04E^{-6}$
Chronic vs. HC	Lopez-Garcia et al., 2015	***1.84***	1.29	2.39	$5.60E^{-11}$
FE Total		**0.75**	0.58	0.92	0
FE vs. HC	Woodward, 2016	0.47	-0.20	1.14	0.17
FE vs. HC	Lesh et al., 2015	**0.58**	-0.024	1.19	0.060
FE vs. HC	Richard et al., 2013	**0.59**	0.19	0.99	0.0035

FE vs. HC	Barch et al., 2001	**0.64**	-0.13	1.40	0.10
FE vs. HC	Yoon et al., 2012	**0.65**	0.26	1.05	0.0012
FE vs. HC	Lesh et al., 2015	**0.76**	-0.055	1.57	0.068
FE vs. HC	Barch et al., 2003	***0.90***	0.52	1.28	$2.88E^{-6}$
FE vs. HC	Braver et al., 1999	***0.93***	0.22	1.65	0.010
FE vs. HC	Fornito et al., 2011	***1.04***	0.45	1.64	$6.07E^{-4}$
FE vs. HC	Yoon et al., 2014	***1.10***	0.31	1.89	0.0066
At-risk Total		0.36	0.15	0.58	0.0011
At-risk vs. HC	Richard et al., 2013	-0.08	-0.54	0.38	0.74
At-risk vs. HC	Sloat, 2007	-0.06	-0.56	0.44	0.82
At-risk vs. HC	Lopez-Garcia et al., 2016	0.20	-0.56	0.97	0.60
At-risk vs. HC	Paavola, 2013	0.35	-0.039	0.73	0.078
At-risk vs. HC	MacDonald, 2002	0.47	-0.40	1.33	0.29
At-risk vs. HC	Barch et al., 2004	**0.53**	0.02	1.04	0.043
At-risk vs. HC	MacDonald, 2002	**0.53**	-0.44	1.50	0.29

At-risk vs. HC	McClure et al., 2008	**0.58**	0.19	0.98	0.0038
At-risk vs. HC	Lopez-Garcia et al., 2015	***0.87***	0.34	1.41	0.0015

FE= first-episode schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics, very large effects in underlined bold italics

Table 10, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within chronic groups?	Chronic	38.93	21	0.010	46.06	0.030
-Is there significant variance in effects within first-episode groups?	First-Episode	4.43	9	0.88	0	0
-Is there significant variance in effects within at-risk groups?	At-Risk	11.64	8	0.17	31.29	0.033
-Does the grouping variable explain	Total within	55.01	38	0.037		

significant variance in the model?					
	Mixed effects				
-Does the effect differ between subgroups?	Total between	8.55	2	0.014	

Table 11

Subgroup Analyses: Random Effects Model of Diagnostic Groups by Illness Length for d'context.

Subgroup	Study Name	Hedges' *g*	Lower CI	Upper CI	*p*-value
Chronic Total		***-0.94***	-1.11	-0.77	0
Chronic vs. HC	Sheffield et al., 2014	***-1.83***	-2.26	-1.40	0
Chronic vs. HC	Cohen et al., 1999	***-1.47***	-1.96	-0.97	$5.25E^{-9}$
Chronic vs. HC	Todd et al., 2014	***-1.39***	-1.86	-0.92	$6.00E^{-9}$
Chronic vs. HC	MacDonald, 2002	***-1.37***	-2.41	-0.32	0.010
Chronic vs. HC	Holmes et al., 2005	***-1.36***	-2.40	-0.31	0.011
Chronic vs. HC	Ceccherini-Nelli et al., 2007	***-1.27***	-2.01	-0.54	$6.67E^{-4}$
Chronic vs. HC	MacDonald, 2002	***-1.25***	-2.23	-0.28	0.012
Chronic vs. HC	Merrill et al., 2017	***-1.20***	-1.77	-0.62	$4.33E^{-5}$
Chronic vs. HC	Dias et al., 2013	***-1.08***	-1.83	-0.33	0.0050
Chronic vs. HC	Gold et al., 2012; Henderson et al., 2012	***-1.03***	-1.28	-0.78	$1.11E^{-15}$
Chronic vs. HC	Chung et al., 2011	***-1.02***	-1.53	-0.51	$8.98E^{-5}$

Chronic vs. HC	Poppe et al., 2015	***-0.98***	-1.60	-0.36	0.0020
Chronic vs. HC	Sheffield et al., 2014	***-0.83***	-1.20	-0.45	$1.60E^{-5}$
Chronic vs. HC	Barch et al., 2008	**-0.71**	-1.13	-0.29	$9.81E^{-4}$
Chronic vs. HC	Poppe et al., 2016	**-0.66**	-1.06	-0.27	$9.95E^{-4}$
Chronic vs. HC	Reilly et al., 2017	**-0.66**	-0.82	-0.51	0
Chronic vs. HC	Delawalla et al., 2006	**-0.65**	-1.15	-0.16	0.0098
Chronic vs. HC	Woodward, 2016	**-0.59**	-1.14	-0.039	0.036
Chronic vs. HC	Jones et al., 2010	**-0.51**	-0.91	-0.10	0.014
Chronic vs. HC	Thoma & Daum, 2008	-0.45	-1.05	0.14	0.13
Chronic vs. HC	MacDonald & Carter, 2003	-0.20	-0.86	0.46	0.55
FE Total		***-0.94***	-1.15	-0.73	0
FE vs. HC	Lesh et al., 2015	***-1.29***	-2.15	-0.44	0.0029
FE vs. HC	Braver et al., 1999	***-1.20***	-1.94	-0.47	0.0014
FE vs. HC	Barch et al., 2003	***-1.18***	-1.57	-0.79	$2.96E^{-9}$
FE vs. HC	Woodward, 2016	***-1.03***	-1.73	-0.32	0.0043

FE vs. HC	Barch et al., 2001	***-0.80***	-1.58	-0.026	0.043
FE vs. HC	Richard et al., 2013	**-0.69**	-1.09	-0.29	$6.71E^{-4}$
FE vs. HC	Lesh et al., 2015	**-0.58**	-1.19	0.027	0.061
At-risk Total		-0.40	-0.56	-0.23	$3.70E^{-6}$
At-risk vs. HC	McClure et al., 2008	**-0.72**	-1.12	-0.32	$4.29E^{-4}$
At-risk vs. HC	MacDonald, 2002	**-0.70**	-1.69	0.29	0.16
At-risk vs. HC	Poppe et al., 2015	-0.49	-1.01	0.034	0.067
At-risk vs. HC	Barch et al., 2004	-0.40	-0.91	0.10	0.12
At-risk vs. HC	Richard et al., 2013	-0.37	-0.84	0.094	0.12
At-risk vs. HC	MacDonald, 2002	-0.34	-1.21	0.52	0.44
At-risk vs. HC	Delawalla et al., 2008	-0.29	-0.71	0.12	0.16
At-risk vs. HC	Sloat, 2007	-0.28	-0.79	0.22	0.27
At-risk vs. HC	Delawalla et al., 2006	-0.09	-0.55	0.37	0.70

FE= first-episode schizophrenia. HC=healthy controls. Total meta-analytic effects for each subgroup in gray. Hedges' *g*: medium effects in bold, large effects in bold italics, very large effects in underlined bold italics.

Table 11, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within chronic groups?	Chronic	55.56	20	$3.38E^{-5}$	64.00	0.086
-Is there significant variance in effects within first-episode groups?	First-Episode	5.60	6	0.47	0	0
-Is there significant variance in effects within at-risk groups?	At-Risk	5.12	8	0.74	0	0
-Does the grouping variable explain significant variance in the model?	Total within	66.29	34	$7.60E^{-4}$		
	Mixed effects					
-Does the effect differ between subgroups?	Total between	24.68	2	$4.40E^{-6}$		

Table 12

Symptom-Task Correlations: Random Effects Model for Authors' Conceptualizations of Positive Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.054	-0.12	0.012	0.11	17.35	16	0.36	7.77	0.0015
Becker, 2012	**-0.36**	-0.59	-0.07	0.016					
Stratta et al., 2000	**-0.32**	-0.68	0.16	0.18					
Woodward, 2016	-0.29	-0.70	0.26	0.30					
Woodward, 2016	-0.19	-0.49	0.16	0.28					
Lesh et al., 2015	-0.14	-0.52	0.29	0.53					
Reilly et al., 2017	-0.13	-0.23	-0.031	0.010					
MacDonald, 2002	-0.10	-0.55	0.38	0.69					
Delawalla et al., 2006	-0.10	-0.46	0.29	0.62					
Sheffield et al., 2014	-0.015	-0.21	0.18	0.88					
Gold et al., 2012	-0.010	-0.18	0.16	0.91					

Henderson et al., 2012									
Poppe et al., 2016	0	-0.28	0.28	1.00					
Barch et al., 2008	0.011	-0.25	0.27	0.94					
Fisher, 2016	0.052	-0.29	0.38	0.77					
Richard et al., 2013	0.064	-0.19	0.31	0.62					
Barch et al., 2003	0.11	-0.18	0.38	0.45					
Sheffield et al., 2015	0.19	-0.11	0.46	0.20					
Ceccherini-Nelli et al., 2007	**0.34**	-0.18	0.71	0.20					
BX_{Err} Total	0.049	-0.021	0.12	0.17	10.98	11	0.45	0	0
Ceccherini-Nelli et al., 2007	**-0.41**	-0.76	0.09	0.13					
Sheffield et al., 2015	-0.15	-0.43	0.14	0.31					
Fisher, 2016	-0.15	-0.46	0.20	0.40					
Barch et al., 2008	-0.010	-0.27	0.25	0.94					
Sheffield et al., 2014	-0.0090	-0.20	0.18	0.93					
Woodward, 2016	0.014	-0.49	0.51	0.96					

Barch et al., 2003	0.057	-0.23	0.33	0.70		
Reilly et al., 2017	0.090	-0.0090	0.19	0.073		
MacDonald, 2002	0.11	-0.38	0.56	0.67		
Woodward, 2016	0.13	-0.22	0.44	0.47		
Stratta et al., 2000	**0.30**	-0.18	0.66	0.22		
Lesh et al., 2015	**0.40**	-0.040	0.71	0.073		

Err=errors. Total meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold.

Table 13

Symptom-Task Correlations: Random Effects Model for Authors' Conceptualizations of Negative Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.15	-0.22	-0.069	0	20.80	16	0.19	23.07	0.0055
Lesh et al., 2015	***-0.78***	-0.94	-0.35	0.003					
Delawalla et al., 2006	***-0.51***	-0.77	-0.11	0.0016					
Richard et al., 2013	**-0.31**	-0.53	-0.059	0.016					
Barch et al., 2003	-0.29	-0.54	0	0.050					
Gold et al., 2012 Henderson et al., 2012	-0.23	-0.39	-0.061	0.008					
Woodward, 2016	-0.20	-0.51	0.14	0.24					
Woodward, 2016	-0.18	-0.63	0.36	0.52					
Stratta et al., 2000	-0.18	-0.57	0.29	0.46					
MacDonald, 2002	-0.14	-0.57	0.35	0.59					

Becker, 2012	-0.13	-0.40	0.16	0.37					
Reilly et al., 2017	-0.13	-0.23	-0.031	0.010					
Sheffield et al., 2014	-0.060	-0.25	0.13	0.55					
Barch et al., 2008	-0.013	-0.27	0.25	0.93					
Poppe et al., 2016	0	-0.28	0.28	1.00					
Ceccherini-Nelli et al., 2007	0.018	-0.45	0.48	0.95					
Sheffield et al., 2015	0.077	-0.22	0.36	0.61					
Fisher, 2016	0.080	-0.26	0.40	0.65					
BX_{Err} Total	0.12	0.017	0.21	0.021	14.64	11	0.20	24.87	0.0068
Sheffield et al., 2015	-0.097	-0.38	0.20	0.52					
Ceccherini-Nelli et al., 2007	-0.087	-0.53	0.40	0.74					
MacDonald, 2002	-0.024	-0.49	0.45	0.93					
Fisher, 2016	0.0020	-0.33	0.34	0.99					
Woodward, 2016	0.048	-0.46	0.53	0.86					
Sheffield et al., 2014	0.061	-0.13	0.25	0.54					

Reilly et al., 2017	0.11	0.011	0.21	0.029		
Barch et al., 2008	0.12	-0.14	0.37	0.37		
Woodward, 2016	0.14	-0.20	0.45	0.42		
Stratta et al., 2000	0.15	-0.30	0.55	0.52		
Barch et al., 2003	**0.45**	0.16	0.66	0.003		
Lesh et al., 2015	***0.70***	0.26	0.90	0.005		

Err=errors. Total meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold, large effects in bold italics.

Table 14

Symptom-Task Correlations: Random Effects Model for Authors' Conceptualizations of Disorganization Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.22	-0.34	-0.094	0.001	28.65	13	0.0073	54.63	0.029
Lesh et al., 2015	***-0.81***	-0.95	-0.38	0.002					
Becker, 2012	***-0.51***	-0.71	-0.23	0.001					
Barch et al., 2008	**-0.42**	-0.63	-0.16	0.003					
Richard et al., 2013	**-0.41**	-0.61	-0.16	0.002					
Barch et al., 2003	**-0.40**	-0.63	-0.11	0.008					
Stratta et al., 2000	-0.27	-0.64	0.20	0.26					
Delawalla et al., 2006	-0.23	-0.56	0.17	0.26					
Gold et al., 2012 Henderson et al., 2012	-0.14	-0.30	0.029	0.10					
Sheffield et al., 2014	-0.11	-0.30	0.085	0.27					

Sheffield et al., 2015	-0.081	-0.36	0.21	0.59					
Ceccherini-Nelli et al., 2007	-0.009	-0.47	0.46	0.97					
Poppe et al., 2016	0	-0.28	0.28	1.00					
Fisher, 2016	0.084	-0.26	0.41	0.64					
MacDonald, 2002	0.28	-0.23	0.67	0.28					
BX_{Err} Total	0.24	0.070	0.39	0.006	19.13	9	0.024	52.96	0.037
MacDonald, 2002	-0.11	-0.55	0.38	0.68					
Fisher, 2016	-0.040	-0.37	0.30	0.82					
Ceccherini-Nelli et al., 2007	0.014	-0.46	0.48	0.96					
Sheffield et al., 2015	0.074	-0.22	0.35	0.63					
Sheffield et al., 2014	0.12	-0.080	0.30	0.25					
Stratta et al., 2000	0.25	-0.22	0.63	0.29					
Barch et al., 2003	**0.38**	0.09	0.61	0.011					
Barch et al., 2008	**0.47**	0.21	0.67	0.001					
Edwards et al., 2010	***0.51***	0.053	0.79	0.030					

Lesh et al., 2015	***0.82***	0.39	0.96	0.002		

Err=errors. Total meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold, large effects in bold italics.

Table 15

Symptom-Task Correlations: Random Effects Model for Our Conceptualizations of Positive Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.02	-0.10	0.055	0.55	10.65	12	0.56	0	0
Becker, 2012	**-0.36**	-0.59	-0.07	0.016					
Stratta et al., 2000	**-0.32**	-0.68	0.16	0.18					
Lesh et al., 2015	-0.14	-0.52	0.29	0.53					
MacDonald, 2002	-0.10	-0.55	0.38	0.69					
Delawalla et al., 2006	-0.10	-0.46	0.29	0.62					
Sheffield et al., 2014	-0.015	-0.21	0.18	0.88					
Gold et al., 2012	-0.010	-0.18	0.16	0.91					
Henderson et al., 2012									
Poppe et al., 2016	0	-0.28	0.28	1.00					
Barch et al., 2008	0.011	-0.25	0.27	0.94					

Fisher, 2016	0.052	-0.29	0.38	0.77					
Richard et al., 2013	0.064	-0.19	0.31	0.62					
Barch et al., 2003	0.11	-0.18	0.38	0.45					
Ceccherini-Nelli et al., 2007	**0.34**	-0.18	0.71	0.20					
BX_{Err} Total	0.021	-0.11	0.15	0.75	8.02	7	0.33	12.69	0.004
Ceccherini-Nelli et al., 2007	**-0.41**	-0.76	0.12	0.13					
Fisher, 2016	-0.15	-0.46	0.20	0.40					
Barch et al., 2008	-0.010	-0.27	0.25	0.94					
Sheffield et al., 2014	-0.009	-0.20	0.18	0.93					
Barch et al., 2003	0.057	-0.23	0.33	0.70					
MacDonald, 2002	0.11	-0.38	0.55	0.67					
Stratta et al., 2000	**0.30**	-0.18	0.66	0.22					
Lesh et al., 2015	**0.40**	-0.04	0.71	0.073					

Err=errors. Total meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold.

Table 16

Symptom-Task Correlations: Random Effects Model for Our Conceptualizations of Negative Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.17	-0.27	-0.062	0.002	18.22	12	0.11	34.15	0.012
Lesh et al., 2015	***-0.78***	-0.94	-0.35	0.003					
Delawalla et al., 2006	***-0.51***	-0.77	-0.11	0.016					
Richard et al., 2013	**-0.31**	-0.53	-0.059	0.016					
Barch et al., 2003	-0.29	-0.54	0	0.050					
Gold et al., 2012 Henderson et al., 2012	-0.23	-0.39	-0.061	0.008					
Stratta et al., 2000	-0.18	-0.57	0.29	0.46					
MacDonald, 2002	-0.14	-0.57	0.35	0.59					
Becker, 2012	-0.13	-0.40	0.16	0.37					
Sheffield et al., 2014	-0.060	-0.25	0.13	0.55					

Barch et al., 2008	-0.013	-0.27	0.25	0.93					
Poppe et al., 2016	0	-0.28	0.28	1.00					
Ceccherini-Nelli et al., 2007	0.018	-0.45	0.48	0.95					
Fisher, 2016	0.080	-0.27	0.40	0.65					
BX_{Err} Total	0.16	-0.003	0.32	0.055	12.33	7	0.090	43.25	0.024
Ceccherini-Nelli et al., 2007	-0.087	-0.53	0.34	0.74					
MacDonald, 2002	-0.024	-0.49	0.45	0.93					
Fisher, 2016	0.0020	-0.33	0.34	0.99					
Sheffield et al., 2013	0.06	-0.13	0.25	0.54					
Barch et al., 2008	0.12	-0.14	0.37	0.37					
Stratta et al., 2000	0.15	-0.30	0.55	0.52					
Barch et al., 2003	**0.45**	0.16	0.66	0.003					
Lesh et al., 2015	***0.70***	0.26	0.87	0.005					

Err=errors. Total meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold, large effects in bold italics.

Table 17

Symptom-Task Correlations: Random Effects Model for Our Conceptualizations of Disorganization Symptoms in Schizophrenia Patients

Study Name	Pearson's *r*	Lower CI	Upper CI	*p*-value	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
d'context Total	-0.23	-0.36	-0.10	0.001	27.91	12	0.006	57.01	0.033
Lesh et al., 2015	***-0.81***	-0.95	-0.38	0.002					
Becker, 2012	***-0.51***	-0.71	-0.23	0.001					
Barch et al., 2008	**-0.42**	-0.63	-0.16	0.003					
Richard et al., 2013	**-0.41**	-0.61	-0.16	0.002					
Barch et al., 2003	**-0.40**	-0.63	-0.11	0.008					
Stratta et al., 2000	-0.27	-0.64	0.20	0.26					
Delawalla et al., 2006	-0.23	-0.56	0.17	0.26					
Gold et al., 2012	-0.14	-0.30	0.029	0.10					
Henderson et al., 2012									
Sheffield et al., 2014	-0.11	-0.30	0.085	0.27					

Ceccherini-Nelli et al., 2007	-0.009	-0.47	0.46	0.97					
Poppe et al., 2016	0	-0.28	0.28	1.00					
Fisher, 2016	0.084	-0.26	0.41	0.64					
MacDonald, 2002	0.28	-0.23	0.67	0.28					
BX_{Err} Total	**0.26**	0.08	0.43	0.006	18.11	8	0.020	55.83	0.044
MacDonald, 2002	-0.11	-0.55	0.38	0.68					
Fisher, 2016	-0.040	-0.37	0.30	0.82					
Ceccherini-Nelli et al., 2007	0.014	-0.46	0.48	0.96					
Sheffield et al., 2014	0.12	-0.080	0.30	0.25					
Stratta et al., 2000	0.25	-0.22	0.63	0.29					
Barch et al., 2003	**0.38**	0.09	0.61	0.011					
Barch et al., 2008	**0.47**	0.21	0.67	0.001					
Edwards et al., 2010	***0.51***	0.05	0.79	0.030					
Lesh et al., 2015	***0.82***	0.39	0.96	0.002					

Err=errors. Meta-analytic effects for each outcome in gray. Pearson's *r*: medium effects in bold, large effects in bold italics.

Table 18

Moderation Analyses: Random Effects Model of Cue Duration for AY Errors in Schizotypy-Spectrum Group versus Healthy Control Group.

AY Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (1000ms)	0.12	-0.01	0.25	0.07
Total Long (≤500ms)	0.38	0.23	0.53	$1.04E^{-6}$

Table 18, Continued

Question Being Answered	Model	*Q*	df(Q)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within the short cue condition?	Short Cue	15.53	17	0.56	0	0
-Is there significant variance in effects within the long cue condition?	Long Cue	20.18	12	0.064	40.55	0.026
-Does the moderator explain significant variance in the model?	Total within	35.72	29	0.18		
	Mixed effects					
-Does cue duration moderate the overall effect?	Total between	6.65	1	0.010		

Table 19

Moderation Analyses: Random Effects Model of Cue Duration for BX Errors in Schizotypy-Spectrum Group versus Healthy Control Group.

BX Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (1000ms)	**0.61**	0.44	0.78	$3.17E^{-12}$
Total Long (≤500ms)	**0.58**	0.48	0.67	0

Hedges' *g*: medium effects in bold.

Table 19, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within the short cue condition?	Short Cue	35.09	20	0.020	43.00	0.065
-Is there significant variance in effects within the long cue condition?	Long Cue	4.62	12	0.97	0	0
-Does moderator explain significant variance?	Total within	39.71	32	0.16		
	Mixed effects					
-Does cue duration moderate the overall effect?	Total between	0.13	1	0.72		

Table 20

Moderation Analyses: Random Effects Model of Cue Duration for d'context in Schizotypy-Spectrum Group versus Healthy Control Group.

d'context	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (1000ms)	**-0.73**	-0.92	-0.54	$8.55E^{-14}$
Total Long (≤500ms)	***-0.86***	-1.07	-0.66	$2.22E^{-16}$

Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 20, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within the short cue condition?	Short Cue	44.36	18	$5.12E^{-4}$	59.42	0.10
-Is there significant variance in effects within the long cue condition?	Long Cue	36.01	13	$5.92E^{-4}$	63.90	0.079
-Does moderator explain significant variance?	Total within	80.37	31	$2.94E^{-6}$		
	Mixed effects					
-Does cue duration moderate the overall effect?	Total between	0.88	1	0.35		

Table 21

Moderation Analyses: Random Effects Model of Cue-Probe Interval for AY Errors in Schizotypy-Spectrum Group versus Healthy Control Group.

AY Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (<3500ms)	0.39	0.27	0.52	$4.90E^{-10}$
Total Long (≥3500ms)	0.16	-0.02	0.33	0.08

Table 21, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within short cue-probe condition?	Short Cue-Probe Interval	20.39	21	0.50	0	0
-Is there significant variance in effects within long cue-probe condition?	Long Cue-Probe Interval	54.90	26	$7.80E^{-4}$	52.65	0.087
-Does the moderator explain significant variance in the model?	Total within	75.30	47	0.0055		
	Mixed effects					
-Does cue-probe interval moderate the overall effect?	Total between	4.74	1	0.029		

Table 22

Moderation Analyses: Random Effects Model of Cue-Probe Interval for BX Errors in Schizotypy-Spectrum Group versus Healthy Control Group.

BX Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (<3500ms)	**0.60**	0.48	0.73	0
Total Long (≥3500ms)	**0.57**	0.47	0.67	0

Hedges' *g*: medium effects in bold.

Table 22, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within short cue-probe condition?	Short Cue-Probe Interval	12.37	21.00	0.93	0	0
-Is there significant variance in effects within long cue-probe condition?	Long Cue-Probe Interval	27.35	27.00	0.45	1.26	$9.99E^{-4}$
-Does moderator explain significant variance?	Total within	39.72	48.00	0.80		
	Mixed effects					
-Does cue-probe interval moderate the overall effect?	Total between	0.13	1	0.72		

Table 23

Moderation Analyses: Random Effects Model of Cue-Probe Interval for d'context in Schizotypy-Spectrum Group versus Healthy Control Group

.

d'context	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total Short (<3500ms)	***-0.82***	-1.03	-0.61	$7.33E^{-15}$
Total Long (≥3500ms)	**-0.71**	-0.85	-0.56	0

Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 23, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within short cue-probe condition?	Short Cue-Probe Interval	60.55	24	$5.34E^{-5}$	60.36	0.15
-Is there significant variance in effects within long cue-probe condition?	Long Cue-Probe Interval	35.62	26	0.099	27.00	0.031
-Does moderator explain significant variance?	Total within	96.17	50	$9.50E^{-5}$		
	Mixed effects					
-Does cue-probe interval moderate the overall effect?	Total between	0.81	1	0.37		

Table 24

Moderation Analyses: Random Effects Model of Task Type for AY Errors in Chronic Schizophrenia Group versus Healthy Control Group.

AY Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total AX-CPT	0.33	0.02	0.64	0.03
Total DPX	**0.50**	0.35	0.65	$4.97E^{-11}$

Hedges' *g*: medium effects in bold.

Table 24, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within AX-CPT results?	AX-CPT	42.28	11	$1.45E^{-5}$	73.98	0.21
-Is there significant variance in effects within DPX results?	DPX	17.95	8	0.022	55.44	0.024
-Does moderator explain significant variance?	Total within	60.23	19	$3.60E^{-6}$		
	Mixed effects					
-Does task type moderate the overall effect?	Total between	0.98	1	0.32		

Table 25

Moderation Analyses: Random Effects Model of Task Type for BX Errors in Chronic Schizophrenia Group versus Healthy Control Group.

BX Errors	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total AX-CPT	**0.68**	0.53	0.82	0
Total DPX	**0.76**	0.53	0.99	$5.97E^{-11}$

Hedges' *g*: medium effects in bold.

Table 25, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within AX-CPT results?	AX-CPT	10.06	13	0.69	0	0
-Is there significant variance in effects within DPX results?	DPX	40.70	8	$2.37E^{-6}$	80.34	0.081
-Does moderator explain significant variance?	Total within	50.76	21	0.00029		
	Mixed effects					
-Does task type moderate the overall effect?	Total between	0.36	1	0.55		

Table 26

Moderation Analyses: Random Effects Model of Task Type for d'context in Chronic Schizophrenia Group versus Healthy Control Group.

d'context	Hedges' *g*	Lower CI	Upper CI	*p*-value
Total AX-CPT	***-1.04***	-1.28	-0.80	0
Total DPX	**-0.78**	-0.98	-0.59	$2.00E^{-15}$

Hedges' *g*: medium effects in bold, large effects in bold italics.

Table 26, Continued

Question Being Answered	Model	*Q*	df(*Q*)	*p*-value	I^2	Tau^2
	Fixed effect					
-Is there significant variance in effects within AX-CPT results?	AX-CPT	37.36	14	$6.51E^{-4}$	62.52	0.13
-Is there significant variance in effects within DPX results?	DPX	10.60	5	0.060	52.85	0.027
-Does moderator explain significant variance?	Total within	47.96	19	$2.60E^{-4}$		
	Mixed effects					
-Does task type moderate the overall effect?	Total between	2.69	1	0.10		

Table 27

Meta-Regression: Random Effects Model of Total Number of Trials for AY Errors in Schizotypy-Spectrum Group versus Healthy Control Group

k=31 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.65	1.03	3.37	0.021
Total Number of Trials	-0.0024	-0.0045	-0.00030	0.025

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	5.01	1	0.03		
Goodness of fit: Test that unexplained variance is zero	62.05	29	0.0003	53.27%	0.06

	R^2 analog
Proportion of total between-study variance explained by model	0.12

Table 28

Meta-Regression: Random Effects Model of Total Number of Trials for BX Errors in Schizotypy-Spectrum Group versus Healthy Control Group

k=34 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.46	0.17	0.74	0.0019
Total Number of Trials	0.0012	-0.0005	0.0028	0.16

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	1.93	1	0.16		
Goodness of fit: Test that unexplained variance is zero	48.29	9	0.012	33.73%	0.026

	R^2 analog
Proportion of total between-study variance explained by model	0.14

Table 29

Meta-Regression: Random Effects Model of Total Number of Trials for d'context in Schizotypy-Spectrum Group versus Healthy Control Group

k=31 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	-0.4974	-0.808	-0.1868	0.0017
Total Number of Trials	-0.0016	-0.0031	-0.0001	0.0329

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	4.55	1	0.033		
Goodness of fit: Test that unexplained variance is zero	63.76	29	0.0002	54.52%	0.076

	R^2 analog
Proportion of total between-study variance explained by model	0.17

Table 30

Meta-Regression: Random Effects Model of Cue-Probe Interval for AY Errors in Schizotypy-Spectrum Group versus Healthy Control Group

k=49 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.3426	0.1299	0.554	0.0016
Total Number of Trials	-0	-0.0001	0	0.3206

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	0.99	1	32.06		
Goodness of fit: Test that unexplained variance is zero	75.97	47	0.0047	38.13%	0.049

	R^2 analog
Proportion of total between-study variance explained by model	0.00

Table 31

Meta-Regression: Random Effects Model of Cue-Probe Interval for BX Errors in Schizotypy-Spectrum Group versus Healthy Control Group

k=50 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.5739	0.4042	0.7435	0
Total Number of Trials	0	0	0	0.9325

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	0.01	1	0.9325		
Goodness of fit: Test that unexplained variance is zero	39/88	48	-/79	0	0

	R^2 analog
Proportion of total between-study variance explained by model	0.00

Table 32

Meta-Regression: Random Effects Model of Cue-Probe Interval for d'context in Schizotypy-Spectrum Group versus Healthy Control Group

k=52 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	-0.8388	-1.0674	-0.6102	0
Total Number of Trials	0	0	0.0001	0.4671

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	0.53	1	0.47		
Goodness of fit: Test that unexplained variance is zero	97.80	50	0.0001	48.88%	0.083

	R^2 analog
Proportion of total between-study variance explained by model	0.00

Table 33

Meta-Regression: Random Effects Model of Schizophrenia Patients' Length of Illness for AY Errors in Schizophrenia Group versus Healthy Control Group

k=9 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.5091	0.0786	0.9397	0.0205
Total Number of Trials	-0.0263	-0.0638	0.0112	0.169

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	1.89	1	0.17		
Goodness of fit: Test that unexplained variance is zero	18.65	7	0.0094	62.46%	0.089

	R^2 analog
Proportion of total between-study variance explained by model	0.42

Table 34

Meta-Regression: Random Effects Model of Schizophrenia Patients' Length of Illness for BX Errors in Schizophrenia Group versus Healthy Control Group

k=11 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	0.7299	0.206	1.2539	0.0063
Total Number of Trials	0.0049	-0.0393	0.0492	0.8278

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	0.05	1	0.83		
Goodness of fit: Test that unexplained variance is zero	32.01	9	0.0002	71.89%	0.17

	R^2 analog
Proportion of total between-study variance explained by model	0.00

Table 35

Meta-Regression: Random Effects Model of Schizophrenia Patients' Length of Illness for d'context in Schizophrenia Group versus Healthy Control Group

k=9 studies	Coefficient	95% Lower CI	95% Upper CI	*p*-value
Intercept	-0.7201	-1.1266	-0.3135	0.0005
Total Number of Trials	-0.0303	-0.0634	0.0028	0.0731

	Q	df(*Q*)	*p*-value	I^2	Tau^2
Test that all coefficients (excluding intercept) are zero	3.21	1	0.073		
Goodness of fit: Test that unexplained variance is zero	9.23	7	0.24	24.19%	0.032

	R^2 analog
Proportion of total between-study variance explained by model	0.58

Table 36

Trim-and-Fill Results: Adjusted Results for All Schizotypy-Spectrum Groups versus Healthy Control Groups Using a Random Effects Model.

	# Studies Trimmed	Hedges' *g* Estimate	Lower Limit	Upper Limit	*Q*-Value
Short AY Errors					
Observed Values		0.33	0.19	0.46	33.59
Adjusted Values	3	0.37	0.23	0.51	40.36
Short BX Errors					
Observed Values		0.61	0.50	0.71	22.22
Adjusted Values	5	0.68	0.55	0.81	41.19
Short d'Context					
Observed Values		-0.84	-1.06	-0.63	96.00
Adjusted Values	4	-0.97	-1.20	-0.75	142.78
Long AY Errors					
Observed Values		0.0034	-0.17	0.18	82.07
Adjusted Values	9	0.24	0.065	0.41	143.33
Long BX Errors					
Observed Values		0.59	0.47	0.72	44.53
Adjusted Values	7	0.47	0.34	0.60	75.74
Long d'Context					

Observed Values		-0.76	-0.91	-0.61	67.11
Adjusted Values	5	-0.66	-0.82	-0.50	99.51

Table 37

Trim-and-Fill Results: Adjusted Results for Schizophrenia Patient Groups versus Psychiatric Comparison Groups Using a Random Effects Model.

	# Studies Trimmed	Hedges' *g* Estimate	Lower Limit	Upper Limit	*Q*-Value
AY Errors					
Observed Values		0.017	-0.12	0.15	1.95
Adjusted Values	3	0.073	-0.049	0.29	5.06
BX Errors					
Observed Values		0.33	0.19	0.46	3.30
Adjusted Values	1	0.31	0.18	0.45	5.49
d'Context					
Observed Values		-0.73	-1.07	-0.40	36.38
Adjusted Values	0	--	--	--	--

Table 38

Trim-and-Fill Results: Adjusted Results for Symptom-Task Performance Correlations Within Schizophrenia Patient Groups Using a Random Effects Model with Authors' Symptom Conceptualizations.

	# Studies Trimmed	Fisher's *z* Estimate	Lower Limit	Upper Limit	*Q*-Value
Disorganized Symptom-BX Error Correlations					
Observed Values		0.29	0.089	0.46	31.32
Adjusted Values	2	0.36	0.17	0.52	43.05
Disorganized Symptom-d'Context Correlations					
Observed Values		-0.25	-0.38	-0.11	41.65
Adjusted Values	3	-0.33	-0.46	-0.19	62.92
Negative Symptom-BX Error Correlations					
Observed Values		0.14	0.023	0.26	20.62
Adjusted Values	3	0.19	0.076	0.30	28.21
Negative Symptom-d'Context Correlations					
Observed Values		-0.17	-0.27	-0.077	31.59
Adjusted Values	3	-0.21	-0.30	-0.12	40.93
Positive Symptom-BX Error Correlations					
Observed Values		0.044	-0.034	0.12	11.68
Adjusted Values	0	--	--	--	--

Positive Symptom-d'Context Correlations					
Observed Values		-0.054	-0.12	0.015	18.10
Adjusted Values	0	--	--	--	--

Table 39

Results of p-curve Analysis for Authors' Symptom Correlations in Short and Combined/Unknown Delay Conditions

	Evidential value present?			Evidential value absent?				
Short Delay Correlation	Binomial Test	Continuous Test, Full (*p*'s<0.05)	Continuous Test, Half (*p*'s<0.025)	Binomial Test	Continuous Test, Full (*p*'s<0.05)	Statistical Power	# studies included/ excluded	Conclusion
Disorganized -d'context	*p*=0.125	*Z*=-3.5 *p*=0.0002	*Z*=-2.91 *p*=0.0018	*p*>0.9999	*Z*=2.14 *p*=0.9837	91%	3/4	Evidential value present
Negative -d'context	*p*=0.125	*Z*=- 2.18, *p*=0.0148	*Z*=-1.4, *p*=0.0806	*p*>0.9999	*Z*=0.92, *p*=0.8222	67%	3/4	Evidential value present
Positive -d'context	*p*=0.5	*Z*=-0.77, *p*=0.2212	*Z*=-0.14, p=0.4424	*p*>0.9999	*Z*=0.12, *p*=0.5478	42%	1/6	Evidential value not present/absent

Disorganized -BX errors	p=0.25	Z=-0.39, p=0.0973	Z=-0.34, p=0.3652	p>0.9999	Z=0.39, p=0.6519	53%	2/2	Evidential value not present/absent
Negative -BX errors	--	--	--	--	--	--	0/3	Could not be generated: all results p>0.05
Positive -BX errors	--	--	--	--	--	--	0/3	Could not be generated: all results p>0.05

Table 40

Results of p-curve Analysis for Authors' Symptom Correlations in Long and Combined/Unknown Delay Conditions

	Evidential value present?			Evidential value absent?				
Long Delay Correlation	Binomial Test	Continuous Test, Full (*p*'s<0.05)	Continuous Test, Half (*p*'s<0.025)	Binomial Test	Continuous Test, Full (*p*'s<0.05)	Statistical Power	# studies included/ excluded	Conclusion
Disorganized -d'context	*p*=0.0625	*Z*=-3.28, *p*=0.0005	*Z*=- 2.46, *p*=0.007	*p*>0.9999	*Z*=1.82, *p*=0.9655	83%	4/2	Evidential value present
Negative -d'context	*p*=0.75	*Z*=-1.59, *p*=0.0562	*Z*=- 2.2, *p*=0.0141	*p*=0.5012	*Z*=0.64, *p*=0.739	65%	2/4	Evidential value present
Positive -d'context	*p*=0.5	*Z*=-0.77, *p*=0.2212	*Z*=-0.14, *p*=0.4424	*p*>0.9999	*Z*=0.12, *p*=0.5478	42%	1/5	Evidential value not present/absent

Disorganized -BX errors	p=0.125	Z=-1.17, p=0.1218	Z=0.13, p=0.5527	p>0.9999	Z=0.12, p=0.5493	39%	3/1	Evidential value not present/absent
Negative -BX errors	p>0.9999	Z=0.6, p=0.7247	--	p=0.2947	Z=-1.05, p=0.1467	5%	1/0	Evidential value not present/absent
Positive -BX errors	--	--	--	--	--	--	0/3	Could not be generated: all results p>0.05

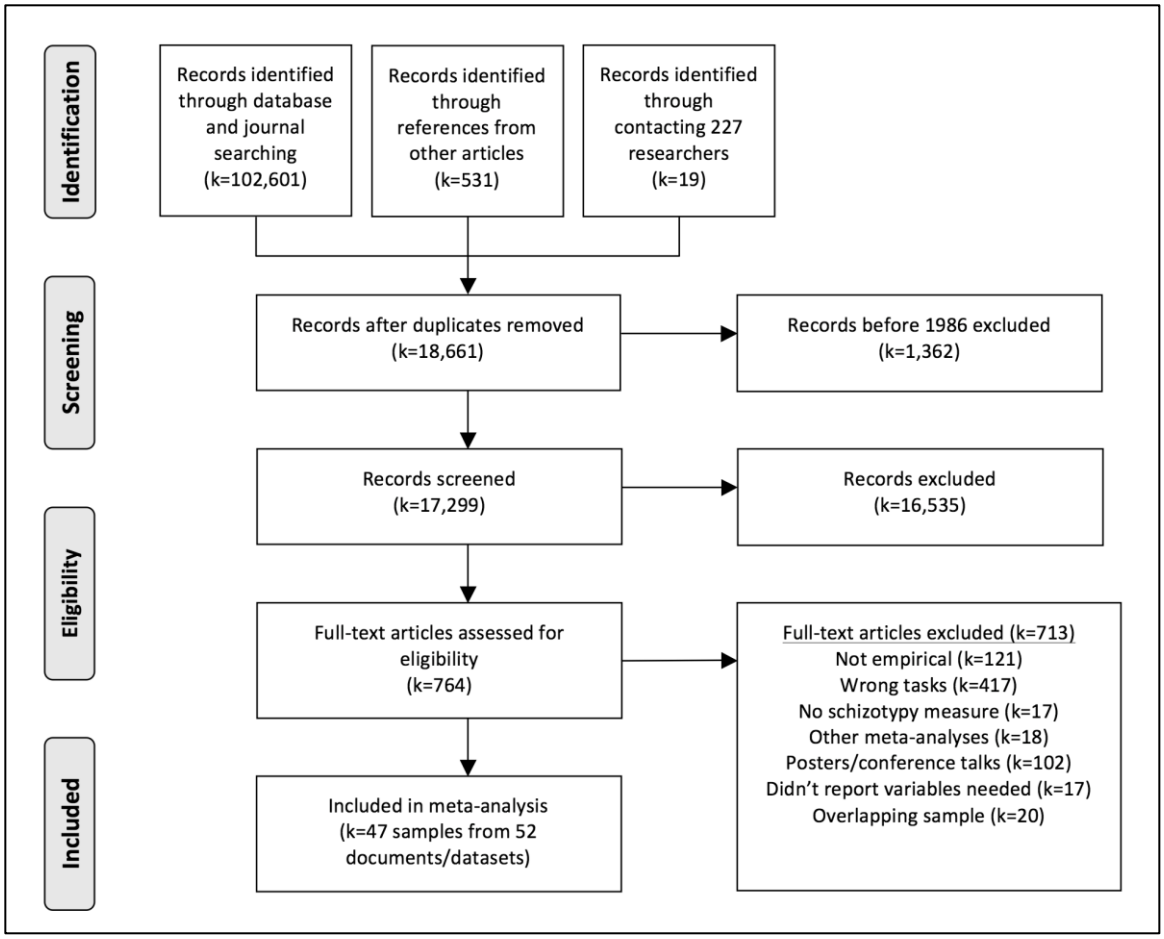

Figure 1. PRISMA Flow Diagram of Literature Search Results.

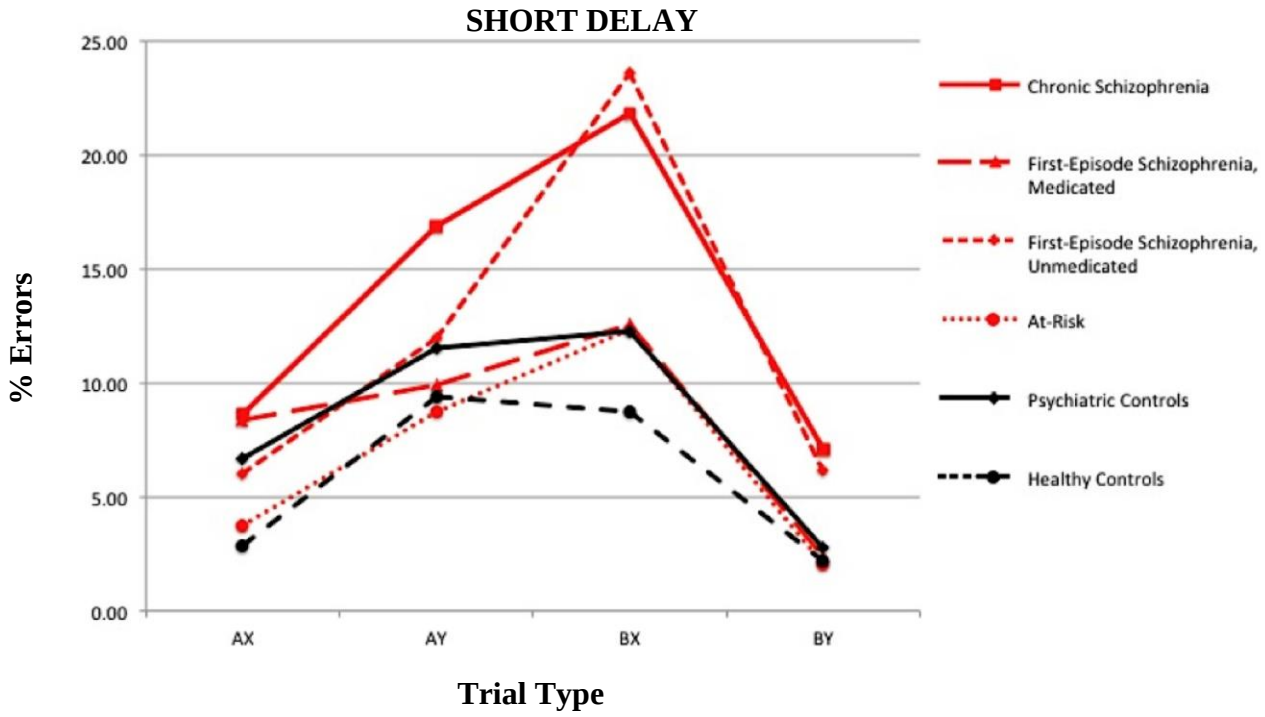

Figure 2. Point Estimate of Percent Errors by Trial Type for Short Delay Trials. Cue-probe interval was less than 3500ms. Points represent weighted means. Schizotypy-spectrum groups are portrayed in red and control groups are portrayed in black. At-risk group includes SPD, first-degree relatives of patients with schizophrenia, and psychometrically-identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, non-psychotic depression, bipolar disorder, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

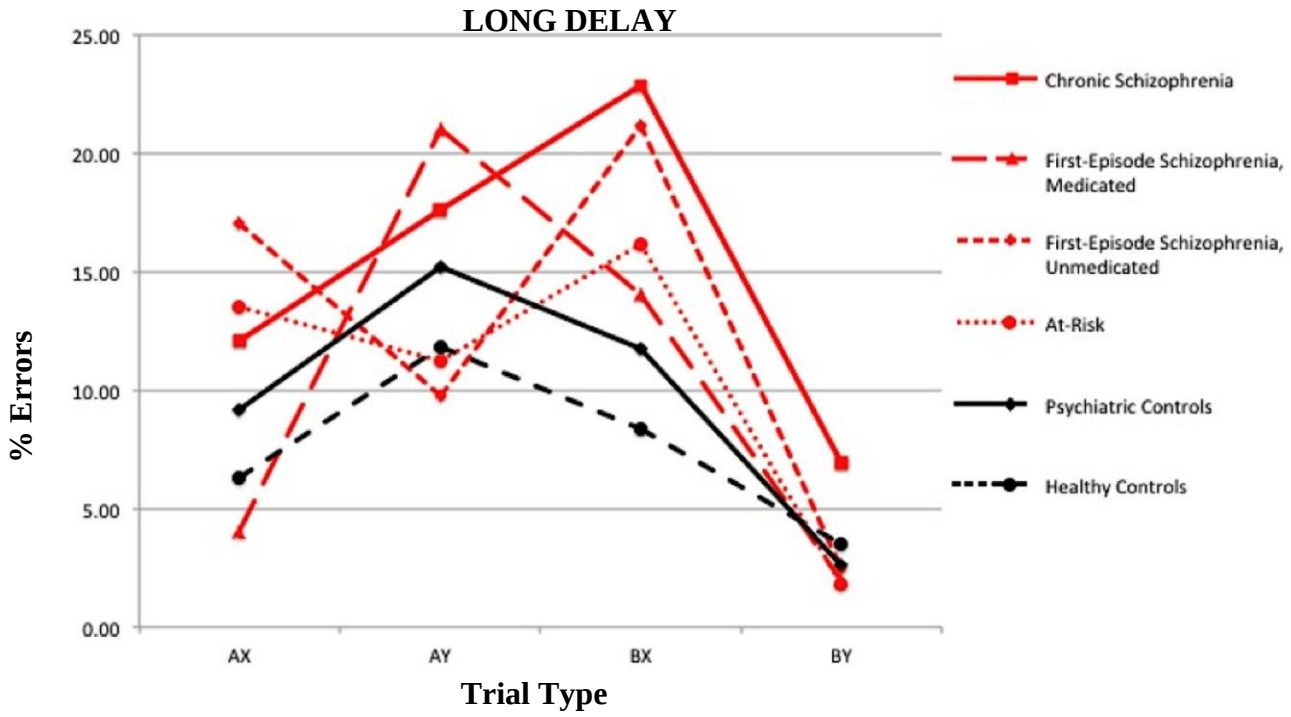

Figure 3. Point Estimate of Percent Errors by Trial Type for Long Delay Trials. Cue-probe interval was 3500ms or longer. Points represent weighted means. Schizotypy-spectrum groups are portrayed in red and control groups are portrayed in black. At-risk group includes SPD, ultra high-risk, first-degree relatives of patients with schizophrenia, and psychometrically Identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, non-psychotic depression, bipolar disorder, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

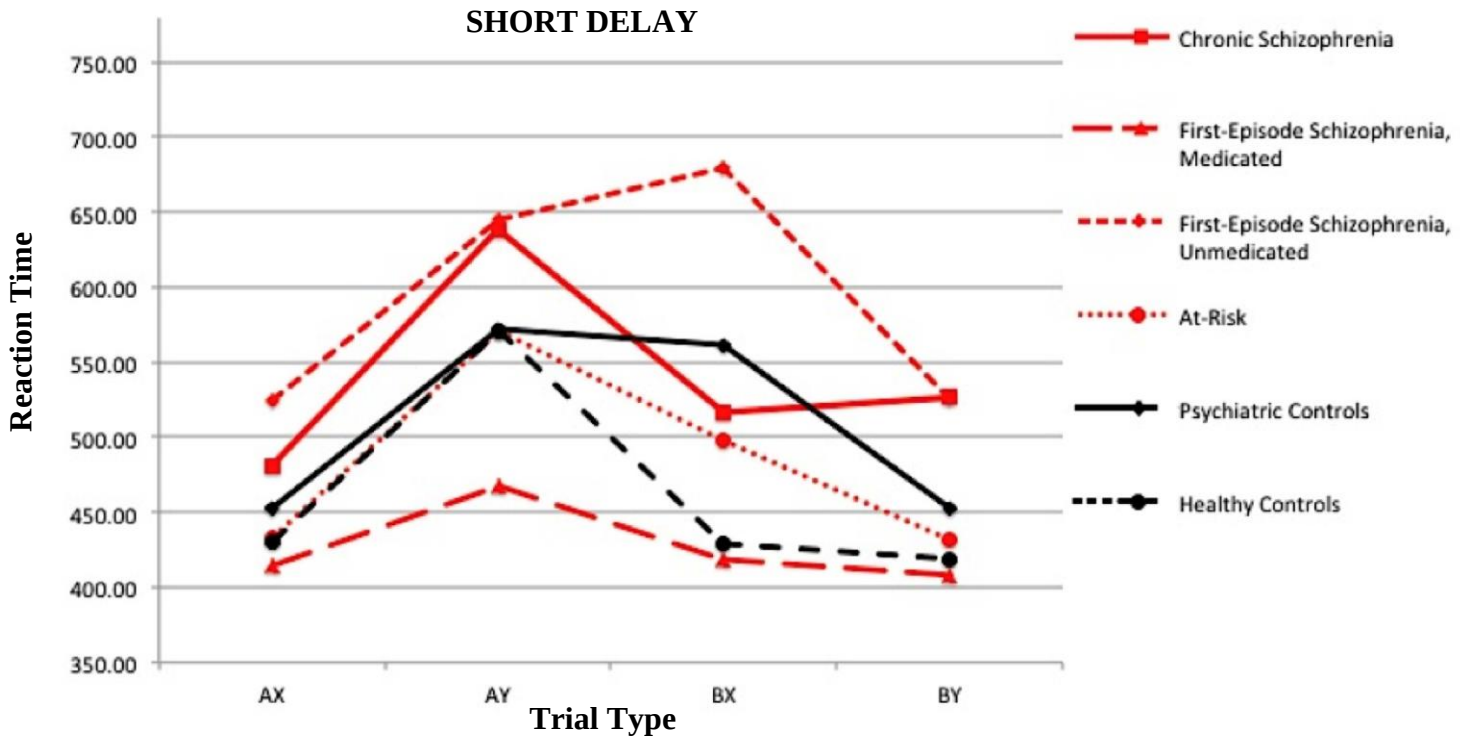

Figure 4. Point Estimate of Reaction Time in Milliseconds by Trial Type for Short Delay Trials. Cue-probe interval was less than 3500ms. Points represent weighted means. Schizotypy-spectrum groups are portrayed in red and control groups are in black. At-risk group includes SPD, first-degree relatives of patients with schizophrenia, and psychometrically-identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, bipolar disorder, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

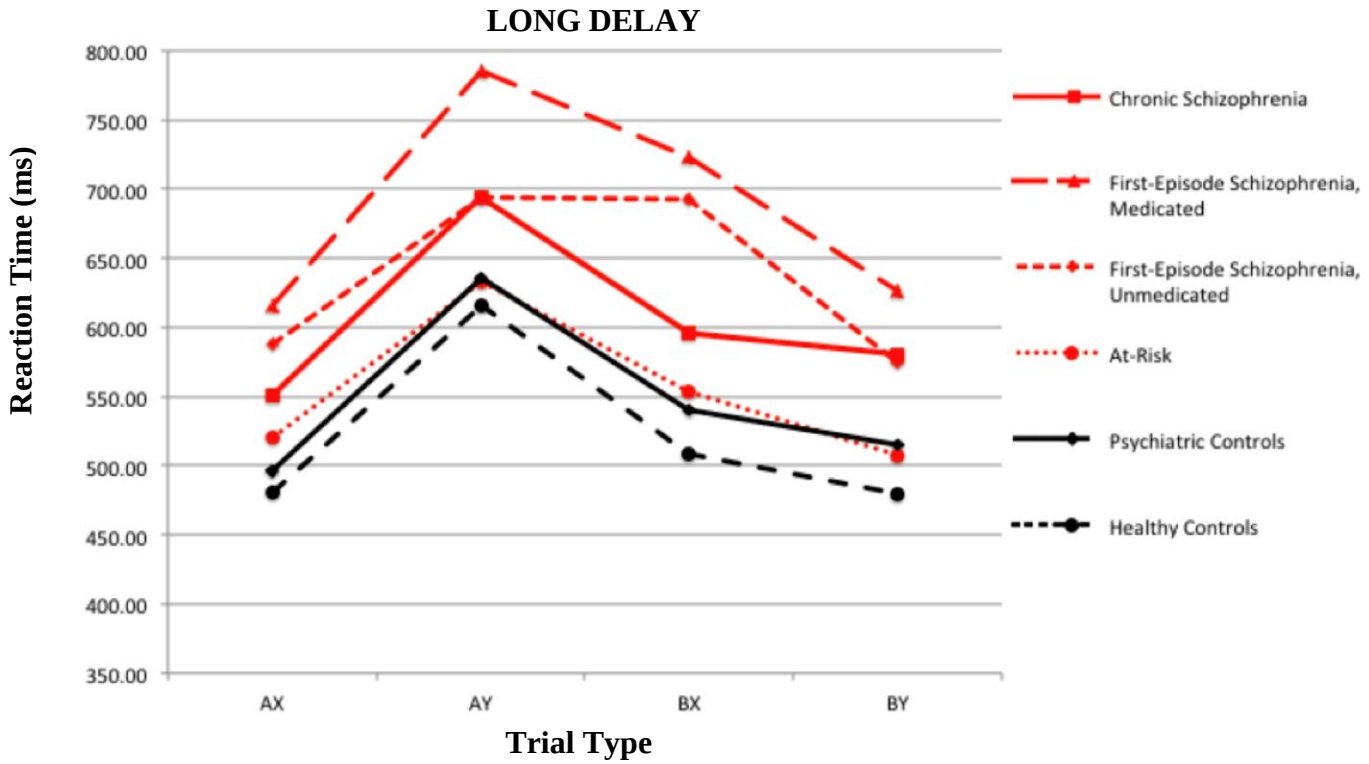

Figure 5. Point Estimate of Reaction Time in Milliseconds by Trial Type for Long Delay Trials. Cue-probe interval was 3500ms or longer. Points represent weighted means. Schizotypy-spectrum groups are portrayed in red and control groups are in black. At-risk group includes SPD, ultra high-risk, first-degree relatives of patients with schizophrenia, and psychometrically-identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, bipolar disorder, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

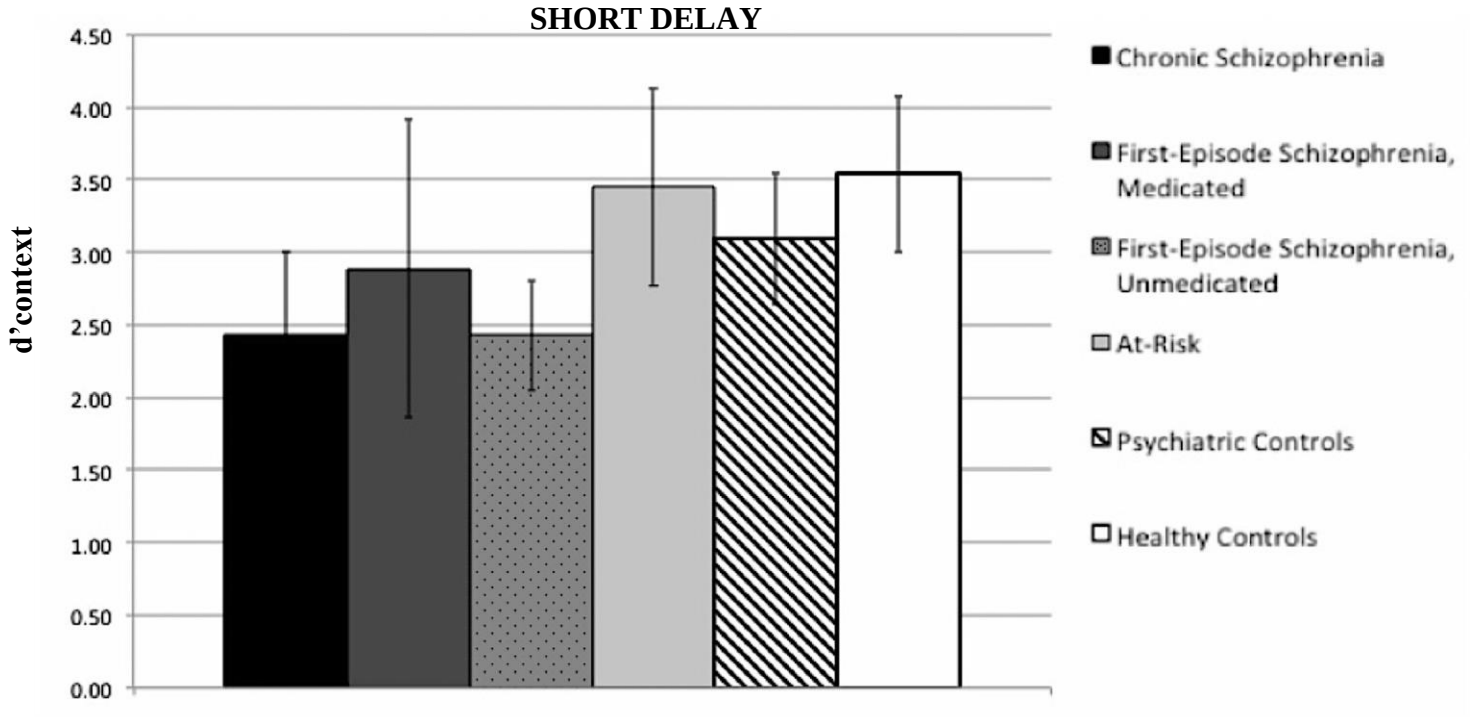

Figure 6. Point Estimate of D'context for Short Delay Trials. Cue-probe interval was less than 3500ms. Points represent weighted means and error bars represent standard deviations. At-risk group includes SPD, first-degree relatives of patients with schizophrenia, and psychometrically-identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, non-psychotic depression, bipolar disorder, bipolar disorder with psychotic features, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

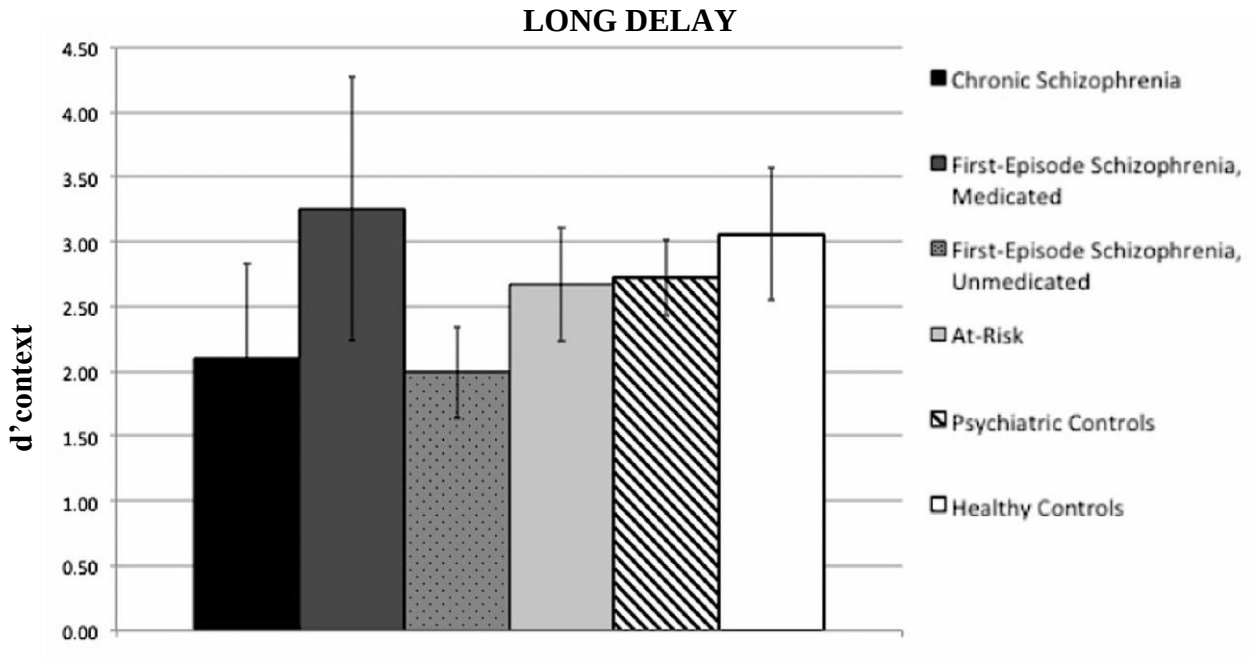

Figure 7. Point Estimate of D'context for Long Delay Trials. Cue-probe interval was 3500ms or longer. Points represent weighted means and error bars represent standard deviations. At-risk group includes SPD, ultra high-risk, first-degree relatives of schizophrenia patients, and psychometrically-identified schizotypy. Psychiatric controls include non-schizophrenia psychosis, non-psychotic depression, bipolar disorder, bipolar disorder with psychotic features, non-Cluster A personality disorders, and psychometrically-identified vulnerability to depression.

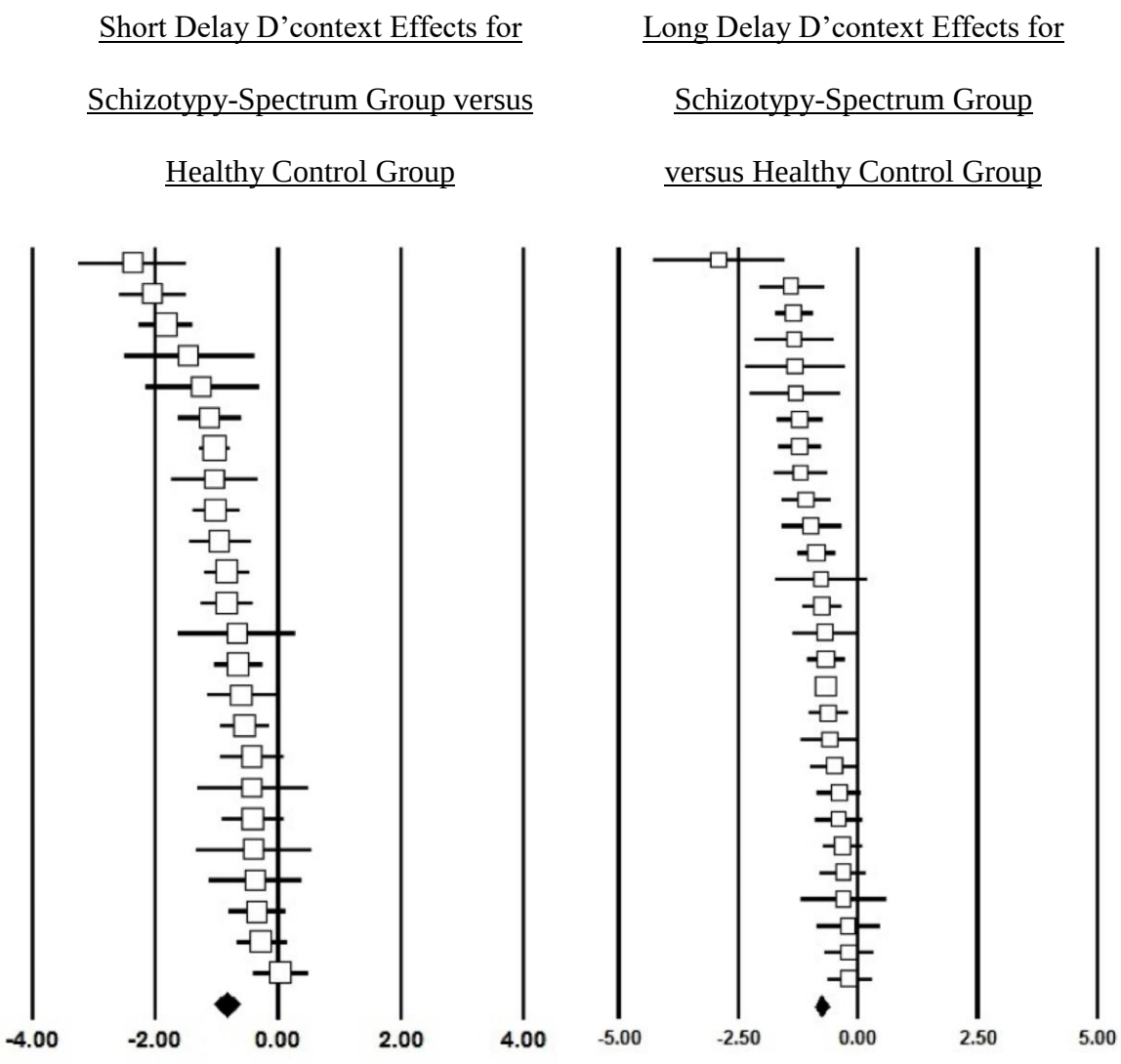

Figure 8. Forest Plots of Short and Long D'context Effects for Schizotypy-Spectrum Groups versus Healthy Control Groups. Short delay condition pictured on left and long delay condition on right. Note that forest plots for the two delay conditions are on different scales. The center of each white box represents the Hedges' *g* estimate for a single study, the size of the white box represents the relative weight given to the study according to inverse variance, and the bars represent the 95% confidence intervals on the Hedges' *g* estimate for that study. Effects in the negative direction indicate that the

schizotypy group performed worse than healthy control group. The black diamond at the bottom of the plot represents the summary statistic, or the estimated Hedges' g for the overall effect, with the width of the diamond showing 95% confidence intervals on the overall effect.

D’context Effects for Schizophrenia Patients

vs. Psychiatric Comparison Patients

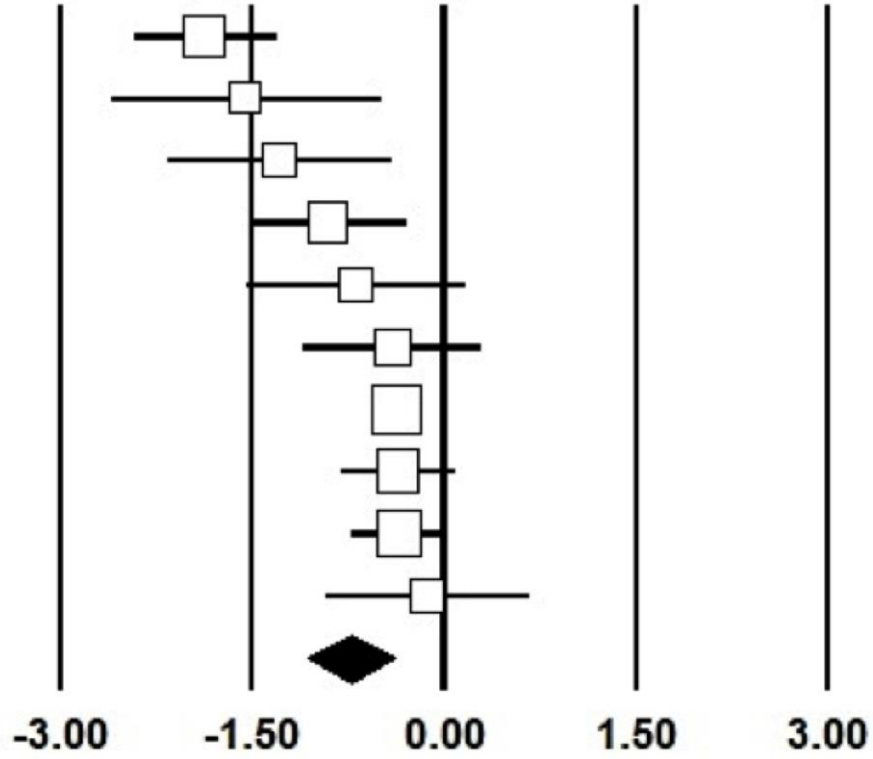

Figure 9. Forest Plot of D’context Effects for Patients with Schizophrenia versus Psychiatric Comparisons. The center of each white box represents the Hedges’ *g* estimate for a single study, the size of the white box represents the relative weight given to the study according to inverse variance, and the bars represent the 95% confidence intervals on the Hedges’ *g* estimate for that study. Effects in the negative direction indicate that the schizophrenia group performed worse than the psychiatric comparison group. The black diamond at the bottom of the plot represents the summary statistic, or the estimated Hedges’ *g* for the overall effect, with the width of the diamond indicating the 95% confidence intervals on the overall effect.

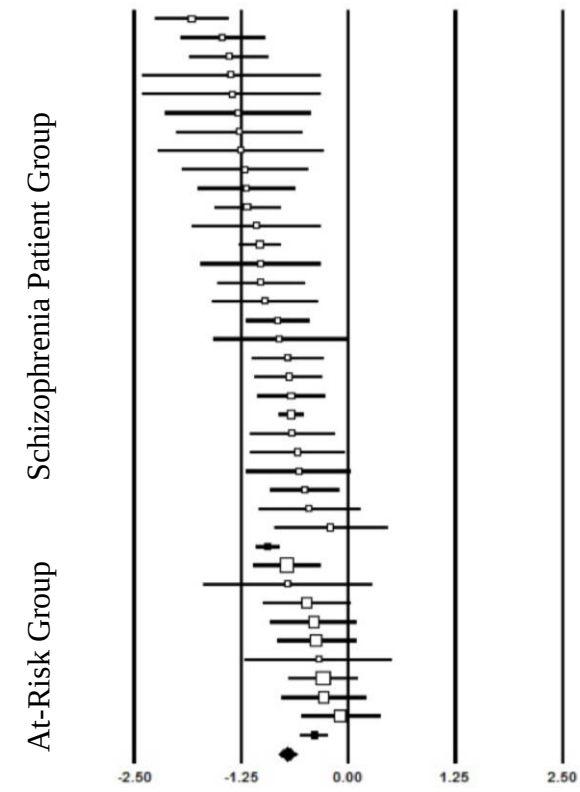

Figure 10. Forest Plot of D'context Effects for Patients with Schizophrenia versus Healthy Controls and At-Risk versus Healthy Controls. Effects for schizophrenia on top and at-risk at the bottom. The center of each white box represents Hedges' *g* estimate for single studies, the size of the white box represents the relative weight given according to inverse variance, and bars represent 95% confidence intervals on Hedges' *g* estimate for that study. Effects in negative direction indicate the schizotypy-spectrum group performed worse. Black squares represent summary statistics, estimated Hedges' *g* for the overall effect, for each subgroup with the width of the square showing 95% confidence intervals on the overall effect. The black diamond represents the overall effect across subgroups.

Correlation of Disorganized Symptoms with

D'context in Schizophrenia Patients

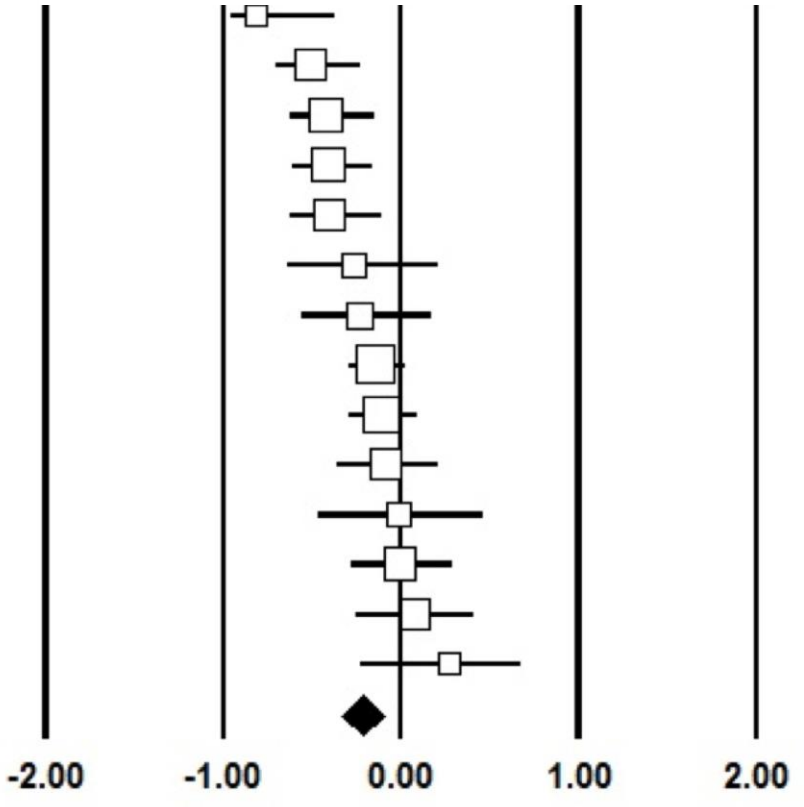

Figure 11. Forest Plot of Disorganized Symptom-D'context Correlations within Schizophrenia Groups. Authors' conceptualizations were used. The center of each white box represents Pearson's *r* estimate for single studies, the size of the white box shows relative weight given according to inverse variance, and bars represent 95% confidence intervals on Pearson's *r* estimate. Effects in negative direction indicate that more severe symptoms are associated with lower d'context scores (stronger CI deficits). The black diamonds represents the summary statistic, estimated Pearson's *r* for the overall effect, with the width of the diamond showing 95% confidence intervals on the overall effect.

Correlation of Negative Symptoms with

D'context in Schizophrenia Patients

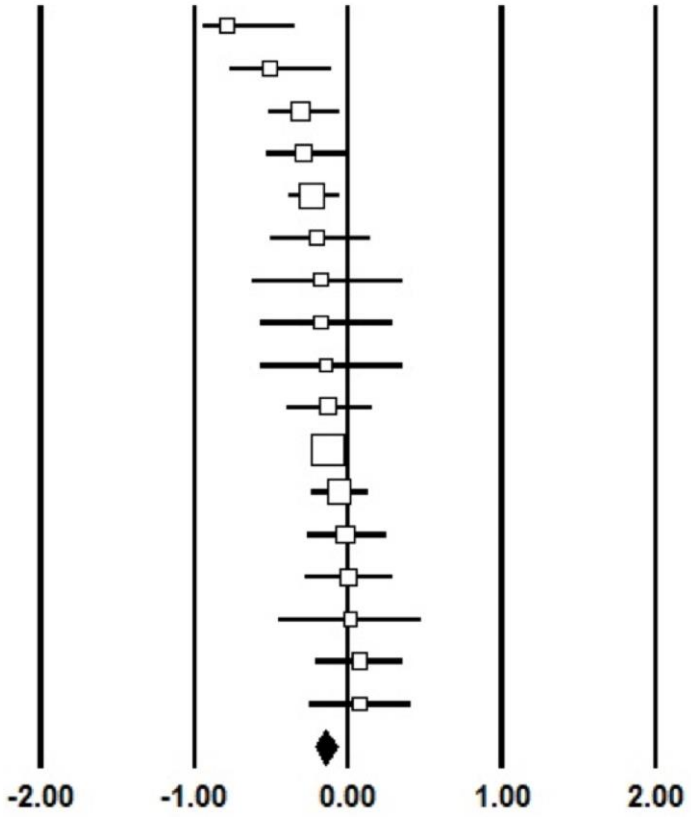

Figure 12. Forest Plot of Negative Symptom-D'context Correlations within Schizophrenia Groups. Authors' conceptualizations were used. The center of each white box represents Pearson's *r* estimate for single studies, the size of the white box shows relative weight given according to inverse variance, and bars represent 95% confidence intervals on Pearson's *r* estimate. Effects in negative direction indicate that more severe symptoms are associated with lower d'context scores (stronger CI deficits). The black diamonds represents the summary statistic, estimated Pearson's *r* for the overall effect, with the width of the diamond showing 95% confidence intervals on the overall effect.

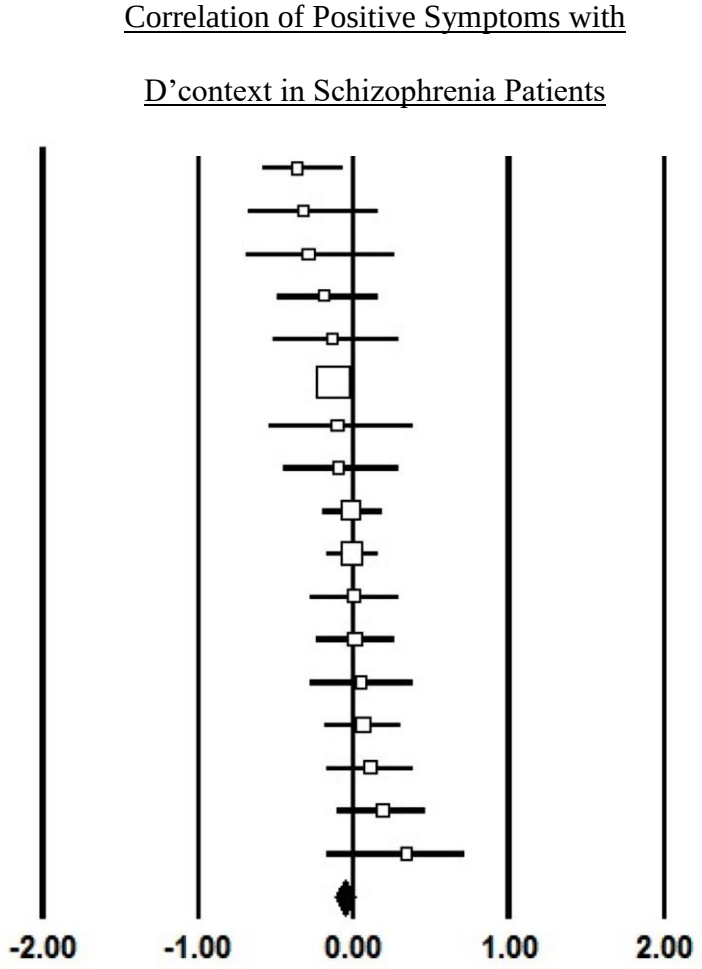

Figure 13. Forest Plot of Positive Symptom-D'context Correlations within Schizophrenia Groups. Authors' conceptualizations were used. The center of each white box represents Pearson's *r* estimate for single studies, the size of the white box shows relative weight given according to inverse variance, and bars represent 95% confidence intervals on Pearson's *r* estimate. Effects in negative direction indicate that more severe symptoms are associated with lower d'context scores (stronger CI deficits). The black diamonds represents the summary statistic, estimated Pearson's *r* for the overall effect, with the width of the diamond showing 95% confidence intervals on the overall effect.

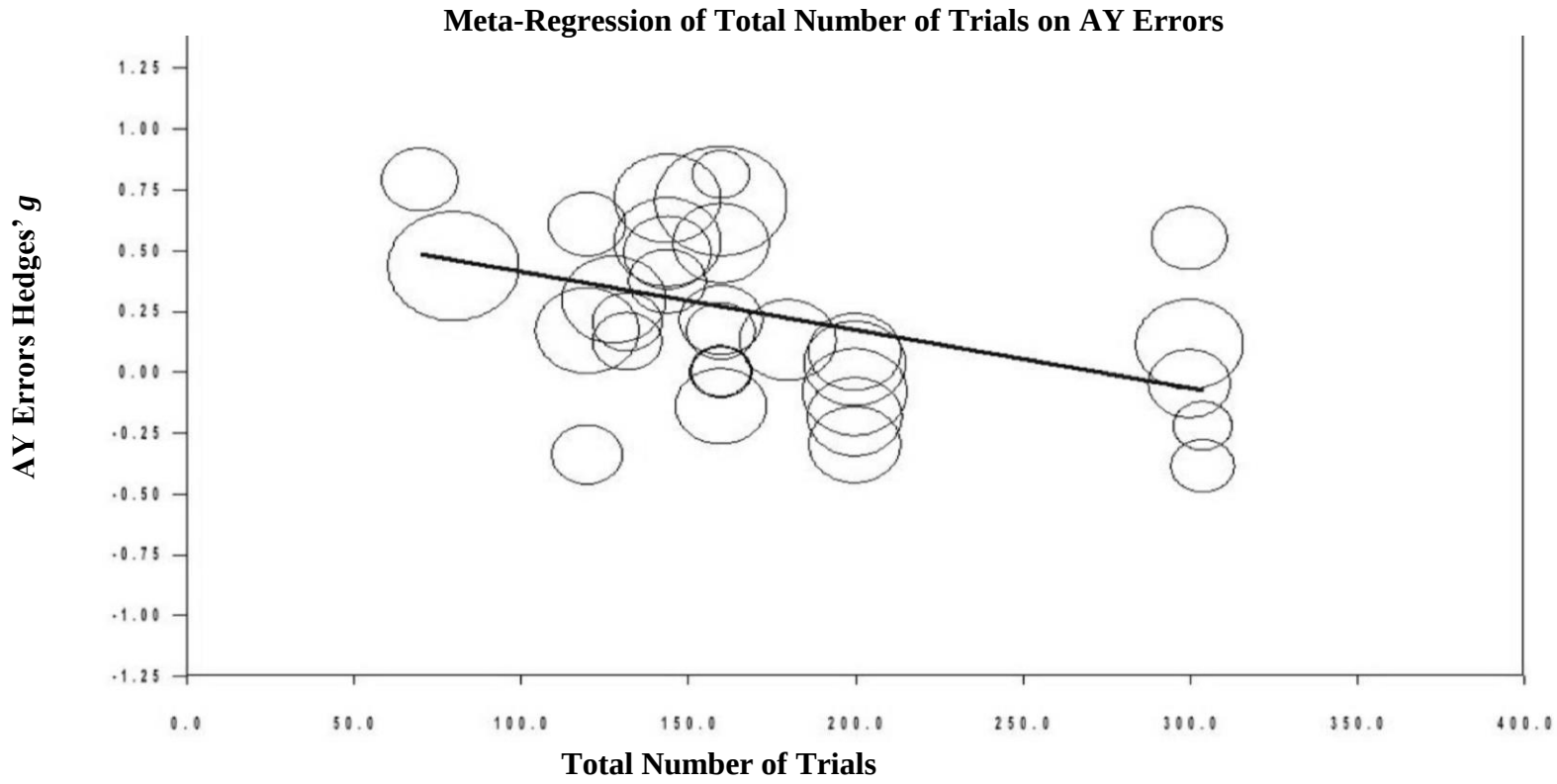

Figure 14. Meta-Regression Plot of Total Number of Trials Predicting Hedges' g for AY Errors in Schizotypy-Spectrum versus Healthy Controls. Circles represent individual studies, with the size of the circle showing relative weight given to the study.

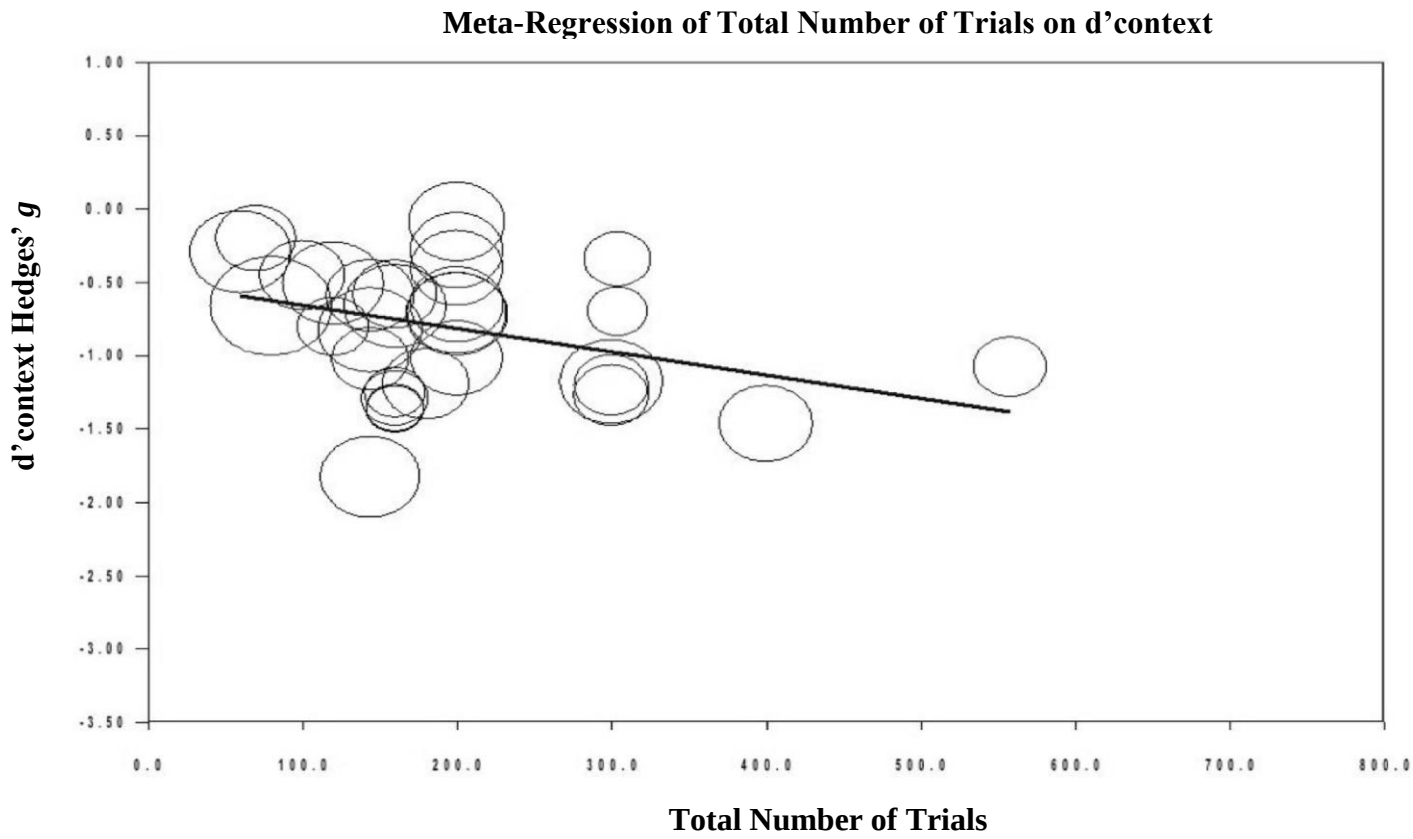

Figure 15. Meta-Regression Plot of Total Number of Trials Predicting Hedges' g for D'context in Schizotypy-Spectrum versus Healthy Controls. Circles represent individual studies, with the size of the circle showing relative weight given to the study.

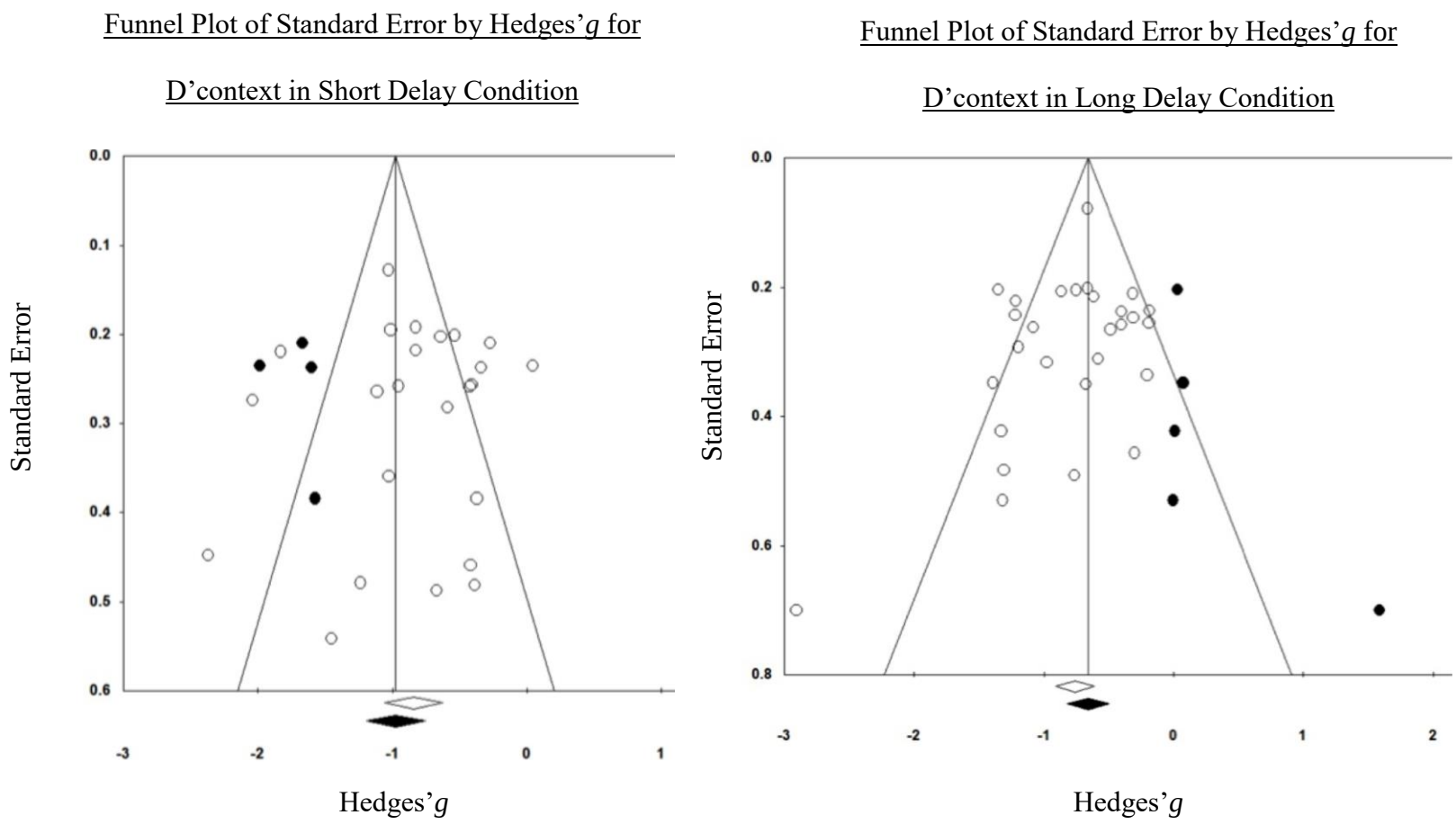

Figure 16. Funnel Plot of Hedges' g Effects for D'context in Short and Long Delay Conditions for All Schizotypy-Spectrum Groups versus Healthy Controls. Short delay condition pictured on left and long delay condition on right. Observed effects

from studies included in the meta-analysis are represented by white circles, and the observed point estimate is represented by a white diamond. Imputed effects from the trim-and-fill method are represented by black circles and the adjusted point estimate is represented by a black diamond.

Funnel Plot of Standard Error by Hedges'*g* for D'context

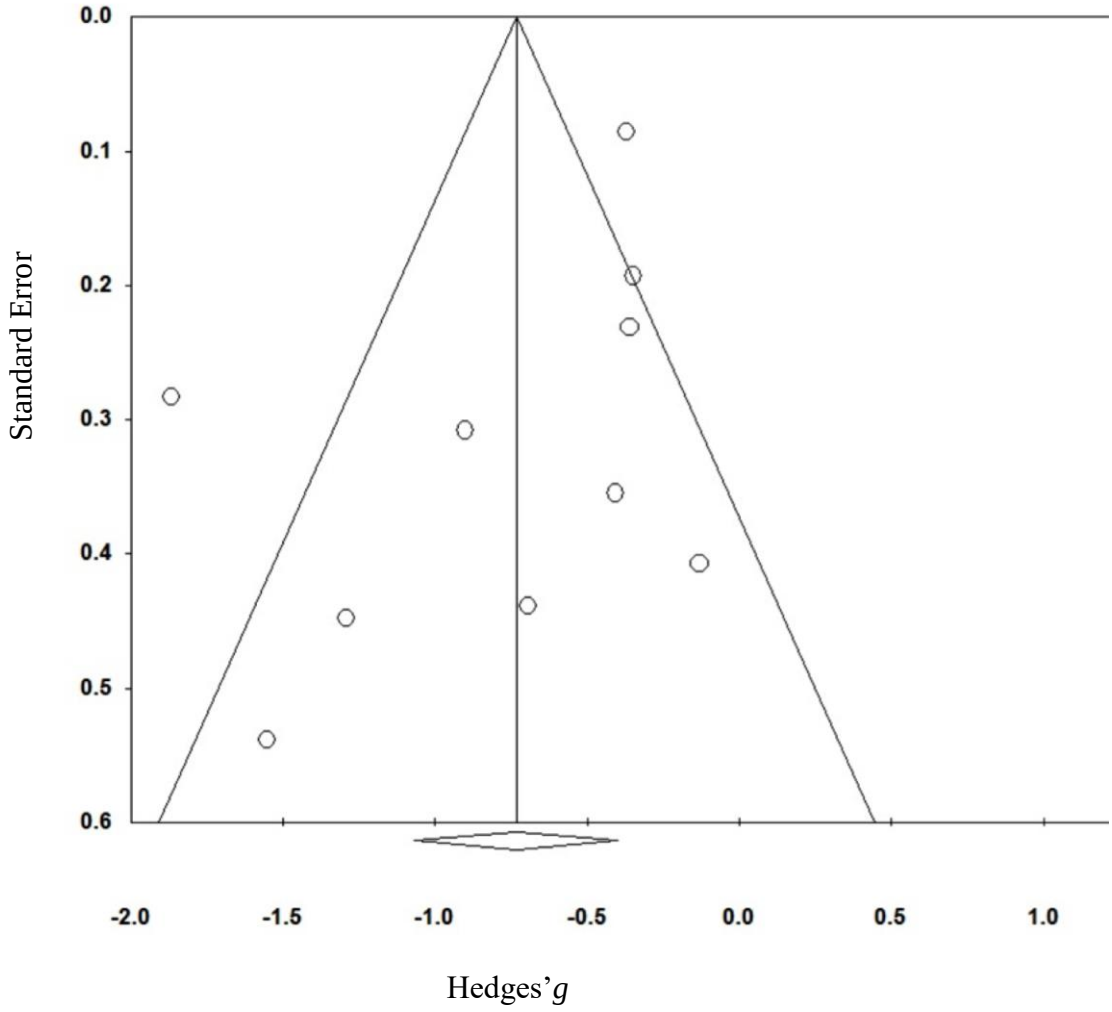

Figure 17. Funnel Plot of Hedges' g Effects for D'context for Patients with Schizophrenia versus Psychiatric Comparisons. Observed effects are represented by circles and the point estimate is represented by a diamond; no values were imputed using the trim-and-fill method.

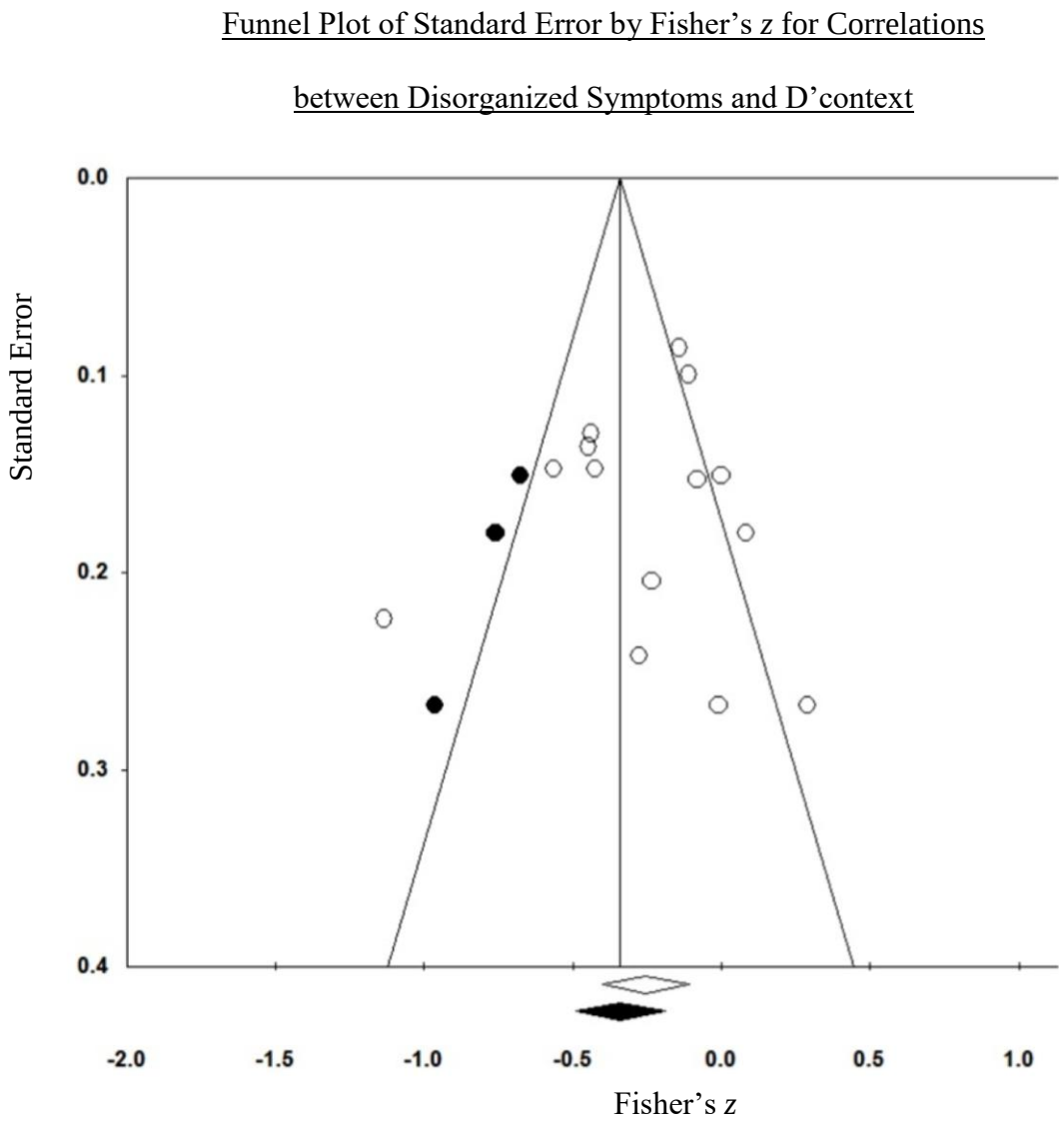

Figure 18. Funnel Plot of Fisher's z Effects for Correlations between Disorganized Symptoms and D'context within Schizophrenia Groups. Observed effects from studies included in the meta-analysis are represented by white circles, and the observed point estimate is represented by a white diamond. Imputed effects from the trim-and-fill method are represented by black circles and the adjusted point estimate is represented by a black diamond.

APPENDIX B

GENERAL METHODOLOGICAL DECISIONS

This appendix describes detailed methodological decisions that were made *a priori*, that is, before the analysis phase began. Whenever possible, decisions were informed by best practices in meta-analysis (Borenstein, 2009; Higgins & Green, 2011), by conventions in the CI literature, and/or by consultation with a clinical psychologist and expert schizotypy-spectrum researcher, Dr. Thomas Kwapil. General decisions that applied to multiple studies are described here. Basic information about included studies is provided in Appendix F and more specific information pertaining to individual studies is provided in Appendix G.

Inclusion/Exclusion of Studies and/or Participants

- When no control group was included, data were collected and summarized for point estimates only.
- Studies that used the correct AX-CPT task but had additional parameters or task goals present simultaneously (e.g., ignoring background noise during the task, or the addition of colored letters to add another level of salience) were excluded.
- Studies assessing CI in analogue conditions, based on psychoactive substances and genotyping were excluded from the proposed meta-analysis because they do not provide direct evidence of CI deficits in schizotypy. Genotyping studies were only included if the article or the authors provided data on schizotypy and control groups separately.

- When multiple documents were available for a study (e.g., a thesis document or unpublished manuscript that became a published article), all documents were searched and evaluated for appropriateness. When conflicting information was present, the document to be included was decided on a case-by-case basis via consultation with Dr. Kwapil. When non-overlapping information was available, all appropriate versions of the study were coded.
- Every published article (including supplemental material, when available) was searched for keywords "subset", "sample", "previously", "reported", and "described." When overlap was suspected, the full text of the article in question was re-read in a separate session from coding. When a study reported overlapping data with another study, the group or groups that had any degree of overlap in participants were excluded; that is, only one study was retained to avoid violating the assumption of independent data. Decisions on which articles to retain were made by consulting with Dr. Kwapil, and factors such as sample size, number and characteristics of groups included, and number of outcomes reported were used to guide selection. When the overlap between studies was unclear (e.g., study A overlaps with study B and study B overlaps with study C but it was not stated whether study A and C shared any participants), the authors were contacted for clarification. When no answer was received, we took a conservative approach: overlap was assumed and only one study was retained. When no overlap was reported in the text of the manuscript but it was apparent that two studies used the same group of participants (e.g., same means and standard deviations for demographic information), only one study was retained.

- Mood disorders with psychotic features and other non-schizophrenia psychoses were included as a psychiatric comparison group, not as a schizophrenia-spectrum group. Accordingly, studies including mood disorders with psychotic features in the schizophrenia group were excluded if data was not reported separately for these two groups.
- When psychotic disorders were present in some relatives of patients with schizophrenia, this group was not included (as it no longer constitutes an examination of unaffected relatives); however, relative groups were still included if the authors did not report screening for psychotic disorders.

Data Collection and Coding

- Quantitative data from articles text, tables, appendices, and supplementary information were recorded. Effect sizes were recorded directly from the text or computed from summary data (means and standard deviations) and from standard, parametric analyses without covariates or other corrections.
- When only significance limits were provided and no precise information was reported (e.g., $p < 0.05$), no data was coded.
- When precise information on magnitude of between-group differences was given but no direction was provided (e.g., patients did not differ from controls on BY errors, t=XXX), the direction was guessed according to theoretical predictions and other findings in the literature. This only happened for 3 analyses and the literature was quite consistent in these cases (4/5, 5/5, and 8/8 of the other studies had effects in the expected direction, respectively, for the 3 analyses).

- Only Pearson correlations were coded. To ensure that Spearman correlations were not included, each article was searched for the terms “Spearman” and “rho.” When a specific range of Pearson correlation values was given (e.g., *r* ranged from 0.005 to 0.18, the average of the range was coded; this was only the case for 2 studies.
- When data was only available in a figure, no attempt was made to extract or estimate values from the figure and instead asked authors for the data.
- When dichotomizing first-episode versus chronic illness status for patients with schizophrenia, a cutoff of 1.5 years was selected, according to precedence in the literature (Ellison-Wright, Glahn, Laird, Thelen, & Bullmore, 2008). If authors did not provide any information on length of illness or number of episodes, it was assumed that patients with schizophrenia were chronic.
- When dichotomizing medicated versus unmedicated schizophrenia groups, a group was considered medicated if any of the patients were taking antipsychotic medications.
- Dichotomization of cue duration at 500ms and cue-probe delay at 3500ms was determined a priori by looking at the distribution of durations for each variable across included studies. Those numbers were near a median split and made sense logically and based on the number of studies using each condition.
- When information for task parameters, error rates, reaction time, d’context, or symptom correlations were missing, authors were contacted to request the information.

- When studies cited another article to describe the task and it was clear the parameters were the same (e.g., “We used an exact replica of XXX’s task.”), information from the cited study was coded. When studies cited another article but the parameters were obviously not the same (e.g., the article in question manipulated short and long delay conditions but the article they cited did not), then information was not coded.
- When it was unclear whether an article’s reported sample size included participants who were dropped or excluded, a conservative approach was taking by subtracting this number out of the total sample.
- Some studies use a correction factor (a’) in the case of absolute error rates (e.g., 100% AX accuracy) when calculating d’context. Studies using d’context versus a’context were not differentiated because the two values correlate so highly, around 0.99 (Ceccherini-Nelli, Turpin-Crowther, & Crow, 2007).
- When the authors did not provide d’context scores, these values were not calculated from AX and BX group summary data. Calculating regular d-prime from averaged versus collapsed data yields different results. Average d' is computed by calculating d' for each participant and then averaging the group results, whereas collapsed d' is computed by first averaging the hits and the false alarms of participants, then calculating d' from the averaged proportions. Collapsed d’ tends to underestimate true average d’ and is particularly unreliable when participants’ average values are 1.5^{+} standard deviations apart, or when hit and false alarm rates are largely asymmetrical (Macmillan & Kaplan, 1985). Since both undesirable cases are true for this current dataset, collapsed d’context would likely underestimate true average d’context to a

large extent. Indeed, this was tested with raw datasets that were available: it was found that collapsed d'context underestimated average d' context by about 20%. Thus, collapsed d'context was not computed from summary data.

- When raw data was available, d'context was calculated for each participant in excel using the following formula: d'context = NORMINV(AX hit rate,0,1) - NORMINV(BX err rate,0,1). Group d'context was then computed by averaging individual participants' d'context values.
- There were a few cases where it seemed clear that researchers did not standardize values when calculating d'context (e.g., their values were less than 1.0 for both patients and controls, and corresponded to the unstandardized AX hit rate minus BX error rate). This was the case for 3 studies and it was decided not to include those d'context values in analyses.
- When calculating information from raw datasets, it was decided *a priori* to exclude participants who had 90% or more errors on any trial type and delete their data for all measures. There is no general consensus in the literature regarding exclusion criteria for low performance but a number of studies used the 90% cutoff.
- For articles that examined healthy controls and their healthy relatives, these two groups were combined into a single healthy control group, as long as diagnostic inclusion/exclusion criteria were similar.
- 5 studies included both patients with schizophrenia and their relatives (whereas 2 other relative studies included relatives of non-participant patients), which presented a possible issue due to non-independence of data. When including data from non-

independent/matched groups in the same analysis, it is appropriate to account for the correlation between the two groups or to estimate this correlation from another study. One study (MacDonald, 2002) reported correlations between schizophrenia patients' performance and their relatives' performance on AY trials, BX trials, and d'context. The performance correlations were low, ranging -0.18 to 0.04. Nonetheless, effect sizes were compared between the groups when accounting for the correlation (using formula from Borenstein, 2009) versus not accounting for the correlation. The adjusted and non-adjusted effect sizes were almost identical, correlating at r=0.999992; therefore, it was decided *a priori* not to adjust for non-independence between patients with schizophrenia and their relatives in analyses.

- When analyzing symptom correlations, data from both AX-CPT and DPX measures was used. For two studies that gave both tasks, the weighted average of the AX-CPT symptom-task correlations and DPX symptom-task correlations was taken (keeping the smaller N to be conservative). For primary, subgroup, moderation, and meta-regression analyses, data from each task was included separately but the sample sizes were halved for the two studies including both tasks.

APPENDIX C

SYSTEMATIC CODING SHEET

A) Coding Characteristics		**Response**
A1	Coder 1 = C.C., 2 = L.C.	
A2	Date coded (mm/dd/yyyy)	
A3	Finished coding? (put "yes" when finished)	
A4	Comments/concerns about coding this study	
B) Study Characteristics		**Response**
B1	Report ID number	
B2	First author's last name	
B3	Second author's last name	
B4	Third author's last name	
B5	Year of publication	

B6	Study number (if multiple studies within article)
B7	Type of report 1 = Journal article 2 = Book 3 = Book chapter 3 = Dissertation 4 = Master's thesis 5 = Private report 6 = Conference paper 7 = Other (specify)
B8	Published document 1 = Published 2 = Unpublished
B9	Peer reviewed document 1 = Peer-reviewed

2 = Not peer-reviewed

C) Location Characteristics		Response	Article pg #	Citation if authors referred to different article for info
C1	Country where study conducted 1 = United States 2 = Not in United States (specify)			
C2	Type of setting where study conducted 1 = Hospital 2 = University 3 = Academic medical center 4 = Community clinic 5 = Veterans Affairs 6 = Private setting 7 = Other (specify)			

D) Qualitative Characteristics of Schizotypy-Spectrum Group #1		Response	Article pg #	Citation if authors referred to different article for info
D1	Type of schizotypy-spectrum group included 1 = Schizophrenia 2 = Schizoaffective 3 = Schizophreniform 4 = Schizotypal Personality Disorder 5 = Clinical high-risk/Ultra high-risk 6 = Unaffected relatives of schz-spectrum group 7 = Psychometrically identified schizotypy 8 = Non-schizophrenia psychotic patients 9 = Other (specify)			
D2	Determination of diagnostic status 1 = Record review 2 = Structured interview (specify which)			

	3 = Unstructured interview 4 = Questionnaire (specify which) 5 = Did not specify
D3	Patient status 1 = Inpatient 2 = Outpatient 3 = Not patients
D4	Status of antipsychotic/neuroleptic medication 1 = Medicated with antipsychotics/neuroleptics 2 = Unmedicated/Not medicated with antipsychotics/neuroleptics
D5	Medication specifics 1 = Medication details not specified 2 = Never medicated/Medication-naïve 3 = Newly medicated

	4 = Chronically medicated 5 = Currently unmedicated, but medicated in past
D5.5	Type of antipsychotics 1 = First generation 2 = Second generation
D6	Duration of illness 1 = First-episode 2 = Chronic
D7	Other schizotypy-spectrum group inclusion/ exclusion criteria (cut and paste from article)

E) Qualitative Characteristics of Schizotypy-Spectrum Group #2	**Response**	**Article pg #**	**Citation if authors referred to different article for info**
[same variables as section D above]			

F) Qualitative Comparison Group Characteristics		Response	Article pg #	Citation if authors referred to different article for info
F1	Comparison groups included 1 = Healthy controls 2 = Non-schizophrenia psychotic patients 3 = Non-psychotic depressed patients 4 = Bipolar patients 5 = Healthy controls' relatives 6 = Other (specify)			
F2	Comparison group diagnostic criteria 1 = No history of psychosis 2 = No current psychosis 3 = No history of mental illness 4 = No current mental illness			

F3	Determination of diagnostic status 1 = Record review 2 = Structured interview (specify which) 3 = Unstructured interview 4 = Questionnaire (specify which)	
F4	Other comparison group inclusion/exclusion criteria (cut and paste from article)	

G) AX-CPT Task Parameters		**Response**	**Article pg #**	**Citation if authors referred to different article for info**
G1	Total number of trials			
G2	Number of trial blocks			
G3	% AX trials			
G4	% AY trials			

G5	% BX trials
G6	% BY trials
G7	Reaction time (RT) measurement 1 = Correct responses only 2 = All responses
G8	Duration of cue (ms)
G9	Duration of probe (ms)
G10	Delay conditions used? 1 = No 2 = Yes
G11	Delay conditions in separate trial blocks? 1 = No 2 = Yes
G12	Duration of short delay cue-probe interval (ms)

G13	Duration of short delay inter-trial interval (ms)
G14	Duration of long delay cue-probe interval (ms)
G15	Duration of long delay inter-trial interval (ms)
G16	Duration of cue-probe interval, if no delay conditions used (ms)
G17	Duration of inter-trial interval, if no delay conditions used (ms)
G18	Duration of response time allowed for cue (ms)
G19	Duration of response time allowed for probe (ms)
G20	Corrective feedback provided? 1 = No 2 = Yes
G21	Other parameters (specify)

H) DPX Task Parameters	Response	Article pg #	Citation if authors referred to different article for info
[same variables as section G above]			

I) Symptom Measures (Authors' Conceptualization)		Response	Article pg #	Citation if authors referred to different article for info
I1	Positive symptom measures			
I2	Positive symptom subscales			
I3	Negative symptom measures			
I4	Negative symptom subscales			
I5	Disorganized symptom measures			
I6	Disorganized symptom subscales			
I7	Other symptom measures: Specify			
I8	Other symptom subscales: Specify			

I9 Dimensions not specified

NOTE: Use these symptom measures for I1, I3, I5, I7

1 = Brief Psychiatric Rating Scale, 16-item scale

2 = Brief Psychiatric Rating Scale, 18-item scale

3 = Brief Psychiatric Rating Scale, 24-item scale

4 = Positive and Negative Syndrome Scale

5 = Scale for the Assessment of Positive Symptoms

6 = Scale for the Assessment of Negative Symptoms

7 = Scale of Prodromal Symptoms

8 = Schizotypal Personality Questionnaire

9 = DSM-5: Schizotypal Personality Disorder

10 = Schneiderian First Rank Symptoms

11 = Huber's pure defect Basic Symptoms

12 = Clinical Language Disorder Rating Scale

13 = Communication Disturbance Index

14 = Other (specify)

NOTE: Use these symptom subscales for I2, I4, I6, I8

1 = Total Score

2 = BPRS Hallucinations

3 = BPRS Unusual Thought Content

4 = BPRS Conceptual Disorganization

5 = BPRS Grandiosity

6 = BPRS Suspiciousness

7 = BPRS Emotional Withdrawal

8 = BPRS Motor Retardation

9 = BPRS Blunted Affect

10 = BPRS Self-Neglect

11 = BPRS Mannerisms and Posturing

12 = BPRS Disorientation

13 = BPRS Bizarre Behavior

14 = PANSS Positive scale

15 = PANSS Negative Scale

16 = SAPS Hallucinations

17 = SAPS Delusions

18 = SAPS Bizarre Behavior

19 = SAPS Positive Formal Thought Disorder

20 = SANS Anhedonia

21 = SANS Avolition

22 = SANS Alogia

23 = SANS Affective Flattening

24 = SANS Attention

25 = SOPS Unusual Thought Content

26 = SOPS Suspiciousness

27 = SOPS Perceptual Disturbances

28 = SOPS Grandiosity

29 = SOPS Social Anhedonia

30 = SOPS Avolition

31 = SOPS Decreased Expression of Emotion

32 = SOPS Decreased Experience of Emotions and Self

33 = SOPS Deterioration in Role Functioning

34 = SOPS Decreased Ideational Richness

35 = SOPS Conceptual Disorganization

36 = SOPS Odd Behavior and Appearance

37 = SOPS Bizarre Thinking

38 = SOPS Trouble with Focus and Attention

39 = SOPS Personal Hygiene

40 = SPQ Odd Beliefs/Magical Thinking

41= SPQ Unusual Perceptual Experiences

42 = SPQ Suspiciousness

43= SPQ Ideas of Reference

44 = SPQ Constricted Affect

45 = SPQ Excessive Social Anxiety

46 = SPQ No Close Friends

47 = SPQ Odd Speech

48 = SPQ Odd Behavior

49 = DSM-5 SPD criterion 1

50 = DSM-5 SPD criterion 2

51 = DSM-5 SPD criterion 3

52 = DSM-5 SPD criterion 4

53 = DSM-5 SPD criterion 5

54 = DSM-5 SPD criterion 6

55 = DSM-5 SPD criterion 7

56 = DSM-5 SPD criterion 8

57 = DSM-5 SPD criterion 9

58 = CDI Disorganized Speech

59 = Other (specify)

60 = Did not specify subscales

J) Quantitative Characteristics of Schizotypy-Spectrum Group #1		**Response**	**Article pg #**	**Citation if authors referred to different article for info**
J1	Sample size (n=__)			
J2	Diagnostic composition: %			
J3	Positive symptoms: M (SD)			
J4	Negative symptoms: M (SD)			
J5	Disorganized symptoms: M (SD)			
J6	Other symptoms (specify Sx type): M (SD)			
J7	Other symptoms (specify Sx type): M (SD)			

J8	Length of illness in years: M (SD)
J9	% Unmedicated
J10	% Medicated with antipsychotic/neuroleptic
J10.5a	% Medicated: first-generation antipsychotic
J10.5b	% Medicated: second-generation antipsychotic
J11	Dose of antipsychotic/neuroleptic (mg): M (SD)
J12	% Taking other meds (if reported, specify type)
J13	Age: M (SD)
J14	Sex: % Female
J15	Race/Ethnicity: % White
J16	Race/Ethnicity: % Black
J17	Race/Ethnicity: % Latin(a/o)
J18	Race/Ethnicity: % Asian
J19	Race/Ethnicity: % Other

J20	Years of education: M (SD)
J21	Socioeconomic status (SES) or Years of father's education: M (SD)
J22	IQ measure #1: Specify + M (SD)
J23	IQ measure #2: Specify + M (SD)
J24	Community/occupational functioning measure #1: Specify + M (SD)
J25	Community/occupational functioning measure #2: Specify + M (SD)
J26	Number of subjects dropped: Specify reason(s)
J27	AX Errors (no delay conditions used): M (SD)
J28	AY Errors (no delay conditions used): M (SD)
J29	BX Errors (no delay conditions used): M (SD)
J30	BY Errors (no delay conditions used): M (SD)

J31	Total Errors (no delay conditions used): M (SD)
J32	AX % Correct (no delay conditions used): M (SD)
J33	AY % Correct (no delay conditions used): M (SD)
J34	BX % Correct (no delay conditions used): M (SD)
J35	BY % Correct (no delay conditions used): M (SD)
J36	Total % Correct (no delay conditions used): M (SD)
J37	AX Reaction time (ms) (no delay conditions used): M (SD)

J38	AY Reaction time (ms) (no delay conditions used): M (SD)
J39	BX Reaction time (ms) (no delay conditions used): M (SD)
J40	BY Reaction time (ms) (no delay conditions used): M (SD)
J41	Total Reaction time (ms) (no delay conditions used): M (SD)
J42	d'context (no delay conditions used): M (SD)
J43	a'context (no delay conditions used): M (SD)
J44	AX Errors (short delay): M (SD)
J45	AY Errors (short delay): M (SD)
J46	BX Errors (short delay): M (SD)
J47	BY Errors (short delay): M (SD)

J48	Total Errors (short delay): M (SD)
J49	AX % Correct (short delay): M (SD)
J50	AY % Correct (short delay): M (SD)
J51	BX % Correct (short delay): M (SD)
J52	BY % Correct (short delay): M (SD)
J53	Total % Correct (short delay): M (SD)
J54	AX Reaction time (ms) (short delay): M (SD)
J55	AY Reaction time (ms) (short delay): M (SD)
J56	BX Reaction time (ms) (short delay): M (SD)
J57	BY Reaction time (ms) (short delay): M (SD)
J58	Total Reaction time (ms) (short delay): M (SD)
J59	d'context (short delay): M (SD)
J60	a'context (short delay): M (SD)
J61	AX Errors (long delay): M (SD)

J62	AY Errors (long delay): M (SD)
J63	BX Errors (long delay): M (SD)
J64	BY Errors (long delay): M (SD)
J65	Total Errors (long delay): M (SD)
J66	AX % Correct (long delay): M (SD)
J67	AY % Correct (long delay): M (SD)
J68	BX % Correct (long delay): M (SD)
J69	BY % Correct (long delay): M (SD)
J70	Total % Correct (long delay): M (SD)
J71	AX Reaction time (ms) (long delay): M (SD)
J72	AY Reaction time (ms) (long delay): M (SD)
J73	BX Reaction time (ms) (long delay): M (SD)
J74	BY Reaction time (ms) (long delay): M (SD)
J75	Total Reaction time (ms) (long delay): M (SD)

J76	d'context (long delay): M (SD)
J77	a'context (long delay): M (SD)
J78	Other (specify): M (SD)

K) Quantitative Characteristics of Schizotypy-Spectrum Group #2	**Response**	**Article pg #**	**Citation if authors referred to different article for info**
[same variables as Section J above]			

L) Quantitative Comparison Group Characteristics	**Response**	**Article pg #**	**Citation if authors referred to different article for info**
[same variables as Section J above]			

M) RESULTS: Between-Group Comparisons #1	**Groups, direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Signifi-cance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**
M1	Positive symptoms							

M2	Negative symptoms
M3	Disorganized symptoms
M4	Other symptoms #1
M5	Other symptoms #2
M6	Age
M7	Sex
M8	Race/Ethnicity
M9	Years of education
M10	SES/Father's education
M11	IQ measure #1
M12	IQ measure #2
M13	Community/ occupational #1
M14	Community/

	occupational #2
M15	# subjects dropped
M16	AX Errors (no delay conditions)
M17	AY Errors (no delay conditions)
M18	BX Errors (no delay conditions)
M19	BY Errors (no delay conditions)
M20	Total Errors (no delay conditions)
M21	AX % Correct

	(no delay conditions)
M22	AY % Correct (no delay conditions)
M23	BX % Correct (no delay conditions)
M24	BY % Correct (no delay conditions)
M25	Total % Correct (no delay conditions)
M26	AX RT (no delay conditions)
M27	AY RT (no delay conditions)
M28	BX RT

	(no delay conditions)
M29	BY RT (no delay conditions)
M30	Total RT (no delay conditions)
M31	d'context (no delay conditions)
M32	a'context (no delay conditions)
M33	AX Errors (short delay)
M34	AY Errors (short delay)
M35	BX Errors (short delay)
M36	BY Errors (short delay)
M37	Total Errors

	(short delay)
M38	AX % Correct (short delay)
M39	AY % Correct (short delay)
M40	BX % Correct (short delay)
M41	BY % Correct (short delay)
M42	Total % Correct (short delay)
M43	AX RT (short delay)
M44	AY RT (short delay)
M45	BX RT (short delay)

M46	BY RT (short delay)
M47	Total RT (short delay)
M48	d'context (short delay)
M49	a'context (short delay)
M50	AX Errors (long delay)
M51	AY Errors (long delay)
M52	BX Errors (long delay)
M53	BY Errors (long delay)
M54	Total Errors (long delay)
M55	AX % Correct (long delay)
M56	AY % Correct (long delay)
M57	BX % Correct

	(long delay)
M58	BY % Correct (long delay)
M59	Total % Correct (long delay)
M60	AX RT (long delay)
M61	AY RT (long delay)
M62	BX RT (long delay)
M63	BY RT (long delay)
M64	Total RT (long delay)
M65	d'context (long delay)
M66	a'context (long delay)
M67	ANOVA: Main effect of

	Trial type, **errors**
M68	ANOVA: Main effect of Group, **errors**
M69	ANOVA: Group x Trial Type interaction, **errors**
M70	ANOVA: Main effect of Trial type, **RT**
M71	ANOVA: Main effect of Group, **RT**
M72	ANOVA: Group x Trial Type interaction, **RT**
M73	ANOVA: Main effect of Trial type, **d'context**
M74	ANOVA: Main effect of

	Group, **d'context**
	ANOVA: Group x Trial
M75	Type interaction,
	D'**context**
	ANOVA: $BX_{\mathbf{errors}}$ -
M76	
	$AY_{\mathbf{errors}}$
M77	ANOVA: $BX_{\mathbf{RT}}$ - $AY_{\mathbf{RT}}$
	ANOVA: Main effect of
M78	
	Delay, **errors**
	ANOVA: Main effect of
M79	
	Delay, **RT**
	ANOVA: Main effect of
M80	
	Delay, **d'context**
M81	ANOVA: Group x

	Delay interaction, **errors**
M82	ANOVA: Group x Delay interaction, **RT**
M83	ANOVA: Group x Delay interaction, **d'context**
M84	ANOVA: Trial Type x Delay interaction, **errors**
M85	ANOVA: Trial Type x Delay interaction, **RT**
M86	ANOVA: Trial Type x Delay interaction, **d'context**

	ANOVA: Group x Trial							
M87	Type x Delay							
	interaction, **errors**							
	ANOVA: Group x Trial							
M88	Type x Delay, **RT**							
	ANOVA: Group x Trial							
M89	Type x Delay							
	interaction, **d'context**							
M90	Other (specify)							
N) RESULTS: Between-Group Comparisons #2		**Groups and direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Signifi-cance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**
[same variables as Section M above]									

N) RESULTS: Between-Group Comparisons #2	**Groups and direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Signifi-cance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**

[same variables as Section M above]

P) RESULTS: Symptom and Function Associations in all Schizotypy-Spectrum Groups		***Bivariate Correlations (r)***	***Significance (p)***	**Article pg #**
P1	Positive symptoms with d'context			
P2	Positive symptoms with BX errors			
P3	Positive symptoms with BX-AY errors			
P4	Positive symptoms with BX reaction time			
P5	Negative symptoms with d'context			
P6	Negative symptoms with BX errors			
P7	Negative symptoms with BX-AY errors			

P8	Negative symptoms with BX reaction time
P9	Disorganized symptoms with d'context
P10	Disorganized symptoms with BX errors
P11	Disorganized symptoms with BX-AY errors
P12	Disorganized symptoms with BX reaction time
P13	IQ with d'context
P14	IQ with BX errors
P15	IQ with BX-AY errors
P16	IQ with BX reaction time
P17	Community/occupational functioning with d'context
P18	Community/occupational functioning with BX errors
P19	Community/occupational functioning with BX-AY errors
P20	Community/occupational functioning with BX reaction time
P21	Other symptoms (specify type) with d'context

P22	Other symptoms (specify type) with BX errors
P23	Other symptoms (specify type) with BX-AY errors
P24	Other symptoms (specify type) with BX reaction time

Q) RESULTS: Symptom and Function Associations in Schizotypy-Spectrum Group #1	***Bivariate Correlations (r)***	***Significance (p)***	**Article pg #**
[same variables as section P above]			

R) RESULTS: Symptom and Function Associations in Schizotypy-Spectrum Group #2	***Bivariate Correlations (r)***	***Significance (p)***	**Article pg #**
[same variables as section P above]			

S) RESULTS: Symptom and Function Associations in Comparison Group	***Bivariate Correlations (r)***	***Significance (p)***	**Article pg #**
[same variables as section P above]			

T) RESULTS: Symptom and Function Associations in All Participants	***Bivariate Correlations (r)***	***Significance (p)***	**Article pg #**
[same variables as section P above]			

U) Follow-Up characteristics		**Response**	**Article pg #**	**Citation if authors referred to different article for info**
U1	Follow-up conducted? 1 = No (if No, stop here: ignore U2&U3, ignore sections V-X) 2 = Yes			
U2	Duration of follow-up			
U3	List changes from baseline to follow-up (e.g., sample size)			

V) Follow-Up: Between-Group Comparisons #1	**Groups and direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Significance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**
[same variables as section M above]								
W) Follow-Up: Between-Group Comparisons #2	**Groups and direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Significance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**
[same variables as section M above]								
X) Follow-Up: Between-Group Comparisons #3	**Groups and direction of findings**	**Analysis** (1 = *t*-test, 2 = ANOVA, 3 = Planned contrasts, 4 = Other	***F***	***t***	**df**	***Significance (p)***	**Measure of Effect Size** (1 = Cohen's *d*, 2 = η^2, 3 = other	**Effect Size**
[same variables as section M above]								

APPENDIX D

OUR CONCEPTUALIZATIONS OF SYMPTOM DIMENSIONS

Positive Symptoms	**Brief Psychiatric Rating Scale:** hallucinations, unusual thought content, grandiosity, suspiciousness **Scale for the Assessment of Positive Symptoms:** hallucinations, delusions **Scale of Prodromal Symptoms, Perceptual Aberration Scale, Magical Ideation Scale:** unusual thought content, suspiciousness, grandiosity, perceptual abnormalities **Positive and Negative. Syndrome Scale:** delusions, hallucinations, grandiosity, suspiciousness/persecution. Based on Liddle (1987) reality distortion**:** delusions, hallucinations. **DSM-5 Schizotypal Personality Disorder criteria:** ideas of reference, odd beliefs/magical thinking, unusual perceptual experiences, suspiciousness/paranoid ideation **Schneiderian First Rank Symptoms:** delusional perception, audible thoughts, voices arguing/discussing, voices commenting on one's actions, thought withdrawal, thought insertion, thought broadcast, passivity of affect, passivity of impulse, passivity of volition, somatic passivity **Schizotypal Personality Questionnaire:** unusual perceptual experiences, referential, paranoia, odd beliefs

Negative Symptoms	**Brief Psychiatric Rating Scale:** emotional withdrawal, motor retardation, blunted affect, self-neglect **Scale for the Assessment of Negative Symptoms:** anhedonia, avolition, alogia, affective flattening, attention **Scale of Prodromal Symptoms, Physical Anhedonia Scale, Social Anhedonia Scale:** social isolation, avolition, decreased expression of emotion, decreased experience of self, decreased ideational richness, deterioration in role functioning **Positive and Negative. Syndrome Scale:** blunted affect, emotional withdrawal, poor rapport, passive/apathetic social withdrawal, lack of spontaneity and flow of
Disorganized Symptoms	**Brief Psychiatric Rating Scale:** conceptual disorganization, mannerisms and posturing, disorientation, bizarre behavior **Scale for the Assessment of Positive Symptoms:** bizarre behavior, positive formal thought disorder **Scale for the Assessment of Negative Symptoms:** attention **Positive and Negative Syndrome Scale:** conceptual disorganization, difficulty in abstract thinking, poor attention. Based on Liddle (1987) disorganisation: inappropriate affect, poverty of content of speech, disturbances in the form of thought. **Standard speech task:** number of speech

This table details the symptom conceptualizations we have categorized as appropriate for running symptom correlations with a subset of studies. These include authors' conceptualizations we agreed with and our own conceptualization when none was provide or when raw data were available for us to calculate our own symptom composites.

APPENDIX E

DEALING WITH MULTIPLE OUTCOMES IN *P*-CURVE ANALYSIS

For studies that gave both the AX-CPT and the DPX, only *p*-values from the AX-CPT were used, which was selected because it is used more frequently than the DPX. For studies that gave multiple measurements of a single symptom dimension or reported on item-level correlations rather than overall scales, the measurement or item was chosen that was deemed to best represent that symptom dimension or to be most strongly related to CI ability in theory. Specifically, one study reported correlations with both the Scale for the Assessment of Negative Symptoms (SANS) and word count on a speech task. The SANS scale was selected because it is more commonly used in the literature to assess negative symptoms. For a different study that reported correlations separately for hallucination and delusion items, the delusion item was selected to represent the positive dimension because it is cognitive–rather than perceptual—in nature and the CI is a cognitive task. The same study reported on the following negative symptom items: affective flattening, alogia, avolition, and anhedonia. The alogia item was selected because it was deemed to represent the cognitive—rather than affective—aspects of the negative symptom dimension. Finally, that same study reported on the following disorganization symptom items: bizarre behavior, formal thought disorder, and attention. The formal thought disorder item was selected because it is considered most specific to disorganization and is a hallmark of disorganized cognition.

APPENDIX F

LIST OF INCLUDED STUDIES WITH BASIC SAMPLE AND TASK INFORMATION

Study	Sample Size and Type	Task
Barch, D., Carter, C., Braver, T., Sabb, F., MacDonald, A., Noll, D., & Cohen, J. (2001). Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Archives of General Psychiatry, 58,* 280-288.	14 first-episode schizophrenia 12 healthy control	AX-CPT
Barch, D., Carter, C., MacDonald, A., Braver, T., & Cohen, J. (2003). Context-processing deficits in schizophrenia: diagnostic specificity, 4-week course, and relationships to clinical symptoms. *Journal of Abnormal Psychology, 112,* 132-43.	49 first-episode schizophrenia 30 non-schizophrenia psychotic control 72 healthy control	AX-CPT
Barch, D., Mitropoulou, V., Harvey, P., New, A., Silverman, J., & Siever, L. (2004). Context-Processing Deficits in Schizotypal Personality	26 schizotypal personality disorder	AX-CPT

Disorder. *Journal of Abnormal Psychology*, *113*, 556-568.	35 healthy control	
Barch, D., Yodkovik, N., Sypher-Locke, H., & Hanewinkel, M. (2008). Intrinsic motivation in schizophrenia: relationships to cognitive function, depression, anxiety, and personality. *Journal of Abnormal Psychology*, *117*, 776–87.	57 chronic schizophrenia 37 healthy control	AX-CPT
Becker, T.. (2012). Understanding the relationship between goal maintenance and disorganized speech. *Unpublished Dissertation.*	49 chronic schizophrenia 28 healthy control	AX-CPT
Braver T., Barch D., Cohen J. (1999): Mechanisms of cognitive control: Active memory, inhibition, and the pre-frontal cortex (Technical Report PDP.CNS.99.1). Pittsburgh, PA: Carnegie Mellon University.	16 first-episode schizophrenia 16 healthy control	AX-CPT
Buchanan, R., Keefe, R., Lieberman, J., Barch, D., Csernansky, J., Goff, D., … & Marder, S. (2011). A randomized clinical trial of MK-0777 for the treatment of cognitive impairments in people with schizophrenia. *Biological Psychiatry*, *69*, 442–449.	47 chronic schizophrenia	AX-CPT

Ceccherini-Nelli, A., Turpin-Crowther, K., & Crow, T. (2007). Schneider's first rank symptoms and continuous performance disturbance as indices of dysconnectivity of left- and right-hemispheric components of language in schizophrenia. *Schizophrenia Research, 90,* 203-213.	17 chronic schizophrenia 14 non-schizophrenia psychotic control 15 depressed control 16 healthy control	AX-CPT
Chung, Y., Mathews, J., & Barch, D. (2011). The effect of context processing on different aspects of social cognition in schizophrenia. *Schizophrenia Bulletin, 37,* 1048-56.	41 chronic schizophrenia 27 healthy control	AX-CPT
Cohen, J., Barch, D., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology, 108,* 120–133.	53 chronic schizophrenia 25 depressed control 31 healthy control	AX-CPT

Delawalla, Z., Barch, D.M., Eastep, J., Thomason, E., Hanewinkel, M., Thompson, P., & Csernansky, J. (2006). Factors mediating cognitive deficits and psychopathology among siblings of individuals with schizophrenia. *Schizophrenia Bulletin, 32,* 525–537	27 chronic schizophrenia 31 relatives of schizophrenia 81 healthy control	AX-CPT
Delawalla, Z., Csernansky, J., & Barch, D. (2008). Prefrontal Cortex Function in Nonpsychotic Siblings of Individuals with Schizophrenia. *Biological Psychiatry, 63,* 490-497	30 relatives of schizophrenia 92 healthy control	AX-CPT
Dias, E., Bickel, S., Epstein, M., Sehatpour, P., & Javitt, D. (2013). Abnormal task modulation of oscillatory neural activity in schizophrenia. *Frontiers in Psychology, 4.*	17 chronic schizophrenia 13 healthy control	AX-CPT
Edwards, B., Barch, D., & Braver, T. (2010). Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Frontiers in Human Neuroscience, 4,* 32.	22 chronic schizophrenia 14 healthy control	AX-CPT
Fisher, M. (2016). *Personal Communication.*	34 chronic schizophrenia	AX-CPT

Fornito, A., Yoon, J., Zalesky, A., Bullmore, E., & Carter, C. (2011). General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance. *Biological Psychiatry, 70,* 64–72.	23 first-episode schizophrenia 25 healthy control	AX-CPT
Gold, J., Barch, D., Carter, C., Dakin, S., Luck, S., MacDonald, A., & ... Strauss, M. (2012). Clinical, functional, and intertask correlations of measures developed by the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia Consortium. *Schizophrenia Bulletin, 38,* 144-152. Henderson, D., Poppe, A., Barch, D., Carter, C., Gold, J., Ragland, J., & ... MacDonald, A. (2012). Optimization of a goal maintenance task for use in clinical applications. *Schizophrenia Bulletin, 38,* 104-113.	138 chronic schizophrenia 136 healthy control	DPX

Holmes, A., MacDonald, A., Carter, C., Barch, D., Andrew, S., & Cohen, J. (2005). Prefrontal functioning during context processing in schizophrenia and major depression: an event-related fMRI study. *Schizophrenia Research, 76,* 2-3.	7 chronic schizophrenia 10 depressed control 9 healthy control	AX-CPT
Jones, J, Sponheim, S., Macdonald, A. (2010, Study 3). The dot pattern expectancy task: Reliability and replication of deficits in schizophrenia. *Psychological Assessment, 22,* 131-141	47 chronic schizophrenia 48 healthy control	DPX
Lesh, T., Tanase, C., Geib, B., Niendam, T., Yoon, J., Minzenberg, M., Ragland, J., ... Carter, C. (2015). A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *Jama Psychiatry, 72*, 226-34.	23 first-episode schizophrenia (medicated) 22 first-episode schizophrenia (unmedicated) 37 healthy control	AX-CPT

Lewis, D., Cho, R., Carter, C., Eklund, K., Forster, S., Kelly, M., & Montrose, D. (2008). Subunit-Selective Modulation of GABA Type A Receptor Neurotransmission and Cognition in Schizophrenia. *American Journal of Psychiatry, 165,* 1585–1593.	15 chronic schizophrenia	AX-CPT
Lopez-Garcia, P., Cristobal-Huerta, A., Espinoza, L., Molero, P., Sanchez-Pedreño, F., & Hernández-Tamames, J. A. (2016). The influence of the COMT genotype in the underlying functional brain activity of context processing in schizophrenia and in relatives. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 71,* 176–182.	15 chronic schizophrenia spectrum patients 16 relatives of schizophrenia 20 healthy control	DPX
López-García, P., Young, L., Marín, J., Molero, P., & Ortuño, F. (2015). The impact of the Val158Met COMT polymorphism on context processing in patients on the schizophrenia spectrum and their relatives. *Schizophrenia Research: Cognition, 2,* 179–184.	40 chronic schizophrenia spectrum patients 26 relatives of schizophrenia 63 healthy control	DPX

MacDonald, A. (2002). A differential deficit in context processing associated with the genetic liability to schizophrenia: A sibling study. ProQuest Information & Learning, US.	**AX-CPT:** 24 chronic schizophrenia, 24 relatives of schizophrenia, 29 healthy control **DPX:** 17 chronic schizophrenia, 16 relatives of schizophrenia, 28 healthy control	AX-CPT, DPX
MacDonald, A., & Carter, C. (2003). Event-related FMRI study of context processing in dorsolateral prefrontal cortex of patients with schizophrenia. *Journal of Abnormal Psychology*, *112*, 689–97.	17 chronic schizophrenia 17 healthy control	AX-CPT
McClure, M., Barch, D., Flory, J., Harvey, P., & Siever, L. (2008). Context processing in schizotypal personality disorder: evidence of specificity of impairment to the schizophrenia spectrum. *Journal of Abnormal Psychology*, *117*, 342–354	63 schizotypal personality disorder 25 non-Cluster A personality disorder 42 healthy control	AX-CPT

McClure, M., Barch, D., Romero, M., Minzenberg, M., Triebwasser, J., Harvey, P., & Siever, L. (2007). The Effects of Guanfacine on Context Processing Abnormalities in Schizotypal Personality Disorder. *Biological Psychiatry, 61,* 1157–1160.	29 schizotypal personality disorder	AX-CPT
McClure, M., Harvey, P., Goodman, M., Triebwasser, J., New, A., Koenigsberg, H., … Siever, L. (2010). Pergolide treatment of cognitive deficits associated with schizotypal personality disorder: continued evidence of the importance of the dopamine system in the schizophrenia spectrum. *Neuropsychopharmacology, 35,* 1356–1362.	24 schizotypal personality disorder	AX-CPT
Merrill, A., Karcher, N., Cicero, D., Becker, T., Docherty, A., & Kerns, J. (2017). Evidence that communication impairment in schizophrenia is associated with generalized poor task performance. *Psychiatry Research, 249,* 172-179.	43 chronic schizophrenia 19 healthy control	AX-CPT

Niendam, T., Lesh, T., Yoon, J., Westphal, A., Hutchison, N., Ragland, J., & ... Carter, C. (2014). Impaired context processing as a potential marker of psychosis risk state. *Psychiatry Research: Neuroimaging, 221,* 13-20	25 ultra high-risk	AX-CPT
Paavola, J. (2013). Personality dimensions and cognitive functioning of relatives of persons diagnosed with schizophrenia and bipolar I disorder: A comparative and predictive study. *Unpublished dissertation.*	47 relatives of patients with schizophrenia 57 healthy control	DPX
Perlstein, W., Dixit, N., Carter, C., Noll, D., & Cohen, J. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry, 53*, 25-38.	16 chronic schizophrenia 15 healthy control	AX-CPT
Poppe, A., Carter, C., Minzenberg, M., & MacDonald, A. (2015). Task-based functional connectivity as an indicator of genetic liability to schizophrenia. *Schizophrenia Research, 162,* 1-3.	19 chronic schizophrenia 33 relatives of schizophrenia 50 healthy control	AX-CPT

Poppe, A., Barch, D., Carter, C., Gold, J., Ragland, J., Silverstein, S., & MacDonald, A. (2016). Reduced Frontoparietal Activity in Schizophrenia is Linked to a Specific Deficit in Goal Maintenance: a Multi-site Functional Imaging Study. *Schizophrenia Bulletin, 42*, 1149–1157.	47 chronic schizophrenia 56 healthy control	DPX
Reilly, J., Hill, S., Gold, J., Keefe, R., Clementz, B., Gershon, E., Keshavan, M., ... Sweeney, J. (2017). Impaired Context Processing is Attributable to Global Neuropsychological Impairment in Schizophrenia and Psychotic Bipolar Disorder. *Schizophrenia Bulletin, 43*, 397-406	402 chronic schizophrenia 304 bipolar with psychotic features control 210 healthy control	DPX
Richard, A., Carter, C., Cohen, J., & Cho, R. (2013). Persistence, diagnostic specificity and genetic liability for context-processing deficits in schizophrenia. *Schizophrenia Research, 147*, 75-80.	63 first-episode schizophrenia 31 relatives of patients with schizophrenia 83 healthy control 47 non-schizophrenia psychotic control	AX-CPT

Sheffield, J., Gold, J., Strauss, M., Carter, C., MacDonald, A., Ragland, J., Silverstein, S., ... Barch, D. (2014). Common and specific cognitive deficits in schizophrenia: relationships to function. *Cognitive, Affective & Behavioral Neuroscience, 14,* 161-74.	104 chronic schizophrenia 132 healthy control	AX-CPT, DPX
Sheffield, J., Repovs, G., Harms, M., Carter, C., Gold, J., MacDonald, A., … Barch, D. (2015). Fronto-parietal and cingulo-opercular network integrity and cognition in health and schizophrenia. *Neuropsychologia, 73*, 82–93	46 chronic schizophrenia 54 healthy control	DPX
Sloat, V. (2007). Context processing in psychometrically defined schizotypes. *Unpublished dissertation.*	25 psychometrically-identified schizotypy 18 psychometrically-identified vulnerable to depression 38 healthy control	AX-CPT

Stratta, P., Daneluzzo, E., Bustini, M., Prosperini, P., & Rossi, A. (2000). Processing of context information in schizophrenia: relation to clinical symptoms and WCST performance. *Schizophrenia Research, 44,* 57–67.	20 chronic schizophrenia 20 healthy control	AX-CPT
Thoma, P., & Daum, I. (2008). Working memory and multi-tasking in paranoid schizophrenia with and without comorbid substance use disorder. *Addiction, 103,* 774–786.	23 chronic schizophrenia 22 depressed control 20 healthy control	AX-CPT
Thoma, P., Zoppelt, D., Wiebel, B., & Daum, I. (2007). Context processing and negative symptoms in schizophrenia. *Journal of Clinical and Experimental Neuropsychology, 29,* 428-435	26 chronic schizophrenia 13 healthy control	AX-CPT
Todd, J., Whitson, L., Smith, E., Michie, P., Schall, U., & Ward, P. (2014). What's intact and what's not within the mismatch negativity system in schizophrenia. *Psychophysiology, 51,* 337–347.	33 chronic schizophrenia 58 healthy control	AX-CPT

Woodward, N., et al. (2016). *Personal Communication*.	15 first-episode schizophrenia 35 chronic schizophrenia 19 bipolar with psychotic features control 39 healthy control	AX-CPT
Yoon, J., Nguyen, D., McVay, L., Deramo, P., Minzenberg, M., Ragland, J., Niendam, T., ... Carter, C. (2012). Automated classification of fMRI during cognitive control identifies more severely disorganized subjects with schizophrenia. *Schizophrenia Research, 135,* 1-3.	51 first-episode schizophrenia 51 healthy control	AX-CPT
Yoon, J., Westphal, A., Minzenberg, M., Niendam, T., Ragland, J., Lesh, T., … Carter, C. (2014). Task-evoked substantia nigra hyperactivity associated with prefrontal hypofunction, prefrontonigral disconnectivity and nigrostriatal connectivity predicting psychosis severity in medication naïve first episode schizophrenia. *Schizophrenia Research, 159,* 521–526.	12 first-episode schizophrenia 15 healthy control	AX-CPT

Zhang, Z., Chen, X., Yu, P., Zhang, Q., Sun, X., Gu, H., … Ji, F. (2015). Evidence for the contribution of NOS1 gene polymorphism (rs3782206) to prefrontal function in schizophrenia patients and healthy controls. *Neuropsychopharmacology, 40,* 2015–2016.	339 chronic schizophrenia 665 healthy control	DPX

APPENDIX G

INFORMATION AND METHODOLOGICAL DECISIONS FOR SPECIFIC STUDIES

Study	Additional Information and Methodological Decisions
Barch, D., Carter, C., Braver, T., Sabb, F., MacDonald, A., Noll, D., & Cohen, J. (2001). Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Archives of General Psychiatry, 58,* 280-288.	Accuracy was normalized using an arcsine transformation.
Barch, D., Carter, C., MacDonald, A., Braver, T., & Cohen, J. (2003). Context-processing deficits in schizophrenia: diagnostic specificity, 4-week course, and relationships to clinical symptoms. *Journal of Abnormal Psychology, 112,* 132-43.	For the signal detection measures, a correction factor was applied in cases of a perfect hit rate or false alarm rate. This correction factor was (hit rate = 2E-(1/N); false alarm = 1 - {2E-(1/N)}; N = number of target or non-target trials

Barch, D., Mitropoulou, V., Harvey, P., New, A., Silverman, J., & Siever, L. (2004). Context-Processing Deficits in Schizotypal Personality Disorder. *Journal of Abnormal Psychology, 113,* 556-568.	For reaction time data, median responses were used to reduce the influence of outlier responses
Becker, T. (2012). Understanding the relationship between goal maintenance and disorganized speech. *Unpublished Dissertation.*	The author reported slightly different cue-probe delays in figures (6100ms) versus text (6000ms). I coded delay as 6000ms.
Braver T., Barch D., Cohen J. (1999): Mechanisms of cognitive control: Active memory, inhibition, and the pre-frontal cortex. Pittsburgh, PA: Carnegie Mellon University.	For measures involving accuracy, results confirmed after using an arcsine transformation. For measures involving reaction time, analyses confirmed using an inverse transformation.
Buchanan, R., Keefe, R., Lieberman, J., Barch, D., Csernansky, J., Goff, D., … & Marder, S. (2011). A randomized clinical trial of MK-0777 for the treatment of cognitive impairments in people with schizophrenia. *Biological Psychiatry*, *69*, 442–449.	The three treatment groups at baseline were combined into a single schizophrenia group to compute outcome data.

Cohen, J., Barch, D., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, *108,* 120–133.	Determination of diagnostic status was unclear for depressed controls. Participants who did not make false alarms in a particular condition (e.g., BX, AY, BY) had no reaction time data for that condition. For dependent measures involving accuracy or response proportions, raw data were normalized using an arcsine transformation. For measures involving reaction time, authors confirmed traditional analyses with an inverse transformation of the data. Significant effects conforming to *a priori* predictions were evaluated with planned orthogonal contrasts adjusted for unequal sample sizes. Other significant effects were evaluated with post hoc contrasts using Tukey's HSD for repeated measures, also adjusted for unequal sample sizes.
Fisher, M. (2016). *Personal Communication.*	Group summary data was computed from raw data shared by Dr. Fisher.

Jones, J, Sponheim, S., Macdonald, A. (2010, Study 3). The dot pattern expectancy task: Reliability and replication of deficits in schizophrenia. *Psychological Assessment, 22,* 131-141	Correlations were reported as Spearman's rho and were thus excluded. For performance accuracy, analyses were conducted on arcsine transformations of square roots of error rates. Demographic data were available for 87 of 95 participants; it was not noted which groups were missing demographic data.
Lesh, T., Tanase, C., Geib, B., Niendam, T., Yoon, J., Minzenberg, M., Ragland, J., ... Carter, C. (2015). A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *Jama Psychiatry, 72,* 226-34.	Arcsine transformed error rates were computed for statistical tests to compensate for data non-normality.
López-García, P., Young, L., Marín, J., Molero, P., & Ortuño, F. (2015). The impact of the Val158Met COMT polymorphism on context processing in patients on the schizophrenia spectrum and their relatives. *Schizophrenia Research: Cognition, 2,* 179-184.	Genotype groups were combined into regular diagnostic groups to compute outcome data.

Macdonald, A. (2002). A differential deficit in context processing associated with the genetic liability to schizophrenia: A sibling study. ProQuest Information & Learning, US.	3 relatives of patients with schizophrenia had Cluster A PDs; this group was still included in meta-analysis because authors showed that sibling effects were similar when re-analyzed without PD siblings. The thesis reported different information for d'context and sex compared to corresponding published articles. The author was contacted to clarify but no response was received; thus, values from the thesis manuscript were used. In cases of a perfect hit or false alarm rate, a small constant (0.999 or 0.001) was used to calculate d'context. Arcsine transformations were calculated to reduce error skew.
MacDonald, A., & Carter, C. (2003). Event-related FMRI study of context processing in dorsolateral prefrontal cortex of patients with schizophrenia. *Journal of Abnormal Psychology, 112*, 689–97.	Authors reported that having only seven trials of each trial type decreased reliability of errors (AY errors: alpha = .25; BX errors: alpha = .67). They also noted that level of symptoms during the past week in patients was low relative to the overall population of patients with schizophrenia.

McClure, M., Barch, D., Romero, M., Minzenberg, M., Triebwasser, J., Harvey, P., & Siever, L. (2007). The Effects of Guanfacine on Context Processing Abnormalities in Schizotypal Personality Disorder. *Biological Psychiatry, 61,* 1157–1160.	Both treatment groups at baseline were combined into a single schizotypal personality disorder group to compute outcome data.
McClure, M., Harvey, P., Goodman, M., Triebwasser, J., New, A., Koenigsberg, H., … Siever, L. (2010). Pergolide treatment of cognitive deficits associated with schizotypal personality disorder: continued evidence of the importance of the dopamine system in the schizophrenia spectrum. *Neuropsychopharmacology, 35,* 1356–1362.	The authors provided their own task timing parameters but cited Barch et al. (2004) for "full description" of the task; thus, it was assumed that the trial proportions were the same between these two tasks.
Merrill, A., Karcher, N., Cicero, D., Becker, T., Docherty, A., & Kerns, J. (2017). Evidence that communication impairment in schizophrenia is associated with generalized poor task performance. *Psychiatry Research, 249,* 172-179.	Correlations were reported as Spearman's rho; therefore, these correlations were not included.

Niendam, T., Lesh, T., Yoon, J., Westphal, A., Hutchison, N., Ragland, J., & ... Carter, C. (2014). Impaired context processing as a potential marker of psychosis risk state. *Psychiatry Research: Neuroimaging, 221,* 13-20	15 first-episode schizophrenia patients and 9 healthy control participants were included in a previous publication by Yoon et al. (2008). Because this previous study was already included in the meta-analysis, data for the schizophrenia patients and healthy control group are not included; therefore, data from ultra high-risk participants was used for point estimates only.
Paavola, J. (2013). Personality dimensions and cognitive functioning of relatives of persons diagnosed with schizophrenia and bipolar I disorder: A comparative and predictive study. *Unpublished dissertation.*	The authors did not screen for psychiatric disorders in the bipolar relatives group so this group was excluded from meta-analyses *a priori.* Participants could indicate more than one racial identity. Therefore, frequencies per each racial category do not sum to 100%. Authors conducted subset analyses but only original analyses from the author's full dataset was included in meta-analyses.

Perlstein, W., Dixit, N., Carter, C., Noll, D., & Cohen, J. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry, 53*, 25-38.	Median reaction time data and mean error rates for the AX-CPT task were analyzed using separate analyses of variance for ACue and BCue types, with group as the between-subjects factor and probe type (X, Y) and delay (short, long) as the within-subject factors.
Poppe, A., Barch, D., Carter, C., Gold, J., Ragland, J., Silverstein, S., & MacDonald, A. (2016). Reduced Frontoparietal Activity in Schizophrenia is Linked to a Specific Deficit in Goal Maintenance: a Multi-site Functional Imaging Study. *Schizophrenia Bulletin, 42*, 1149–1157.	Data for healthy controls and healthy controls' relatives were combined by computing outcomes for a joint healthy control group.
Reilly, J., Hill, S., Gold, J., Keefe, R., Clementz, B., Gershon, E., Keshavan, M., ... Sweeney, J. (2017). Impaired Context Processing is Attributable to Global Neuropsychological Impairment in Schizophrenia and Psychotic Bipolar Disorder. *Schizophrenia Bulletin, 43*, 397-406	Schizophrenia and schizoaffective groups were combined into a single patient group to compute outcome data. All relatives' data were excluded from meta-analysis since some relatives had psychotic disorders.

Richard, A., Carter, C., Cohen, J., & Cho, R. (2013). Persistence, diagnostic specificity and genetic liability for context-processing deficits in schizophrenia. *Schizophrenia Research, 147,* 75-80.	Relatives included in this sample were not relatives of actual participants, avoiding the non-independent data issue.
Thoma, P., & Daum, I. (2008). Working memory and multi-tasking in paranoid schizophrenia with and without comorbid substance use disorder. *Addiction, 103,* 774–786.	Dual diagnosis and alcoholism groups were not coded, only paranoid schizophrenia and major depressive disorder groups.
Thoma, P., Zoppelt, D., Wiebel, B., & Daum, I. (2007). Context processing and negative symptoms in schizophrenia. *Journal of Clinical and Experimental Neuropsychology, 29,* 428-435	Authors reported data on participants across median split high and low negative symptom groups; these groups were combined into a single schizophrenia group for meta-analysis.
Todd, J., Whitson, L., Smith, E., Michie, P., Schall, U., & Ward, P. (2014). What's intact and what's not within the mismatch negativity system in schizophrenia. *Psychophysiology, 51,* 337–347.	Young and matched-age healthy control groups were combined into a single healthy control group.

Woodward, N., et al. (2016). *Personal Communication.*	Group summary data was computed from raw data shared by Dr. Neil Woodward.
Zhang, Z., Chen, X., Yu, P., Zhang, Q., Sun, X., Gu, H., … Ji, F. (2015). Evidence for the contribution of NOS1 gene polymorphism (rs3782206) to prefrontal function in schizophrenia patients and healthy controls. *Neuropsychopharmacology, 40,* 2015–2016.	Genotype groups were combined into regular diagnostic groups to compute outcome data. The authors noted, "In the end, one strength of the study (i.e., the two groups matched in education) is also a weakness because the mean IQ of our SCZ sample was neither matched with that of the controls nor as low as typical SCZ patients (pg. 1392)."

APPENDIX H

ADDITIONAL POINT ESTIMATE DATA

Table H1

Point Estimates of Errors by Group and Trial Type.

	First-Episode, Unmedicated	First-Episode, Medicated	Chronic Schizophrenia	At-Risk Group	Healthy Controls	Psychiatric Comparison
Short Delay	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
% AX Errors	6.01 (1.63)	8.40 (6.95)*	8.63 (5.00)	3.70 (1.96)	2.83 (1.09)	6.66 (4.49)
% AY Errors	12.00 (1.00)	9.90 (9.67)*	16.87 (5.25)	8.76 (4.40)	9.37 (3.66)	11.50 (4.54)
% BX Errors	23.62 (2.71)	12.53 (10.53)*	21.83 (3.36)	12.36 (5.71)	8.71 (2.95)	12.30 (5.22)
% BY Errors	6.18 (2.02)	2.40 (4.97)*	7.08 (2.90)	2.07 (0.87)	2.21 (1.37)	2.74 (1.81)
Long Delay	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
% AX Errors	17.07 (5.82)	4.00 (5.00)*	12.10 (5.68)	13.50 (5.67)	6.32 (3.05)	9.17 (4.68)
% AY Errors	9.76 (6.24)	21.00 (19.00)*	17.61 (5.89)	11.25 (9.38)	11.83 (4.16)	15.22 (5.00)

% BX Errors	21.19 (1.46)	14.00 (13.00)*	22.87 (8.18)	16.17 (6.44)	8.39 (2.83)	11.74 (3.15)
% BY Errors	2.51 (1.25)	2.00 (4.00)*	6.92 (9.25)	1.85 (1.95)	3.49 (7.38)	2.63 (1.10)

*Unweighted standard deviation is displayed when only one study provided data.

Percent errors by trial type for short (cue-probe interval <3500ms) and long (cue-probe interval ≥3500ms) delay trials. Means and standard deviations are weighted by sample size. Data also presented in Figures 2 and 3.

Table H2

Point Estimates of Reaction Times by Group and Trial Type.

	First-Episode, Unmedicated	First-Episode, Medicated	Chronic Schizophrenia	At-Risk Group	Healthy Controls	Psychiatric Comparison
Short Delay	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
AX_{RT}	524.89 (13.96)	414.17 (66.75)*	481.28 (42.38)	433.34 (38.04)	430.11 (37.12)	452.45 (42.79)
AY_{RT}	644.83 (27.87)	467.22 (60.63)*	638.47 (53.84)	571.32 (38.37)	570.22 (38.15)	571.96 (50.58)
BX_{RT}	679.83 (65.46)	418.62 (104.68)*	516.58 (54.29)	498.60 (45.21)	429.46 (64.74)	561.13 (64.16)
BY_{RT}	525.94 (34.53)	408.25 (109.39)*	526.51 (52.31)	431.81 (34.90)	419.08 (42.05)	452.41 (34.22)
Long Delay	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
AX_{RT}	588.24 (26.25)	616.00 (204.00)*	550.72 (82.49)	520.37 (41.37)	480.11 (45.19)	496.07 (31.40)
AY_{RT}	694.36 (64.03)	785.00 (152.00)*	694.20 (63.79)	633.71 (57.97)	616.04 (50.91)	635.77 (29.53)
BX_{RT}	692.42 (64.03)	723.00 (286.00)*	595.84 (91.63)	553.58 (52.87)	508.29 (90.29)	540.26 (43.13)
BY_{RT}	574.36 (60.49)	626.00 (203.00)*	580.69 (53.06)	507.78 (56.69)	479.80 (49.84)	515.12 (34.09)

*Unweighted standard deviation is displayed when only one study provided data.

RT=reaction time. Reaction time by trial type for short (cue-probe interval <3500ms) and long (cue-probe interval ≥3500ms) delay trials. Means and standard deviations are weighted by sample size. Data also presented in Figures 4 and 5.

Table H3

Point Estimates of d'context by Group.

	First-Episode, Unmedicated M (SD)	First-Episode, Medicated M (SD)	Chronic Schizophrenia M (SD)	At-Risk Group M (SD)	Healthy Controls M (SD)	Psychiatric Comparison M (SD)
Short delay d'context	2.43 (0.38)	2.89 (1.03)*	2.43 (0.57)	3.45 (0.68)	3.54 (0.54)	3.09 (0.45)
Long delay d'context	1.99 (0.35)	3.26 (1.02)*	2.10 (0.73)	2.67 (0.44)	3.06 (0.51)	2.72 (0.29)

*Unweighted standard deviation is displayed when only one study provided data.

D'context for short (cue-probe interval <3500ms) and long (cue-probe interval ≥3500ms) delay trials. Means and standard deviations are weighted by sample size. Data also presented in Figures 6 and 7.