STABILITY AND CHANGE IN SYMPTOMS, COGNITION, AND COMMUNITY OUTCOME IN SCHIZOPHRENIA

ASHLEY A. A. MILES

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY

YORK UNIVERSITY

TORONTO ONTARIO

AUGUST 2012

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-92791-5

Our file Notre référence
ISBN: 978-0-494-92791-5

Abstract

It has been well established that neurocognitive deficits are a core feature in schizophrenia and predict difficulties in functional independence. However, few studies have assessed the longitudinal stability of cognition and key aspects of functional outcome concurrently. Instead, research has either focused on the stability of cognition alone, the stability of real world outcome alone (though this has received only very recent interest), or explored baseline cognition as a predictor of later community functioning. Accordingly, this study will assess the extent to which significant changes in cognition and community status are independent or related. As a point of comparison, the stability of clinical symptom status and the relationship between symptom change and outcome change will be included. Assessments were conducted evaluating clinical status, cognitive abilities, and functional outcome in 128 patients with schizophrenia at baseline and again one year later. Intraclass correlation coefficients were used to index stability and reliable change index analyses quantified the prevalence of significant improvement or deterioration in each of the three illness features. These data were then used to identify and compare the ability of either symptom change or cognitive change to predict concurrent changes in community status. Results from these analyses revealed that symptom status, cognitive functioning, and community outcome are all similarly stable features of schizophrenia in an outpatient sample receiving treatment as usual. A small proportion of the sample demonstrated significant improvement or deterioration in these domains, with only weak evidence that such change was predicted by changes in symptoms or cognition. Further, there was no strong evidence of a preferential

relationship for cognition (relative to symptoms) in relation to functional outcome. These results shed light on the strength and nature of the cognition-real world outcome relationship in schizophrenia and have implications for pharmacological interventions aimed at improving functional status with cognitive enhancing medications.

Keywords: Schizophrenia; Symptoms; Cognition; Functional Outcome; Stability; Change

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Walter Heinrichs, for accepting me into the Clinical Psychology program at York University. I am grateful for his mentorship, patience, and encouragement to grow as an independent clinician and researcher. The countless opportunities provided under his supervision have played a pivotal role in my success throughout graduate studies.

I am indebted to Dr. Joel Goldberg for hiring me as summer student at the Hamilton Program for Schizophrenia. This invaluable experience laid the groundwork for the rest of my career as a Clinical Neuropsychologist. I am grateful for his unwavering optimism, kindness, and belief in me. Many thanks also to Dr. Stephanie McDermid Vaz, the remaining member of my dissertation committee, who offered insightful comments and advice throughout the process.

This research would not have been possible without the contributions of many people. Thanks to our research assistants on these projects (Diana Smith and Ashley Oman) and our many volunteers and thesis students who provided assistance with data collection and quality control along the way. Our team is grateful for the cooperation from the case managers and staff at each community clinic and to the patients for contributing to these research projects. I am grateful to the Canadian Institutes of Health Research (CIHR) and Ontario Mental Health Foundation (OMHF) for providing financial support.

Finally, my deepest thanks and appreciation are extended to my family and friends who have kept me motivated, distracted, and supported me along the way.

Specific thanks to: Chris, there really are no words to express my gratitude; my parents for their unwavering and often blind faith ("I know you will get the scholarship, match to an internship site, pass your defense, etc. I'm the Mom. I just know"); Kimberley for absolutely everything the past six years, but especially her friendship; Tina for her support (and her food!); and to Michael, Narmeen, Stephanie, and Cooper.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Stability and Change in Symptoms, Cognition, and Community Outcome in Schizophrenia

Introduction

Overview

Schizophrenia is a debilitating illness affecting approximately 1% of the population (Government of Canada, 2006) or over 300 000 adults across Canada. It has been ranked among the top 10 leading causes of disability in the world (World Psychiatric Association, 2001), and in Canada alone, the 2004 estimated prevalence-based cost of illness was $6.85 billion annually including direct and indirect health care costs, family benefits, social support services, and productivity loss due to morbidity or early mortality (Goeree et al., 2005). Individuals with schizophrenia suffer from a range of distressing clinical symptoms, struggle with poverty and homelessness, and they often fall victim to stigmatization and social marginalization (Sartorius, 2002). Moreover, in spite of thousands of published studies investigating schizophrenia, there remains no definitive knowledge surrounding the etiology and pathophysiology of the disease making it challenging to provide patients and their families with definitive information about the cause of illness, likelihood of responses to specific treatment regimens, and prognosis. As such, considerable effort has been invested into attempting to identify specific genetic and environmental factors that lead to the schizophrenia syndrome, as currently described by the two leading classification systems, International Classification of Diseases-10th revision (World Health Organization, 1993) and Diagnostic and Statistical Manual of Mental Disorders-4th edition Text Revision (American Psychiatric

Association [APA], 2000).

Evidence from decades of rigorous scientific exploration reveals that current psychopharmacological interventions are relatively effective at treating positive symptoms and to a lesser extent negative symptoms (Tandon, Nasrallah, & Keshavan, 2010). Yet, schizophrenia has a complex clinical presentation and many aspects of the disorder remain little understood. Accordingly, significant research and resources are being directed toward understanding and treating other correlates of the illness including cognitive functioning, subjective well-being, quality of life, and community independence. For example, extensive evidence has demonstrated that cognitive disability is a key predictor of functional outcome in schizophrenia (Green, 1996). In view of the putative strength of this relationship, a new wave of treatment has emerged with the goal of improving community independence through increasing cognitive functioning.

Significant progress has been made toward understanding the nature of the cognition-functional disability relationship. This is at least in part attributable to a surge of large, multisite initiatives investigating the prevalence, severity, treatment of, and relationship between cognition and community outcome. More recent efforts directed at understanding these relationships include: National Institute of Mental Health Measurement and Treatment Research to Improve cognition in Schizophrenia (NIMH-MATRICS; Marder & Fenton, 2004), NIMH - Cognitive Neuroscience Approaches to the Treatment of Impaired Cognition in Schizophrenia (CNTRICS; Carter et al., 2008), and Validation of Everyday Real World Outcomes (VALERO; Leifker, Patterson, Heaton, &

Harvey, 2011). Yet, little is known about the prevalence of cognitive and functional change in typical outpatient populations receiving treatment as usual. Moreover, there is little if any information on the contingency of functional improvement on cognitive change. Hence, as resources continue to be directed toward improving cognition in schizophrenia based largely on the argument that such improvement will lead to enhanced functionality and community independence, this dissertation will explore the prevalence of significant cognitive and functional change in typical schizophrenia outpatients as well as the nature and strength of this relationship. As a comparison, the stability and prevalence of change in clinical symptom severity will also be considered, including their relationship with change in community independence status.

Background Information on Key Diagnostic Criteria

Clinical symptoms. At present, the current syndrome description, according to DSM-IV-TR (APA, 2000) allows for considerable heterogeneity in illness presentation among patients and thus may contribute to continued uncertainty surrounding treatment options for symptoms, etiology of the disorder, and long-term prognosis. This variability in illness presentation results from the structure of the current edition of the DSM which permits the first diagnostic criterion (i.e., Criterion A) to be satisfied by the presence of a single bizarre symptom or a combination of two or more other symptom types. The five possible symptom categories include i) hallucinations, which are sensory experiences that do not exist outside one's mind (e.g., hearing voices), ii) delusions or implausible beliefs that are firmly held (e.g., paranoid thoughts that one is being watched or monitored), iii) disorganized speech (e.g., tangential or nonsensical discourse), iv) catatonic behaviour

(e.g., motoric immobility), although due to low prevalence rates this is being considered for removal in DSM-V, and v) negative symptoms, which generally involve a deficit in "normal" affect or thought processes and include flat or blunted affect, poverty of speech (alogia), and lack of motivation (avolition). Since the introduction of the DSM-III criteria for schizophrenia (APA, 1980), this collection of symptoms and the diagnosis of the disorder itself has been valid, reliable, and stable over time (Harvey et al., 2012a). However, parsing schizophrenia patients by symptom presentation has not significantly improved understanding of the illness. Hence, in recent decades, other correlates of schizophrenia have been investigated and captured the interest of researchers.

Functional impairment. In addition to satisfying the clinical symptom criterion described above, the DSM also requires the presence of significant dysfunction in one or more major areas of functioning (Criterion B). Over the past decade, it has been repeatedly demonstrated that patients with schizophrenia spectrum disorders exhibit marked functional impairments, which are expressed in daily living skills across vocational, educational, and residential environments (Jaeger, Berns, & Czobor, 2003). These deficits have also been reported in social and community settings (e.g., effectively communicating with others; orienting and traveling within one's community) and contribute individually and collectively to the long term disability seen in the illness (Couture, Penn, & Roberts, 2006; Green, Kern, Braff, & Mintz, 2000; Keefe, Poe, Walker, & Harvey, 2006). For example, many patients are unable to implement or maintain basic self-care activities, including a variety of personal hygiene skills (Evans et al., 2004), and many are unable to fulfill basic social roles (e.g., as a parent, spouse, or

employee; Bellack et al, 2007). Furthermore, research has shown that, at most, one-third of individuals with this illness participate regularly in the work force (Bellack et al., 2007). However, workplace participation in the schizophrenia population drops to less than 15% when considering only competitive employment environments (Dickinson, Bellack, & Gold, 2007) and has been reported to be as low as 4% in some UK samples (see Marwaha & Johnson, 2004 for a review).

The effect of functional impairment has become more visible since the 1960s with progressive deinstitutionalization of people with mental illness (Mausbach et al., 2008). As a consequence, many patients with severe psychiatric illnesses are no longer residents of psychiatric hospital settings, and instead rely on alternative outpatient treatment facilities in community settings (Barnard-Thompson & Leichner, 1997; Lamb & Bachrach, 2001). Unfortunately, the resources available to the large number of individuals with this illness are limited and patients struggle to function independently with only modest support from outpatient rehabilitation and community programs. Consequently, it has become critical to identify key correlates of functional impairment and to integrate these findings with rehabilitation and treatment programs in order to maximize each patient's community independence potential.

Cognitive impairment. Although not a diagnostic criterion in the current edition of the DSM, impaired cognitive function has been established as a core and enduring feature of schizophrenia (Heinrichs & Zakzanis, 1998) and it is being considered as an additional criterion in DSM-V. Specifically, these deficits are evident in varying degrees during the prodromal period (Keefe et al., 2006, Seidman et al., 2010), deteriorate up to

the first episode, and then remain stable throughout the illness, as shown both through cross-sectional comparisons of first episode and chronic patients (Mesholam-Gately, Giuliana, Goff, Faraone, & Seidman, 2009; Sponheim et al., 2010) and through longitudinal studies (Kurtz, 2005; Kurtz, Seltzer, Ferrand, & Wexler, 2005). A review of cognitive impairment among geriatric patients with schizophrenia revealed the same pattern of persistent deficits (Irani, Kalkstein, Moberg, & Moberg, 2011). Further, prospective studies of first episode patients show stability in cognitive functioning for up to 10 years, with possible deterioration noted only in some aspects of verbal memory (Bozikas & Andreou, 2011). Thus, impaired cognition is a persistent deficit among individuals with the illness.

In addition to the enduring nature of these deficits, research on performance across many neurocognitive measures demonstrates that patients differ from control populations by up to 1.5 standard deviation units on indicators of processing speed (Dickinson, Ramsey, & Gold, 2007) and verbal memory, and averages of 1.0 standard deviation units on tests of attention, executive function, language, motor skills, spatial abilities, and general intelligence (e.g., Keefe & Fenton, 2007). While these impairments have been shown to affect virtually all aspects of cognition, specific attention has been given to demonstrated ability on measures assessing working memory, verbal learning and memory, visual learning and memory, attention and concentration, reasoning and problem solving, speed of information processing, executive functions, and social cognition (Elvevag & Goldberg, 2000; Green, Kern, & Heaton, 2004; Heinrichs & Zakzanis, 1998; Matza et al, 2006). These cognitive deficits are considered the most

stable aspect of schizophrenia, with test-retest coefficients ranging from .70 to .85.

Unlike the presence of positive and/or negative symptoms and functional deficits that are observed in all patients upon receiving a diagnosis of schizophrenia, cognitive impairments do not necessarily exist among all individuals. For example, Heinrichs and colleagues (2008) identified a subset of patients with superior-range verbal abilities on the vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997). These patients were compared to healthy adults with similarly superior-range vocabulary scores in terms of performance on other cognitive domains including processing speed, non-verbal reasoning abilities, working memory, verbal learning, word fluency, and response inhibition. There were no group differences on any task among the two groups, suggesting that a subset of patients can be free from cognitive impairment. The investigation of the validity of average-to-superior range performance among schizophrenia patients has been demonstrated in other independent samples (e.g., Kurtz, Donato, & Rose, 2011; MacCabe et al., 2012).

Research on these deficits in schizophrenia has extended to evaluating the relationship between cognition and functional outcome. However, prior to reviewing this extensive literature, it is critical to also understand research on clinical symptoms in schizophrenia including temporal stability and their role in predicting cognitive abilities and community independence.

Clinical symptoms and their relationship with cognitive functioning and outcome.

Stability of symptoms in schizophrenia. Research on the temporal stability of symptoms in schizophrenia is relatively scarce. Instead, focus appears to be on diagnostic

stability within the disorder (e.g., Harvey et al., 2012a; Helgeland & Torgersen, 2005), and on the dissociability (e.g., Dollfus & Petit, 1995) or stability of different symptom factor structures (e.g., Goldman, Tandon, Liberzon, Goodson, & Greden, 1991; Lançon, Auquire, Nayt, & Reine, 2000; Reichenberg, Rieckmann, & Harvey, 2005). However, research on the temporal stability of symptoms does exist. For example, Rey and colleagues (1994) examined mean stability of five symptom factors over a three-year follow-up and showed that both positive and negative symptoms are stable over time. However, positive symptoms were assessed using the Present State Examination and Psychological Impairment Rating Scale, which are not the current convention for assessment. Malla and colleagues (1993) used more typical assessment instruments (i.e., Scale for the Assessment of Negative Symptoms and Scale for the Assessment of Positive Symptoms) and found moderate stability of most symptoms over 12 months. In contrast, while Arndt and colleagues (1995) reported that negative symptoms were stable over two years, positive and disorganized symptoms declined. Wolter et al (2010) also demonstrated instability in positive symptoms, but showed these changes to be bidirectional. Other findings have supported the notion that negative symptoms are stable over both shorter and longer follow-up periods (Fennig, Bromet, Galambos, & Putnam, 1996; Johnstone, Owens, Frith, & Crow, 1987). This trend was not supported by Park and colleagues (2004) who demonstrated insufficient stability in negative symptoms over a 5 year follow-up study or Lindenmayer and colleagues (1986) who found that neither positive nor negative syndromes were stable over a two-year window.

Taken together, research on the temporal stability of symptoms in schizophrenia

is mixed. Further, existing studies have been limited by assessing stability in inpatient or first episode samples (e.g., Goldman et al., 1991) where symptoms are likely to change following treatment (Kay & Lindenmayer, 1991). Some of these studies have used test-retest windows that are quite short (e.g., four weeks) and sample sizes that are too small to generalize to the schizophrenia population (e.g., n=37, 40). The measures used in several of the older studies (e.g., Rating Scale for Chronic Psychotic Patients, Psychological Impairment Rating Scale) are no longer the convention in current research. Lastly, studies often rely on less sophisticated statistical analyses including correlational measures (e.g., Pearson product-moment correlations), which are not ideal for assessing stability and reliable change over time. Thus, further research is necessary regarding the longitudinal stability of symptoms in chronic schizophrenia patients using global and specific scales of more conventional instruments.

Symptoms and cognition. The relationship between psychopathology and neurocognition in schizophrenia continues to provide relatively consistent findings within the different symptom clusters. For example, O'Leary and colleagues (2000) considered three symptom domains: positive, negative, and disorganized symptoms, which were confirmed as separable clusters from previous factor-analytic work. Correlations were obtained for each symptom domain and 20 different cognitive abilities. Nine of 20 correlations between negative symptoms and cognitive scores primarily indexing verbal memory, attention, and fluency measures were significant, but with a very modest average coefficient of $r = .26$. Three correlations were significant in the disorganization domain, also with an average of $r = .26$. Only 2 out of 20 correlations were significant in

the psychotic symptoms data, with an average of $r = .24$. Leung and colleagues (2008) approached the question in a different way, comparing neuropsychologically "normal" and impaired patients to those who were symptomatic and those who were deemed to be in remission. Their findings revealed no significant interaction between neuropsychological and symptom status. In another study published in the same year, Heinrichs et al (2008) found no differences in positive, negative, or mood-related symptoms between patients with and without superior-range verbal abilities. More recently, a systematic review of the symptom-cognition relationship was conducted by Dominguez and colleagues (2009), which evaluated findings from a 21-year period spanning 1986-2007. Their analyses revealed that negative and disorganized symptoms had significant but modest correlations with cognitive functioning (i.e., ranging from $r = -.29$ to $r = -.12$), whereas positive and mood symptoms were not significantly related to cognition. Ammari, Heinrichs, and Miles (2010) also considered the symptom-cognition relationship and largely confirmed this pattern of findings in that no differences in positive symptoms were found between groups of patients with average-range cognitive abilities and those with both intellectual and memory deficits. However, significant differences were noted in negative symptoms. These findings were also demonstrated in a study where patients who showed a decline in IQ from premorbidly high average levels had significantly greater negative symptom pathology than non-declining patients (MacCabe et al., 2012). Taken together, these data show that negative and disorganized symptoms range in their relationship to cognitive tasks from non-significant to moderate associations, whereas positive and mood symptoms tend to be independent of cognitive

functioning.

Symptoms and functional outcome. Similar to the symptom-cognition literature, understanding of the relationship between symptoms and functional outcome is complicated by inconsistent findings. However, like the research discussed previously, there may be a trend pointing toward a relationship between negative but not positive symptoms and outcome. For example, in a recent article reviewing patient data (n=1447) from the NIMH CATIE trial of chronic schizophrenia (Rabinowitz et al., 2012), negative symptoms were found to have a distinctive and independent effect on quality of life relative to other symptom domains (e.g., positive, excited, disorganized, and depressive symptoms). These findings differ somewhat from a comprehensive meta-analysis of symptom-outcome data published between 1966 and 2005 (Eack & Newhill, 2007). Their findings revealed weak, negative relationships between psychiatric symptoms and quality of life, with general psychopathology having the strongest relationship. However, they noted that positive and negative symptoms were not associated with quality of life in all patients equally, but that the relationship varied with particular demographic characteristics (e.g., inpatients compared with outpatients; early versus later course of illness).

Perhaps the best understood relationship between psychopathology and outcome is that of negative symptoms predicting successful work outcome. This association has been demonstrated across many different studies with results suggesting that negative symptoms predict successful work performance, in some cases up to 18 months later (e.g., Erickson, Jaafari, & Lysaker, 2011; Marwaha & Johnson, 2004). As further

evidence, a review article by Tsang and colleagues (2010) that considered 62 studies published after 1998 also provided strong evidence of a relationship between negative but not positive symptoms and vocational outcome.

Taken together, although there are promising results with regard to the potential predictive ability of negative symptoms in determining employment success, the role of other symptoms and community outcome domains is less clear. Further, even in the event of successful treatment of clinical symptoms in the disorder, many patients continue to experience reduced quality of life and difficulties in several real world settings including social, occupational, and residential domains (Harvey, Green, Keefe, & Velligan, 2004). In contrast, there is overwhelming evidence across several decades to support the strength of a cognition-outcome relationship in schizophrenia.

Cognition and Functional Outcome in Schizophrenia

Since the publication of Michael Green's influential manuscript in 1996, titled "What are the Functional Consequences of Neurocognitive Deficits in Schizophrenia?" a substantial increase was noted in the number of publications investigating the relationship between cognitive impairments and decreased community independence. Evidence from this past decade of research has consistently shown that neurocognitive ability is a significant predictor of critical functional abilities, such as work performance and independent living skills. Moreover, research has demonstrated that these community deficits are better predicted by impaired cognition than by positive or negative symptoms (Keefe & Fenton, 2007). In fact, in some studies, researchers failed to find a relationship between symptoms and functional status, but were able to demonstrate a significant link

between specific cognitive domains (i.e., verbal learning) and outcome (Kurtz, Wexler, Fujimoto, Shagan, & Seltzer, 2008).

Previous review articles have shown strong relationships between cognition and outcome, and have demonstrated that deficits in cognition account for up to 60% of the variance in social skills, community independence, and skill acquisition in outpatient rehabilitation programs (e.g., Green et al., 2000; Green et al., 2004). A more recent meta-analysis of predictors of community outcome from cognitive measures revealed somewhat more modest findings. Fett and colleagues (2011) considered 52 studies published between 1991 and 2008 and reported that the variance in outcome accounted for by cognitive measures ranges from 4% to 23%.

Nonetheless, the relationship between cognition and outcome continues to receive attention from researchers and findings consistently show significant associations between the two. For example, McGurk and Mueser (2004) found that impaired mental flexibility, learning and memory, and processing speed were associated with poor work performance. Other research (Niekawa, Sakuraba, Uto, Kumazawa, & Matsuda, 2007) showed a relationship between impaired cognitive functioning and poor financial competence. Schutt, Seidman, Caplan, Martsinkiv, and Goldfinger (2007) found that higher executive functioning abilities predicted enhanced self-care, better verbal memory predicted more positive social contacts, and a higher capacity for sustained attention predicted better communication skills. Previous reviews of the literature (e.g., Green, 1996; Green et al., 2000) have reported effect sizes of the relationship between these individual cognitive abilities and functional outcome in the medium range (i.e., Cohen's

d = .50), and even larger effects when composite indicators of cognitive functioning are calculated (Green et al., 2004). Accordingly then, given the statistical strength of the relationship between a diverse range of functional domains and neurocognitive abilities as well as the enduring nature of these deficits, extensive scientific and financial resources are continually being directed toward the discovery and evaluation of cognitive enhancing medications. In view of the strong relationship between cognitive abilities and community independence, the ultimate goal of these pharmacological interventions is to improve real world community independence by increasing cognitive functioning.

In light of this emerging interest in improving functional outcome in patients via pharmacological (e.g., NIMH-MATRICS; Marder & Fenton, 2004) and new behavioural interventions (McGurk et al., 2007), it is important to understand specific aspects of the cognition-functional outcome relationship, including the stability of each and their relationship with each other.

Stability of Cognitive Functioning

Over the past decade, several studies have assessed the longitudinal stability of different cognitive abilities in schizophrenia, albeit using a wide range of methods and patient populations (e.g., chronicity of illness or severity of clinical symptoms). For example, Heinrichs, Ruttan, Zakzanis, and Case (1997) used a subtype approach based on impaired or intact performance across different cognitive domains and assessed the stability of the different groups over a 3-year period. Heaton and colleagues (2001) assessed the stability of specific cognitive ability domains (e.g., verbal abilities, psychomotor abilities, etc.), using composite scores from several subtests, and compared

the performance of schizophrenia patients and normal controls over short and long follow-up intervals. Other researchers have focused on specific tests of cognitive functioning in groups of first episode patients (e.g., Leeson et al., 2009) or compared performance of first episode to chronically ill patients to assess the stability of cognition in schizophrenia across the illness (Sponheim et al., 2010). More recent investigations of cognitive stability have relied on statistical comparisons of estimates of premorbid IQ (e.g., through word reading tests) with current performance on different tests of cognition. For example, Leeson and colleagues (2011) looked at stability of IQ in a first episode sample using the Wechsler Test of Adult Reading as their estimate of premorbid IQ and a prorated IQ estimate indexing current IQ from the short form of the WAIS-III (information, block design, arithmetic, and digit symbol coding). Half of their sample demonstrated stable IQ up to their first episode whereas 44% demonstrated significant decline.

In spite of these investigations, the assessment of stability of cognitive ability domains in schizophrenia remains relatively novel and continued empirical investigations are necessary to determine prevalence of cognitive stability, especially in the absence of pharmacological or behavioural interventions, and for individual ability structures rather than global composite scores. Indeed, in a recent paper summarizing the official position of the American Academy of Clinical Neuropsychology (AACN) on serial neuropsychological assessments (Heilbronner et al., 2010) it was recommended that data be collected on "normal change trajectories for all types of measures with all types of... patient groups" (p.1274). This would ideally identify which tests are most vulnerable to

change with repeated administration in a particular clinical sample, and thus indicate which tests might be most suitable for intervention-based studies.

Stability of Community Independence

In contrast, research on the stability of functional independence has received only recent, albeit minimal interest (e.g., Leifker, Patterson, Bowie, Mausbach, & Harvey, 2010) and such studies have for the most part approached the question of stability from a psychometric standpoint (e.g., test-retest reliability of new or existing measures of community independence) or retrospectively, rather than investigating a priori hypotheses using prospective studies to specifically investigate the progression of different functional independence domains in patients with schizophrenia over time. Another shortcoming of research in this area is that most studies assessing intervention effectiveness have employed measures that more accurately assess capacity for functional independence rather than using instruments that capture an individual's real world functional status (for further discussion see Miles, Heinrichs, & Ammari, 2011).

One example of a widely used measure of functional capacity is the University of California San Diego (UCSD) Performance Based Skills Assessment (UPSA; Patterson, Goldman, McKibbon, Hughs, & Jeste, 2001). This measure assesses functional skills in standardized role-playing situations across five domains including comprehension and planning, finances, transportation, household management, and communication. The UPSA has demonstrated good psychometric properties (e.g., Leifker et al., 2010) and strong correlations with performance on several cognitive tests (e.g., Vella et al., 2011). However, these types of measures are at best simulations of real world functioning, and

recent research has shown that the UPSA shares extensive variance with cognitive test performance and provides only marginal new validity in predicting community outcomes (Heinrichs, Ammari, Miles, & McDermid Vaz, 2010). Accordingly, future research using measures of real world functioning with good reliability and validity is imperative in order to determine the occurrence of significant change in functional status, especially in the absence of an active intervention.

Previous research on the relationship between change in cognition and change in outcome. In spite of the wide range of interest in and research on cognitive functioning in chronic and first episode patients and the emerging interest in the stability of community independence, studies on the relationship between cognition and outcome have typically been conducted at only one point in time (Matza et al., 2006). Few if any studies have explored the important relationship between changes in cognitive abilities over time as predictors of change in community independence, which is critical to evaluate the possible functional implications of interventions that target cognition in schizophrenia (Matza et al., 2006). Nonetheless, several studies have begun to approach this question by first examining the ability of baseline cognitive abilities to predict functional status at follow-up intervals ranging from a few months to 20 years (in a retrospective study). Results of these longitudinal studies (see Green et al., 2004 for a review) have largely revealed that patients with better baseline cognition show greater changes (i.e., improvements) in community functioning over time. In fact, one study noted that neuropsychological functioning predicts up to 74% of the variance in improvements in work performance (Bell, Bryson, & Kaplan, 1999). The study of

changes in *both* cognition and outcome by assessing each of these variables at multiple points in time is not completely novel. However, the generalizability of existing work is significantly limited by methodological shortcomings and inconsistent findings between studies.

One research group reviewed the relationship between cognitive remediation training and improvements in work outcomes and showed preliminary results suggesting that improved cognition results in better job functioning, as well as improved social and interpersonal functioning, which they argue to be critical for sustaining employment (Wexler & Bell, 2005). Further, Bowie, McGurk, Mausbach, Patterson, & Harvey (in press) recently demonstrated that a brief cognitive remediation intervention (two-hour weekly sessions for 12 weeks) led to robust improvements in cognition. However, significant effects were noted in household activities and work skills only when cognitive remediation therapy was combined with a functional skills training program.

A comprehensive review of this topic was conducted by Matza and colleagues (2006) who evaluated nine studies that investigated the association between changes in cognition and changes in functional outcome. They concluded that preliminary findings offer support for a link between changes in the two. However, there are several limitations among the studies reviewed. For example, several investigations relied on a composite cognitive score (Friedman et al., 2002; Hogarty et al., 2004; Wykes, Reeder, Corner, Williams, & Everitt, 1999) or a measure that is not sensitive or comprehensive enough for use in this population (e.g., Mini-Mental State Examination in Harvey et al., 1999; Alzheimer's Disease Assessment Scale in Harvey et al., 2003). Further, in addition

to difficulties with the assessment of cognitive change, the functional measures included in this study (e.g., Quality of Life Scale: Heinrichs, Hanlon, & Carpenter, 1984) have been criticized and are not among the measures currently being evaluated for use in schizophrenia research (e.g., Leifker et al., 2011; Miles et al., 2011). In addition, six of the eight measures used were not designed specifically for use in schizophrenia and therefore do not assess key domains relevant to this population including independent living ability, occupational limitations, and personal care difficulties (Matza et al., 2006). The measures employed varied across studies and therefore different domains of outcome were being assessed (e.g., social functioning, occupational functioning, general quality of life). This complicates the generalizability across studies and again speaks to the need for a gold standard measure for assessing real world outcome in schizophrenia. Other limitations of these existing studies include the use of only geriatric (Friedman et al., 2002; Harvey et al., 1999; Harvey et al., 2003) or inpatient samples (Spaulding et al.,1999), or using follow-up assessments that are too short to effectively comment on longitudinal change in schizophrenia (e.g., a 3- or 6-month test-retest interval in four of the studies). Of the nine studies discussed, only three looked at treatment as usual, but were limited to the geriatric samples noted above, whereas the remainder were divided between pharmacological (Buchanan, Holstein, & Breier, 1994; Galletly, Clark, McFarlane, & Weber, 1997; Velligan et al., 2003) or psychosocial interventions (Spaulding et al., 1999; Hogarty, et al., 2004; Wykes et al., 1999).

More recent evaluations of the relationship have been largely focused on cognitive or psychosocial interventions and have shown promising findings. For example,

Wykes and colleagues (2007) used a cognitive remediation intervention and found a relationship between improvements in working memory and improved social functioning. Fiszdon and colleagues (2008) employed a psychosocial therapy and reported associations between memory and executive functioning and changes in quality of life. Further, cognitive enhancement therapy and functional outcome were shown to be related by Eack and colleagues (2010), though this was demonstrated in a group of early-course schizophrenia patients with a mean age of 25 years. In a study combining neurocognitive enhancement therapy (NET) and supported employment, patients receiving both interventions worked significantly more hours (Bell, Zito, Greig, & Wexler, 2008). As Matza and colleagues (2006) found, these and other intervention-focused studies have been limited by their choice of outcome measures, use of a patient group that is not generalizable to the typical schizophrenia patient population, and assessment of only one functional domain (e.g., work, social functioning, etc.).

Summary of limitations in previous studies. The existing literature on the association between changes in both cognition and functional outcome remains relatively scarce and plagued with methodological shortcomings. These include, but are not limited to the following: poor generalizability of the sample (e.g., geriatric or first episode; inpatient only), too short of a test-retest window, use of composite rather than individual test scores for both cognitive and functional measures that may mask important differences in specific domains, employment of functional measures not sensitive to the schizophrenia population or that do not tap critical community independence domains; as well, few studies have focused on chronic, outpatient populations receiving treatment as

usual. Among these shortcomings, the issue of appropriate measure selection is especially important.

The functional measures chosen in more recent assessments of outcome are often those assessing functional capacity (e.g., Kurtz et al 2008) rather than real world outcome. These instruments assess practical cognition, that is, what one is capable of doing rather than actual community performance (Gupta, Bassett, Iftene, & Bowie, 2012) and thus are related more closely to cognitive performance than to real world outcome (Heinrichs et al., 2010). Therefore, their use in cognitive change-functional change studies may not be ideal. Moreover, the use of composite measures of either cognition or functional outcome deserves further discussion. For example, the MATRICS cognitive consensus battery (MCCB; Nuechterlein & Green, 2006) is among the most frequently used cognitive batteries in schizophrenia research. It provides a composite indicator of cognitive ability that is frequently reported by researchers where an aggregate of performance on tests from seven cognitive ability domains collectively contribute to the total score. However, these types of summary scores, like full scale IQ estimates, have the potential to conceal potential cognitive deficits or exceptionality. As a case in point, Wilk and colleagues (2005) studied schizophrenia patients matched to healthy people with average full scale IQs and found that subtest profiles differed significantly, with patients showing relative deficiencies on memory and processing speed tasks and relative superiority in verbal comprehension and non-verbal reasoning. Accordingly, patients may score in a norms-defined "average range" on a battery of measures or on a composite score like IQ, but still demonstrate domain specific abnormalities or deficiencies.

Although previous studies have presented promising preliminary findings based on assessments at one or multiple points in time, it remains difficult to generalize about the potential functional implications of treatments that target cognitive abilities in schizophrenia. Further, the assessment of this relationship is correlational in the vast majority of studies, which does not ensure that an intervention which increases key cognitive abilities will necessarily translate into improved community independence. This highlights the importance of assessing the ability of cognitive change to predict actual changes in functional independence. Further, it is critical then to evaluate this relationship in terms of changes in specific cognitive and functional domains, rather than using composite indicators of performance. This will aid in elucidating the extent to which improvements in particular cognitive ability domains are required for similarly significant changes in specific indicators of functional status.

Few if any studies have assessed the cognition-outcome change relationship in the absence of specifically targeted cognitive interventions (i.e., in patients receiving treatment as usual) in the chronic, outpatient schizophrenia population. This evaluation is vital if future clinical trials on pharmacological or behavioural interventions in schizophrenia proceed under the assumption that functional outcome improvements occur primarily in the face of cognitive enhancement. Indeed, Matza and colleagues (2006) concluded that further research on this relationship is required in order to "identify the functional outcome domains and measures that will be most sensitive to current change in cognition" (p.674). Further, prevalence data on the stability and propensity to change in both cognitive and community outcome domains is essential to provide a necessary

benchmark against which the potential value of cognitive rehabilitation can be evaluated. That is, if cognitive change cannot be shown to be significantly associated with changes in community status, then the enormous resources being directed at cognitive enhancement via pharmacological and behavioural interventions may not be justified.

In view of the above considerations, this study will seek to answer the following questions.

Research Questions

1. How prevalent is change in symptom severity over a one-year period in the schizophrenia outpatient population receiving treatment as usual?

   Expected results: In view of conflicting findings, it is unclear whether positive and negative symptoms will change in a chronic, outpatient sample involved in community treatment settings. Further, there appears to be little if any research on subscales assessing symptoms other than those measured by positive, negative, and disorganized syndrome scales. Thus, this question is more exploratory and will provide valuable information on a cluster of symptoms, which include anergia, thought disturbance, activation, paranoid belligerence, and depression.

2. How prevalent is change in cognitive abilities, that is, improvement or deterioration in cognitive functioning over a one-year period in the schizophrenia outpatient population?

   Expected results: In light of the paucity of research in this area, the prevalence of significant cognitive change is unclear. However, previous research has demonstrated that cognitive impairments deteriorate up to the first episode and then remain stable throughout the illness (e.g., Mesholam-Gately et al., 2009; Sponheim et al., 2010). Thus,

in view of the chronically ill population in this study, it is anticipated that the assessed cognitive domains will remain stable across the assessment period. If changes exist, it is expected that only a small percentage of patients will demonstrate improvement or deterioration.

3. How prevalent is change in community independence, that is, improvement or deterioration in functional status over a one-year period in the schizophrenia outpatient population?

   Expected results: In keeping with the strong predictive relationship between neurocognition and functional outcome, it is expected that, given the hypothesized stability of cognitive impairment, patients will show similarly stable patterns of impairment in community independence across the assessment period.

4. Is there evidence of associations between symptom and community independence change and cognitive and community independence change? Further, if these relationships exist, is the relationship stronger for the cognition change-outcome change association?

   Expected results: Given the demonstrated relationship between static assessments of cognition and community independence, it is expected that improvements in cognitive abilities will be associated with greater functional independence in the community. For example, improvements in outcome may involve patients incurring additional vocational responsibilities or requiring less support across different community environments.

It is expected that the symptom-community outcome change relationship will also replicate findings from cross-sectional assessments. That is, changes in negative symptoms will have modest relationships with change in outcome status, whereas

changes in positive symptoms and depression will correlate weakly with community independence change, or have no relationship at all. It is less clear how the cluster symptoms will be related to outcome.

In view of the putative relationship between cognition and outcome, it is expected that a stronger association will be shown between these variables as compared to symptom-outcome change variables.

Overview of Reliability Change Theory and Application

Classic test theory. In contrast to other fields of science, the measurement of human behaviour is imprecise and does not lend itself to perfectly stable or reliable results (Lineweaver & Chelune, 2003). In classic test theory, an individual's actual performance or score on any given test (i.e., observed score) is a combination of a true score and error (Streiner & Norman, 2008). Accordingly, when conducting a single assessment or observing change in test scores, one must account for error. Systematic error and random error are the two types considered.

Within the former category, systematic effects influence most individuals in the group and can include practice effects (e.g., explicit memory for test content, procedural learning, general familiarity with the testing process and/or examiner), aging effects (e.g., improvements in the developmental population or deterioration in the geriatric population), and other demographic factors such as education, gender, and ethnicity. These types of error typically affect the entire sample in the same direction (e.g., all participants decline). Random error in contrast comprises other types of error, including measurement error and regression to the mean and is not necessarily uniform among

participants. Measurement error results from imperfect instruments employed in serial assessments. Because no test of human behaviour is likely to have perfect reliability over time, it is expected that an individual's test score will vary to a degree across different assessment points (Lineweaver & Chelune, 2003). Regression to the mean results from the differences between an individual's true score and their observed score. If baseline scores are below the mean, it is expected that follow-up scores will increase, and vice versa for observed baseline scores that are above the mean (Streiner & Norman, 2008). However, observed scores from tests that are highly reliable are less susceptible to this form of error. Other contributors to random error could include an individual's mood or fatigue level at the time of testing or the examiner's level of alertness. Taken together then, the different effects of error make the assessment and interpretation of meaningful change more complicated. The ultimate goal of reliable change assessment is to determine when significant, meaningful change has occurred independent of any of the above error types (Duff, 2012; Hinton-Bayre, 2011).

Methods for assessing change. The assessment of change is becoming increasingly popular, especially among neuropsychologists who are interested in recovery from injury (e.g., traumatic brain injury) or response to treatment (e.g., cognitive rehabilitation; Hinton-Bayre, 2010). Considerable debate exists among neuropsychologists regarding the most effective method for determining significant change in an individual over time (Hinton-Bayre, 2011), with the most suitable methods accounting for the different error types discussed above (Lineweaver & Chelune, 2003). In spite of several decades of work, a consensus model has not been reached. However,

extensive effort has been put forth within the last decade in an attempt to rigorously compare existing methods in order to work toward that end (e.g., Duff, 2012; Hinton-Bayre, 2011; Hinton-Bayre, 2010; Maassen, Bossema, & Brand, 2009). An overview of the different models to assess for significant change will be briefly reviewed here.

Numerous approaches have been used over the years to calculate change over the course of serial assessments. The earlier, simpler models have tended to be flawed and more recently developed methods are favoured among neuropsychologists interested in change. The simplest of these calculations is a basic discrepancy score, which is computed by subtracting time 1 scores from those achieved at time 2. However, these resulting discrepancy scores then need to be compared with normative data that indicates the observed frequency of these scores (Duff, 2012). Discrepancy tables are not always available for a given test, and if they are available, often they are not relevant to the population of interest. Further, simple discrepancy scores do not account for error in measurement (Streiner & Norman, 2008). A slight improvement in change calculation is the Standard Deviation Index (SDI) which involves obtaining a *z*-score by dividing the simple discrepancy score (i.e., time 2 minus time 1) by the standard deviation (SD) of the sample at time 1. In this case, significant change is detected when *z*-scores are above or below 1.645. The next development in the calculation of change scores was the original Reliable Change Index (RCI) equation which was developed by Jacobson and Truax (1991). Rather than relying exclusively on the SD at time 1, it instead uses the standard error of the difference (SE_{diff}), which is a calculation of the SD of the difference scores. In addition to improving on the estimation of the error term (i.e., using the SE_{diff} rather

than SD_1), this was the first change index to incorporate the reliability of the measure, calculated using a Pearson product-moment correlation of the scores at time 1 with those at time 2. However, like the SDI method, the error term in this original RCI calculation relies only on the SD at time 1. Thus, the implied assumption in this model is that the variance across time points is equal. Accordingly, Iverson and colleagues (2003; see also Iverson, 2001) modified the SE_{diff} aspect of the RCI calculation to include variance in scores at both time points. This model is also frequently attributed to Maassen (2004) (refer to Method section for equation).

Variations of this revised RCI model have been developed to account for practice effects within the sample (e.g., Chelune, Naugle, Lüders, Sedlak, & Awad, 1993). Indeed, the RCI method using SE_{diff} that accounts for practice effects is the preferred equation in a recent comparison of numerous approaches for calculating reliable change (Maassen et al., 2009). In this case, a practice effect value, which is simply the mean discrepancy scores (i.e., time 2 minus time 1), is added to the numerator.

There is considerable debate in the literature regarding the inclusion of practice effects in the calculation of cognitive change scores. This is due to the fact that memory for specific items, developed strategies for problem solving tasks, and experience or comfort gained from repeated testing contribute to better performance at retest (McCaffrey, Duff, & Westervelt, 2000). Thus, changes are attributable to prior exposure to testing materials, rather than true change in performance (Heilbronner et al., 2010). Numerous attempts to control for this have been used, including increasing the test-retest window and using alternate forms. However, a recent meta-analysis by Calamia and

colleagues (2012) summarized 1600 effect sizes and concluded that practice effects persisted in spite of excellent test-retest reliability, use of alternate forms, and large retest windows. They noted that practice effects varied within and across cognitive domains, and reported one result where practice effects on the Wisconsin Card Sorting Task, a novel task indexing mental flexibility, only dropped to acceptable levels (i.e., no longer having an effect) after a 7-year test-retest window. Accordingly, these and other researchers have concluded that practice effects must be accounted for to ensure that Type I and Type II errors are being controlled. That is, detecting significant change that is in fact only attributable to this type of systematic error or failing to account for changes such as cognitive decline because practice effects might make performance look stable or even slightly improved.

Previous work has found that the revised RCI calculation including practice effects is comparable to two standardized regression based (SRB) change formulae in terms of the ability to detect reliable and clinically meaningful change (e.g., Hinton-Bayre, 2010). Further, it has been shown that the error term (i.e., SE_{diff}) in the revised RCI calculation consistently falls between more liberal and conservative estimates of change using SRB calculations which can be at either extreme depending on the particular data being analyzed (e.g., Hinton-Bayre, 2011; Maassen, Bossema, & Brand, 2009). The AACN (Heilbronner et al., 2010) indicated that both RCI and SRB approaches to change are well suited for use in longitudinal studies looking at cognitive functioning in natural disease course. Thus, evidence supports both RCI and SRB, and perhaps indicates a slight preference in favour of RCI models.

The RCI calculations above represent single-point estimates of change. An alternate way of calculating reliable change, while still using the SE_{diff} is through confidence intervals. In this way, a z-score represents the desired level of confidence, which in the change literature is typically 90% or ±1.645. This value is then multiplied by the calculated SE_{diff}, and the individual's obtained score at time 1 is subtracted from or added to that value, as in Equation 1.

$$CI = X_1 \pm (Z \times SE_{diff}) \quad (1)$$

Thus, a 90% confidence interval is obtained and it would be expected that 10% of cases would fall outside of this range by chance alone. For example, in a sample of 100, five individuals would be expected to score above the 1.645 cut-off and five below -1.645 due to chance. Although the merits of point- and interval-based estimates of determining significance continue to be examined, data in peer-reviewed journals are increasingly being required to be presented in confidence interval format (Hinton-Bayre, 2010). Indeed, in the case of estimating reliable change using an interval-based estimate, one would be provided with a range of expected outcome scores, rather than a single value to which actual retest data can be compared. Confidence intervals will be narrower when test-retest reliability of a measure, a component of the SE_{diff} equation, is high and wider intervals attained when an instrument is considered less reliable overtime. In view of these considerations, Equation 1 will be utilized in the present study with SE_{diff} accounting for practice effects (see Method section for further discussion).

Method

Participants

Community-dwelling patients with a DSM-IV diagnosis of schizophrenia or schizoaffective disorder were recruited from different outpatient clinics in the greater Toronto area in order to maximize generalizability of findings. These clinics included Community Schizophrenia Service (St. Joseph's Healthcare Hamilton), Hamilton Program for Schizophrenia, Cleghorn Early Intervention in Psychosis Program (St. Joseph's Healthcare Hamilton), Canadian Mental Health Association (Toronto Branch), and the Challenging Directions program (Whitby Mental Health Centre). Exclusionary criteria included any history of neurological or endocrine disorder, including head trauma, epilepsy, Cushing's disease or thyroid disorder, a diagnosed learning or developmental disability, and current DSM-IV diagnosis of substance abuse. All participants were required to be willing and able to sign informed consent. These criteria yielded 128 patients that completed testing at both time points in the study.

All participants signed informed consent and received remuneration for each completed module. The larger research project was approved by the institutional review board at each research setting and by the research and ethics board at York University.

Measures

Diagnostic and clinical measures. Each participant was administered the Structural Clinical Interview for DSM-IV Axis I Disorder-Patient Version (SCID; First, Spitzer, Gibbon, Williams, 1996). The SCID is a structured interview given by trained clinical researchers, and was used in the present study in order to confirm inclusionary diagnostic criteria (i.e., schizophrenia or schizoaffective disorder). The mood, psychotic, and substance modules were administered for this purpose. Within each module, items

directly related to the DSM-IV-TR criteria are rated as "Absent/False", "Subthreshold", or "Threshold/True", and each disorder (e.g., Major Depressive Disorder) is rated as present or absent, in some cases for both the present month and lifetime occurrence. This was the only measure excluded from the follow-up assessments.

Current clinical symptoms were assessed using the Positive and Negative Syndrome Scale (PANSS; Kay, Opler, & Fiszbein, 2005). The PANSS is also a clinician administered, structured interview that allows for the assessment of positive and negative symptom severity, as well as some indicators of general psychopathology. Specifically, the positive scale assesses symptoms such as delusions, hallucinations, and conceptual disorganization; the negative scale addresses blunted affect, emotional and social withdrawal, and poor rapport; and the general scale considers symptoms such as somatic preoccupation, anxiety and tension, depression, and insight. Each of the 30 items that constitute these three broad scales is scored on a 7-point scale, with detailed rating anchors ranging from the absence of symptoms to extreme psychopathology, corresponding to scores of 1 and 7, respectively. These items are aggregated and converted into standardized domain scores (i.e., t-scores with a mean of 50 and standard deviation of 10), thus allowing for consideration of the generalizability of a given sample to the schizophrenia population. In addition to the syndrome scores (positive, negative, general psychopathology), cluster scores are available in five domains to depict the nature of the patient's psychopathology, including anergia, thought disturbance, activation, paranoid belligerence, and depression.

The PANSS has shown high internal reliability and homogeneity among its items,

with coefficients ranging from .73 to .83 (Kay, Opler, & Fiszbein, 2000). Furthermore, test-retest reliability indexes for unremitted patients ranged from .77 to .89 on the core scales. The positive and negative scales are inversely correlated with each other (e.g., r = -.62, $p < .01$), supporting their mutually exclusive dimensions. Inter-rater reliability has been shown to vary between .83 and .87. An intraclass correlation, including ratings from the author and research assistant, revealed that inter-rater reliability (n = 89) for this study was high (ICC= .95, $p < .001$).

Cognitive measures. A selection of neuropsychological tests that represented the most commonly impaired cognitive domains in schizophrenia was administered to all participants. Specifically, four subtests were selected from the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997), which is the most widely used measure of intellectual functioning in adults. The Vocabulary subtest assesses word knowledge and was used as an index of verbal ability. A test of perceptual reasoning skills (i.e., Matrix Reasoning) was used to index non-verbal analytical skills. Working memory was indexed using the Letter-Number Sequencing subtest which required sequencing of verbally presented alphanumerical strings. The Symbol Search task assessed mental efficiency using a pencil and paper format, thus providing an overall index of information processing speed. Age-corrected standardized scores were used which are based on raw score performance with mean of 10 and standard deviation of three. Higher scores indicate better performance. Data from the WAIS-III standardization sample revealed that estimates of test-retest reliability (n = 394) were acceptable for all measures used, ranging from r = .75 to r = .91, with the vocabulary subtest demonstrating the strongest

reliability over time (M = 34.6 days; Psychological Corporation, 1997).

The California Verbal Learning Test (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000) is a clinician administered task involving verbal presentation of a 16-item word list over five successive trials, with a distracter word list, free and cued recall trials, and a forced recognition trial. Accordingly, the test provides several indexes of verbal memory, including short delay free and cued recall, long delay free and cued recall, and recognition memory, therefore allowing for inferences related to encoding and/or retrieval memory deficits. Estimates of learning and perseveration are also available. Age and sex corrected standardized scores (i.e., t-scores with a mean of 50 and standard deviation of 10 or z-scores with a mean of zero and standard deviation of one) are calculated based on raw score performance. The following scores were included in the present study: 1) trials 1-5, which indexes total immediate recall across five trials of the same list of words; 2) short delay free recall, which is the total number of words recalled from the first list of words after hearing the distractor word list; 3) long delay free recall, which indexes memory for the original list after a delay of 20-25 minutes; and 4) intrusions, which includes any word that was not on the word list being tested. Higher scores indicate better performance (i.e., more words recalled) for all standardized scores with the exception of total intrusions where higher z-scores indicate worse performance.

The alternate form of the CVLT-II was used in the second year of the study, which included 16 different words from four alternative semantic categories. There were no significant differences in word frequency between the two forms of the test (Delis et al., 2000). Estimates of test-retest reliability (n = 288; M = 21 days) in the standardization

study using the combination of standard and alternate forms ranged from .72-.79 on the primary measures, and was somewhat lower for the process measures (e.g., total intrusions r = .55). A more recent reliability study (n = 115; M = 29 days) demonstrated similar, though slightly smaller test-retest reliability coefficients for the standard-alternate form primary measures ranging from .61 to.73 (note: total intrusions r = .57) and also showed no evidence of practice effects on any of these measures (Woods, Delis, Scott, Kramer, & Holdnack, 2006).

Oral fluency was assessed using the phonemic and semantic trials of the Controlled Oral Word Association Test (i.e., as adapted by Benton, Hamsher, &Sivan, 1994). Alternate stimuli were used at each time point, with FAS and animal names representing the phonemic and semantic trials at baseline, and CFL and fruits/vegetables at retest. Raw scores were used at both time points where higher scores indicate better performance. The different forms of the instrument have been shown to be strongly correlated among healthy adults (r = .92) and in clinical samples (r = .87 to r = .94; Lacy et al., 1996). Test-retest reliability for the measure has been indexed at values ranging from r = .70 to r = .84 (Dikmen, Heaton, Grant, & Temkin, 1999; Ross, 2003; Ross et al., 2007; Ruff, Light, & Parker, 1996; Tombaugh, Kozak, & Rees, 1999) with a variety of test-retest stimuli combinations (e.g., FAS-FAS, CFL-CFL, FAS-BDT, CFL-PRW).

Functional measure. Real world community functioning was assessed using the Multidimensional Scale of Independent Functioning (MSIF; Jaeger et al., 2003). The MSIF is a structured interview and self-report measure with verification of information provided by history, proxy reports, and informant interviews. Global functioning ratings

are calculated for three environments (i.e., work, education/training/rehabilitation, and residential) and for each of the three domains (i.e., role position, support, and performance). Broadly, the role position domain assesses the actual responsibilities that an individual has in his or her environment. The support rating reflects the amount of support an individual requires in order to perform their specific role responsibilities. Performance refers to an individual's quality, timeliness, and reliability of the specific tasks for which they are responsible in each domain. Thus, this measure assesses six domains of community independence, including role position, support required, and level of performance in work, education, and residential settings. In addition, an overall global independent functioning rating is calculated, which reflects an individual's role functioning in at least one of work, education/training/rehabilitation, and residential environments while correcting for the degree of responsibility, level of support utilized, and actual performance success in those environments. The education domain was not included in these analyses (see Miles et al., 2011 for a discussion).

The developers' validation study revealed that the MSIF had high criterion validity (.78 to .86) relative to the Social Adjustment Scale (SAS-II; Schooler, Hogarty, & Weissman, 1979), which was a widely used measure in treatment trials to assess functional outcome. The MSIF also showed high external validity. For example, work role position ratings were highly correlated with number of hours worked and hourly wages, and work support ratings were highly correlated with amount of professional support received through transitional or supported employment programs. Inter-rater reliability was assessed by the developers using intraclass correlations, which yielded

reliability coefficients ranging from .72 to 1.00. In comparison, inter-rater reliability for this study (n = 101) was high ($r = .913, p < .001$), with coefficients ranging from .75 to .92, indicating comparable agreement among our raters. The raters included the author of this manuscript and the research assistant for the project.

Statistical Analyses

The absolute consistency of the scores for each participant or stability of performance over one year was calculated using intraclass correlations (ICC) for measures of cognitive ability, clinical symptom severity, and community independence. ICCs are a preferred method of analysis to Pearson product-moment correlations when assessing the consistency of multiple observations of the same variable rather than the relationship between two different variables (Streiner & Norman, 2008). The ICC calculations in the present study were conducted using a 2,1 model, which refers to a two-way, random model with a single measure ICC coefficient (Shrout & Fleiss, 1979).

Two-way rather than one-way models are utilized when the study design is one assessing a form of test-retest reliability (Weir, 2005). The model is considered to have random versus fixed effects given that the design is intended to generalize beyond the confines of the study, rather than the factor of interest being related only to the study itself (e.g., designing a new measure and determining it's reliability prior to its inclusion in a larger study). The random effects model accounts for both systematic and random error, whereas a fixed effects model (i.e., ICC 3,1) considers only random error, which may result in an inflated reliability coefficient (Weir, 2005). This is important because systematic error provides an estimate of bias which can result from factors like research

participation fatigue or improvement in scores due to learning (i.e., practice effects). Thus, random effects models provide more conservative estimates of reliability (Weir, 2005). A single measure versus average measures ICC coefficient was used because only a single score was collected for each participant on each variable considered. In contrast, the average measures coefficient might be used if a researcher used the mean height of 10 jumping trials as the estimate of baseline jumping height or if multiple raters provided a score for each individual and each rater was used in the test-retest reliability calculation (McGraw & Wong, 1996).

ICC values are unitless and range from 0 to 1.0, where an ICC of 0 would indicate no reliability and an ICC 1.0 denotes perfect reliability across a minimum of two observations. Given that ICCs are an indicator of strength of association, which is related to effect size measures (Tabachnick & Fidell, 2007), interpretation of the magnitude of the relationship is based on Cohen's (1988) descriptions where .10-.29 is considered small or weak, 30-.49 is considered medium or moderate, and .50-1.00 is considered large or strong. However, when absolute agreement is important, for example in the case of interrater or test-retest reliability, a correlation coefficient of .70 or higher is typically desirable (Cohen, 2001).

Paired samples *t*-tests were used to provide corroborative evidence of whether performance was consistent across the two testing points.

In order to more closely examine the question of significant change in performance over one year among patients with schizophrenia, change scores were determined using a reliable change index (RCI) confidence interval calculation.

Accordingly, group mean and standard deviation estimates were calculated for each measure of interest at initial and follow-up testing. Correlations between performance at the two time points were obtained (i.e., Pearson product-moment correlation, *r*) to provide the index of test-retest reliability. Pearson's *r* closely approximates Shrout and Fleiss' ICC 3,1 or a two-way, *fixed* model and is considered appropriate when assessing test-retest reliability where the results will not generalize beyond the reliability of that particular calculation (McGraw & Wong, 1996). In the case of reliable change calculations, the use of Pearson's *r* within the calculation is the standard practice. Standard error of the difference (SE_{diff}) was then calculated for each measure using Equation 2 (Iverson, Lovell, & Collins, 2003; Maassen, 2004). SD_1 is the standard deviation at time 1, SD_2 the standard deviation at time 2, and r_{12} the correlation between scores on any given measure at time 1 and time 2.

$$SE_{diff} = \sqrt{(SD_1^2 + SD_2^2)(1 - r_{12})} \quad (2)$$

The SE_{diff} score was multiplied by a factor consistent with 90% confidence intervals (i.e., ±1.645) to determine the range of scores that would be expected if no actual change in performance had occurred over the one year testing period (see Equation 3). Accordingly, this provides the "theoretical distribution of a respondent's score distribution under the null hypothesis that no true change occurred" (Maassen, 2004, p.890). As described earlier, 90% confidence intervals were chosen in order to provide a more conservative estimate of change and to keep with previous research in this area (e.g., Harvey et al., 2005). Further, practice effects were included for all cognitive tests to account for systematic error in the assessment of significant change. In this case, X_1

refers to an individual's score at time 1, M_D is the practice effect, which is the average of time 2 scores minus time 1 scores, and SE_{diff} is the score obtained from Equation 2.

$$90\%\ CI = (X_1 + M_D) \pm (1.645 \times SE_{diff}) \quad (3)$$

According to Equation 3, 90% of time 2 scores by chance alone should fall within the confidence interval obtained using this calculation. Actual scores obtained above this interval would be expected to occur less than 5% of the time and thus are indicative of statistically significant improvement in performance. Similarly, actual scores below the confidence interval represent statistically significant decline in performance. It is important to note that on some measures, a negative change signifies improved functioning, such as on estimates of symptom severity where high scores equate with greater psychopathology. Where relevant, these distinctions are clearly articulated in the results section.

This procedure was conducted for each of the aforementioned symptom severity, cognitive performance, and real world outcome variables. Results from these calculations allowed for the development of three groups of patients: those who demonstrated statistically significant improvement in performance, those whose performance significantly declined over the year, and patients who did not demonstrate significant change. Of the patients who significantly improved or deteriorated over the course of the study, it was then evaluated whether they also showed significant increases or decreases in their real world independent functioning.

The relationships between symptom-outcome change and cognition-outcome change were further explored using Pearson product-moment correlations. That is,

correlational analyses were conducted on each of the symptom and cognitive change variables with each indicator of real world functional independence. Multiple regressions were then conducted where the outcome variables were regressed onto the symptom and cognitive change variables separately to control for covariance among the measures. Lastly, multiple regression models were examined where all cognitive and symptom change variables were entered simultaneously into the model, as well as using separate blocks of symptoms and cognitive change variables to determine if adding one set of predictors (e.g., change in symptom scores) improves the prediction of outcome relative to the cognitive predictors alone.

Results

Demographic and Clinical Characteristics

One hundred and twenty eight schizophrenia patients completed testing at both time points. The mean test-retest window across all assessments was 10.30 months (SD=1.68), where the delay between the functional outcome assessment was shortest (M=10.21, SD=1.60) and cognitive battery the longest (M=10.43, SD=1.78).

The schizophrenia patients ranged in age from 21 to 65 years, with a mean age of 41.45 (SD=9.0). The group was predominantly male and Canadian-born with English as their first language learned. Level of engagement with the workforce was varied, with 52% of patients employed in full-time or part-time positions and the remainder in volunteer positions, enrolled in school, or unemployed. The majority (71%) of patients completed high school or equivalency. Demographic characteristics are presented in Table 1.

The mean age of illness-onset, comprising either first psychiatric hospital admission or first psychiatric contact and report of psychotic symptoms, was 20.90 years (SD = 5.45). All patients, with the exception of one were being treated with antipsychotic medication with the majority (85%) taking atypical neuroleptics. Data from the PANSS collected at baseline indicated that patients in this study were experiencing average levels of positive (M=50.10, SD=9.46), negative (M=46.11, SD=9.32), and general (M=52.41, SD=8.75) symptoms (Opler et al., 1999), and that symptom severity ranged from levels much below average up to levels very much above average. All cluster scores of the PANSS, which assess more specific symptoms within the disorder (e.g., anergia, thought disturbance, etc.), were also average and consistent with the normative sample, with the exception of the depression subscale score (M=57.56, SD=9.92) which was mildly elevated (i.e., t-scores between 56-61 are considered slightly above average; Opler et al., 1999). The cluster scores also spanned the range of severity levels from much below average to very much above average. Patients showed a similar pattern of psychopathology at follow-up on the global and subscale indicators of symptom severity on the PANSS (see Table 2).

Cognitive Performance

Performance on cognitive measures of interest is reported in Table 3. Based on age corrected scaled scores, schizophrenia patients scored in the average range relative to their peers on the vocabulary, matrix reasoning, and letter-number sequencing subtests of the WAIS-III at baseline. Low average performance was observed on symbol search, the indicator of processing speed. Performance ranged from extremely low to very superior

on each of the WAIS-III subtests. All indicators of memory, including learning, immediate and delayed free recall, and source memory difficulties were estimated to be low average, with performance ranging from severely impaired at the lower end and from high average to very superior at the upper end. Oral fluency indicators ranged from low average (phonemic) to average (semantic) ranges with performance also spanning all ability ranges. At follow-up, the schizophrenia patients demonstrated no change in ability range on the following tests: vocabulary, matrix reasoning, letter-number sequencing, learning, immediate and delayed recall. In contrast, patients improved from the low-average to average range on symbol search, source memory, and phonemic oral fluency. However, only the change in symbol search performance was statistically significant ($t_{127} = -4.79, p < .0001$).

Real World Outcome Status

Baseline and follow-up data from the MSIF are presented in Table 4. The mean score on the overall global indicator of functional independence was in the moderately disabled range, indicating adequate performance in non-mainstream, specialized environments or difficulty in mainstream environments in spite of regular support or assistance. Role position and global work environment ratings were also reflective of moderate disability, indicating moderate role expectations and responsibilities and moderate overall disability in terms of vocational independence. Patients scored in the modest but definite disability range on the global support indicator of independence suggesting either small to moderate support in all environments or significant supports in one or two. Global residential environment ratings were considered somewhat disabled,

indicating overall performance to be adequate given regular support or impaired without those supports in mainstream environments. Mean ratings on the global performance domain indicated that deficits are mild or occasional but noticeable. Performance on all real world outcome domains spanned the range of the scale from normalized functional independence to "total" disability. Functional independence status was statistically equivalent over the course of one year.

Stability Data

Stability of clinical symptoms. Results from the assessment of stability of clinical symptoms are presented in Table 2. On the global indicators of clinical symptom severity (i.e., positive, negative, and general scales), ICCs ranged from .44 to .69, which is consistent with moderate to strong relationships. The general ($t_{127} = 2.82, p = .006$) and positive ($t_{127} = 2.96, p = .004$) scales were significantly different between baseline and follow-up assessments, with patients demonstrating reduced psychopathology in these domains at follow-up. There was no significant difference in negative symptoms across the two testing points. The remaining subscales had ICCs ranging from .39 to .76, which also indicated modest to strong reliability. Paired samples *t*-tests showed that all subscales, with the exception of thought disturbance ($t_{127} = 2.10, p = .038$), were statistically indistinguishable and thus stable across the baseline and follow-up assessments. If a Bonferroni correction is employed to correct for the multiple comparisons ($p < .006$; [.05/8]), only the general symptom change remains significant.

Reliable change index (RCI) calculations were used to determine the stability and frequency of change in symptom severity of the course of the study. Results of these

analyses are presented in Figure 1. On average, 8 patients (6.3% of the sample) showed reduced symptom severity, ranging from 4 patients (3.1%) on the overall symptom severity scale up to 11 (8.6%) patients on the positive, general, and thought disturbance scales. In contrast, an average of 5 patients (4.2%) demonstrated increased symptom severity, ranging from 4 (3.1%) on the positive symptom scale up to 9 (7.0%) on the activation subscale. The remaining 89% of patients demonstrated stable psychopathology over one year. Overall, total symptom severity was considered the most stable (93%) domain and each of general psychopathology, thought disturbance, and activation were the least stable (87.5%).

Stability of cognitive functioning. Table 3 contains data on the stability of cognitive performance among schizophrenia patients in this sample. On the subtests of the WAIS-III, ICCs ranged from .72 on letter-number sequencing to .95 on vocabulary, indicating acceptable consistency over the year on these tests. Paired samples *t*-tests revealed no statistically significant differences over the year on the vocabulary measure. However, differences were noted in non-verbal analytical skills ($t_{127} = -3.00, p = .003$), working memory ($t_{127} = 2.13, p = .035$), and processing speed ($t_{127} = -4.79, p < .0001$), with both matrix reasoning and processing speed showing slight increases in performance and letter number sequencing declining over the year. Stability estimates from the learning and memory measure (i.e., CVLT-II) ranged from .59 on total intrusions to .75 on the learning indicator. Patients showed significant improvements in learning (i.e., trials 1-5) over the year ($t_{127} = -4.02, p < .0001$), but no differences in the remaining indicators of memory ability. ICCs on the oral fluency measure were .50 for semantic

fluency, which declined over the year ($t_{127} = 3.17, p = .002$), and .79 on phonemic fluency, which remained stable. If a Bonferroni correction is applied ($p<.005$; [05/10]), all but the letter-number change remain significant.

A visual representation of the stability and change frequency data is presented in Figure 2. On average, cognitive functioning was highly stable over the course of one year with 90% of patients (n=115) demonstrating statistically indistinguishable performance between baseline and follow-up assessments. Processing speed performance was the most stable (93.8% of patients showed no change) and verbal knowledge the least stable (84.4%). In terms of those who demonstrated significant change, a mean of 6 (4.8%) patients declined over the year, which ranged from 4 (3.1%) patients on processing speed and short delay free recall up to 11 (8.6%) on a measure of working memory. In comparison, 7 patients on average (5.6% of the sample) showed significant improvements in their cognitive functioning, ranging from 3 (2.3%) on the semantic fluency task to 11 (8.6%) on the verbal knowledge task.

Stability of real world outcome. Paired samples *t*-tests revealed no significant differences between baseline and follow-up testing on any indicator of functional independence. Stability estimates (i.e., ICC) ranged from .31 on the global performance domain to .63 on the global support indicator of independence suggesting modest consistency among the two sets of scores. These data are given in Table 4.

Results from the RCI calculations of stability and frequency of change among schizophrenia patients on real world outcome are presented in Figure 3. On average, 10 patients (7.5% of sample) demonstrated improved community independence over the

year, with the lowest frequency of improvement noted in the global performance domain (5 patients, 3.9%) and largest gains made in the global support domain (17 patients, 13.3%). Deterioration in functioning or reduced independence was found on average in 8 patients (6.3 %) and ranged from 4 patients (3.1%) deteriorating on the performance domain and 13 (10.2%) becoming less independent in the residential domain. Overall, an average of 86.2% of the patients showed stable community outcome over the year, with the global performance domain (93%) showing the highest level of stability and the global support domain the lowest (82%)

Change data

Relationship between symptoms and real world outcome. Results from these analyses are presented in Table 5. Taken together, of the patients that demonstrated statistically significant improvement in their symptom status (i.e., symptom scores decreased over time), an average of 14.7% (range 0-50%) also showed significant improvements in their real world functioning. Of the patients who demonstrated significant increases in their symptom scores (i.e., symptoms became more severe), an average of 6% (range 0-40%) also showed significant deterioration in their functional status. However, the group of patients who had statistically significant change on both a symptom subscale and at least one community outcome domain represented an average of 0.4% of the total sample (ranging from 0-3.1%).

There were 121 observations of change across all of the symptom severity scales from 62 different patients. Of the 121 observations of symptom change, 59 (48.8%) also showed significant change on an average of 1.89 real world outcome domains.

Pearson product-moment correlations were conducted on the 9 symptom domain change scores and 6 outcome domain change scores, resulting in 54 correlations. Of these, there were 7 small (Cohen, 1988), significant positive correlations ranging from r = .18, p = .034 up to r = .22, p = .012. This indicates that decreases in symptom severity associate with better functional status (recall that lower MSIF scores correspond to more independent functioning or less overall disability). If a Bonferroni correction is applied (p < .0009; [05/54]), all correlations become statistically insignificant.

A multiple regression was performed to observe the combined and unique effects of change in the symptom subscales on change in the outcome variables. Perhaps not surprisingly given the structure of the PANSS (i.e., items load onto multiple scales), there were significant issues with multicollinearity in this model, rendering the model non-interpretable. Two of the change variables (positive and general symptom change) had a variance inflation factor (VIF) above four indicating unacceptable covariance with other variables (Miles & Shelvin, 2001). The remaining VIF and tolerance diagnostic statistics were not ideal for the majority of the remaining symptom change variables. Thus, the regressions were re-run by regressing the functional outcome change scores onto the syndrome (i.e., positive, negative, and general symptoms) and cluster (anergia, thought disturbance, activation, paranoid belligerence, and depression) change scores separately. The total symptom score was not included in either set of analyses due to issues of multicollinearity.

The multiple regression analyses were not significant when outcome change scores were regressed onto the block of syndrome change scores (i.e., positive, negative,

and general psychopathology). There were two trend level significant models when global support ($F = 2.201, p = .058$) and global performance ($F = 2.038, p = .078$) were regressed onto the set of cluster change scores. In the former, the beta coefficient for thought disturbance was significant ($B = .187, t = 1.98, p = .05$) indicating that decreases in the severity of thought disturbance associate with reduced support needs in the community. In the case of the latter, reduced activation ($B = .222, t = 2.46, p = .015$) was related to reduced performance deficits in the community. Diagnostic statistics in these models were near 1.0 and therefore acceptable.

Relationship between cognition and real world outcome. The average percentage of the overall sample that demonstrated significant change on both a cognitive ability test *and* a real world outcome domain was 0.5% and ranged from 0-3.1% (see Table 6). Specifically, of the patients who demonstrated significant improvement in their cognitive performance using RCI calculations, an average of 7.3% (range 0-33%) also demonstrated significant improvement in their functional outcome status. Similarly, of the patients who showed significant deterioration in their cognitive performance, an average of 8.3% (range 0-33%) also demonstrated a decline in at least one real world outcome domain.

Significant improvement or decline in cognitive functioning was shown on 130 separate occasions from 81 different patients. Significant change was also demonstrated in community outcome domains in 74 (57%) of the 130 cases and on an average of 1.85 functional independence domains.

Pearson product-moment correlations were conducted to assess the relationship

between change in cognitive abilities and change in real world outcome, resulting in 60 correlations. There were three significant but small (Cohen, 1988) relationships among the variables. Changes in phonemic fluency and short delay free recall had negative correlations with change in the global performance domain (r = -.225, p = .011; r = -.192, p = .03), suggesting that improvements in fluency and short delay memory are related to better community performance (recall that a decrease in MSIF scores corresponds with less disability). The other significant relationship was a positive correlation between changes in delayed memory and global support on the MSIF (r = .229, p = .009), suggesting that a positive change (i.e., improvement) in long term memory associated with positive change scores (i.e., increased need for assistance) in community support requirements. The remaining 57 correlations were non-significant. However, if a Bonferroni correction is applied to correct for the multiple comparisons (p < .0008; [05/60]), all correlations become non-significant.

Further, results from the multiple regression where each outcome change variable was regressed on the group of cognitive change predictors revealed one significant model for the global performance domain (F = 2.042, p = .035). Significant predictors in this model included short delay free recall (B = -.333, t = -3.05, p = .003), long delay free recall (B = .212, t = 2.02, p = .046), and phonemic fluency (B = -.212, t = -2.41, p = .018), such that decreases (i.e., worse performance) in short delay memory and letter fluency associate with positive change scores on global performance indicating increased disability. In contrast, improved performance on long delay memory correlated with increased disability. Tolerance and VIF diagnostic statistics were near 1.0 and therefore

acceptable for this model.

Combined models. The final set of analyses involved regressing each real world outcome domain onto the collection of symptom and cognitive variables as a whole. This was then repeated using hierarchical regression to determine if adding one set of variables to another significantly improved the model.

The overall model with the cognitive variables and syndrome change scores (i.e., positive, negative, and general symptoms) showed a trend level of significance (F = 1.726, p = .065, R^2 = .164). The significant predictors in this model included short delay free recall (B = -.315, t = -2.84, p = .005) and phonemic fluency (B = -.209, t = -2.35, p = .02) change scores. Thus, declined performance in short delay memory and phonemic fluency associated with increased performance deficits. The model was significant when global support was regressed onto all cognitive variables and cluster change scores (i.e., anergia, thought disturbance, activation, paranoid belligerence, and depression), F = 1.823, p = .04, R^2 = .196. Changes in long delay free recall (B = .239, t = 2.26, p = .026), thought disturbance (B = .202, t = 2.05, p = .042), and depression (B = -.196, t = -2.15, p = .034) were all significant predictors of change in community support requirements, such that improved performance in long delay memory associated with increased support needs, and decreased thought disturbance but increases in depression associated with less support utilized over time.

Three significant models emerged when hierarchical multiple regression analyses were employed. The symptom change variables were entered into the model first, followed by the block of cognitive change variables. Results showed that adding a set of

cognitive predictors to either a set of syndrome or cluster symptom change scores improves the prediction of global performance deficits by 14.3% ($F = 1.954$, $p = .045$, $R^2 = .164$) and 15.5% ($F = 2.09$, $p = .031$, $R^2 = .169$), respectively. There was also one significant hierarchical model when change in global support requirements was regressed first onto the set of cognitive variables and then onto the block of cluster symptom change scores, $F = 2.85$, $p = .019$, $R^2 = .196$. These results demonstrated that adding a set of symptom predictors to cognitive change scores improves the prediction of change in community support needs by 10.2%.

Discussion

This study is the first to provide comparative data on the temporal stability and prevalence of change across three critical features of the schizophrenia illness including symptom severity, cognitive functioning, and community independence. Furthermore, the investigation of change in functional status as predicted by changes in clinical symptoms or cognition was also presented for the first time.

Across the three broad domains of symptoms, cognition, and community outcome, schizophrenia patients demonstrated performance spanning both extremes of each measure. That is, symptom scores ranged from severe psychopathology to asymptomatic presentations, cognitive scores ranged from severe impairment to very superior performance, and real world outcome ranged from normalized independence to complete disability. Thus, the present sample indexed all levels of functioning expected within schizophrenia outpatient populations. On average, patients tended to demonstrate average range symptom and cognitive profiles and moderate disability in the community,

indicating that these data are representative of typical, chronic outpatients with schizophrenia. It is noteworthy, however, that this sample was affiliated with relatively enriched community rehabilitation clinics. Thus, for these patients, "treatment as usual" included access to case management, vocational services, domestic skills training, medication management, and computer tutoring, among other opportunities. Although it is unlikely that all patients accessed all available services, it is likely that their treatment environment was more enriched than some Canadian and many American facilities (e.g., Patterson et al., 2001)

Using change score analyses (i.e., reliable change index; RCI) accounting for practice effects, the data demonstrated that approximately 90% of patients show statistical stability across 10 different cognitive ability domains. These data support previous reports (e.g., Heaton et al., 2001) that cognitive functioning is stable over one year in a sample of chronic, schizophrenia outpatients. The comparatively similar reliability estimates across several domains of community independence (i.e., an average of 86% demonstrate stability) and symptom severity (i.e., an average of 89% have stable symptoms) suggest that these features of schizophrenia are also stable, trait-like dimensions of the disorder in treated patients affiliated with outpatient community rehabilitation services. Unlike previous studies, all features of the illness were assessed across multiple domains, rather than relying on a single aggregate score or single indicator of functioning. This allowed for meaningful investigation and comparison within and across specific cognitive, symptom, and community outcome variables.

Among the cognitive variables, processing speed was the most stable over time,

which is consistent with findings from at least one other study (Harvey et al., 2005). This domain of functioning is considered the most sensitive to neurological insult, including schizophrenia (Dickinson et al., 2007), and has been identified as a prime candidate for intervention (Marder, 2011). Vocabulary was considered the least stable domain, although 84% of the sample demonstrated temporal stability across the two testing points. This may be attributable to the exceptionally high test-retest reliability, which allows for small changes in performance to be detected as statistically significant. The total score from the PANSS was the most stable symptom variable, with the paranoid belligerence subscale being the least likely to change among the domain specific indicators. It is notable, however, that this scale had the smallest test-retest coefficient. Tests with lower reliability require greater changes in order to reach criteria for statistical significant using RCI analyses (Hinton-Bayre, 2010). General psychopathology, thought disturbance, and activation were all among the scales most likely to change, though within each, 88% of patients demonstrated stable scores over the year. The global performance domain had the greatest temporal stability among real world outcome domains, yet also the smallest test-retest coefficient. Further, echoing results from the symptom measure, the global indicator of community status had the highest overall stability prevalence rate.

Taken together, all features of the illness are highly stable over one year among treated outpatients. However, a note of caution is offered to readers when interpreting stability data, as it is highly sensitive to the magnitude of test-retest coefficients; lower reliability estimates require greater change in order to be deemed significant, whereas higher coefficients require only small changes in the variable. Further, aggregate scores

often do not explain the full picture as they may mask domains more or less sensitive to change. Thus, although stability estimates were highly similar within and across different features of the illness, a closer examination of individual domains provided additional information about prevalence and susceptibility to change.

In spite of the considerable stability of these illness features, a small subset of patients demonstrated statistically significant improvement or deterioration in functioning over one year. Indeed, significant improvement and deterioration in symptoms was shown in 6% and 4%, respectively. Cognitive change was evident in 6% of patients who improved over the year and 5% who declined. Statistical improvement in community independence occurred in 7.5% of patients and deterioration was noted in 6%. However, a significant majority of the change observations (75%) did not surpass the percentage of the sample expected to change from chance alone. That is, with a 10% confidence interval for follow-up change scores, it would be expected that 5% of the sample would improve and 5% would decline. The majority of significant improvement and deterioration observations in this study occurred in less than 5% of the sample. Accordingly, these data provide further evidence that significant change in any feature of the schizophrenia illness is relatively uncommon among patients receiving treatment as usual.

Nonetheless, the presence of this subset of patients who showed improvements or declines in performance allows for the examination and comparison of the symptom-outcome and cognition-outcome change relationships. At best, 3% of the sample or 4 patients showed significant change in both a predictor (i.e., symptoms or cognition) and

at least one of the community outcome variables. In many cases, however, no relationship was present. In those who did demonstrate significant improvement or deterioration in their symptom or cognitive status, the concurrent change in community functioning was bidirectional (e.g., improvements in cognition were associated with both improvements and declines in community outcome status). Further, in comparing the prevalence of significant symptom-outcome and cognition-outcome change data as well as hierarchical regressions, there was no convincing evidence of a preferential relationship between cognitive ability and functional outcome change compared to symptom severity and community outcome change.

The weak relationship between changes in cognition specifically and changes in community status have important implications for the movement toward improving real world outcome via cognitively enhancing medications and/or behavioural interventions. On the one hand, the failure to find change relationships may suggest a methodological limitation in that an intervention directed at cognitive improvement is required to adequately examine these change relationships. In contrast, the current results may call into question the appropriateness of intervening if a consistent, significant relationship between cognitive and functional change cannot be demonstrated. These considerations will be examined in further detail.

In contrast to the design of this study, it is possible that the relationship between changes in both cognition and functional outcome must be examined within a treatment context. That is, an active pharmacological or behavioural intervention may be required for any possible cognitive improvements to reach a certain magnitude before functional

changes occur (Matza et al., 2006). Thus, the largely null findings in terms of change-change relationships in the present study may simply be attributable to an insufficient effect size. As an extension of this stance, perhaps the null findings are related to the fact that baseline scores were not low enough to allow for significant improvement over the year. In this view, patients with lower baseline cognitive abilities might be less subject to ceiling effects and therefore have greater room to improve overall. It would be expected that corresponding estimates of baseline community independence would also be lower and have similarly greater room to improve. Although floor and ceiling effects were not evident in the present study, this concept was examined by Heaton and colleagues (2001). Their findings revealed no differences in the capacity for change between schizophrenia patient groups with high or low baseline neuropsychological performance.

In a related vein, it could be argued that a specific threshold of cognitive improvement might be required for significant functional change to occur. Thus, regardless of the magnitude of the change, retest cognitive abilities must be above a certain level for simultaneous meaningful change to occur in real world functioning. This was the case in one study reviewed by Matza and colleagues (2006). Wykes et al (1999) failed to find significant group differences in social functioning after patients and controls received a neurocognitive remediation intervention. However, improved social functioning was noted when cognitive flexibility improved beyond a certain level. This concept was also recently evaluated through the comparison of several models that investigated how cognitive improvements change real world functioning (Wykes et al., 2012). The authors showed promising findings in support of a moderated mediation

model where cognition is thought to drive outcome only after a certain performance level is achieved. However, sample sizes were small and they noted that further evaluation is required. Taken together then, it is conceivable that if cognition is actively targeted and significantly larger gains are identified, then corresponding unidirectional improvement in real world outcome may occur.

On the other side of this argument, the data from the present study can be taken as evidence that in typical, chronic schizophrenia patients receiving treatment as usual, changes in cognitive abilities do not associate with similar changes in community independence. That is, when the data are examined more closely in the few patients who do demonstrate significant cognitive change, the corresponding change in real world outcome is bidirectional. For example, when significant improvements are noted in a given cognitive domain, the corresponding significant change in community outcome is either an improvement or deterioration in their level of independence, without a consistent trend for the preferred improvement-improvement relationship. Accordingly then, one interpretation is that cognitive change over one year is in fact unrelated to functional status change. If these change data are then taken to represent the true state of this relationship in schizophrenia, it begs the question of whether interventions aimed at improving outcome via enhanced cognition warrant continued implementation. Instead, it can be argued that significantly more research is required on the cognition-outcome change relationship in typical patients receiving treatment as usual. If significant cognitive-outcome change relationships are consistently demonstrated in this context, investigations should then move toward examining first behavioural interventions, and

pending clinically significant results, pharmacological interventions.

Taken further, the results call into question whether the contribution of cognitive improvement in predicting functional outcome has been oversold. Although older reviews of the cognition-community outcome literature reported that cognitive abilities account for up to 60% of variance in social skills or community independence (e.g., Green et al., 2000), more recent meta-analyses provide perhaps a sobering correction to that number, with variance accounted for ranging from 4-23% (Fett et al., 2011). Further, from a practical standpoint, it seems doubtful that recalling an additional few words on a list recall task or increasing word production on a fluency test will result in functionally meaningful improvements in one's community. Indeed, Miles (2008) investigated a group of verbally gifted patients with schizophrenia. These patients were statistically indistinguishable from verbally gifted healthy adults in terms of their cognitive and functional skillset profiles. However, the gifted patients were significantly disadvantaged in terms of their community outcome status relative to gifted and non-gifted healthy adults. In view of these results and the lack of a symptom-outcome or cognition-outcome relationship in the present study, other mediators or predictors of community adjustment should be considered to determine where the missing validity lies. Future research may benefit from a closer examination of environmental constraints in order to better understand the key ingredients in accurately predicting community tenure.

Yet, in view of disappointing efforts to improve cognition and outcome with pharmacological approaches, new research is being directed at exploring alternative intervention approaches. For example, investigators from the Cognitive Remediation in

the Schizophrenia Trials Network (CRSTN) have proposed that the enhancing potential of newer medications may be limited by a need to first engage patients in cognitive remediation activities (Keefe et al., 2012). As an analogy, they propose that just as physical exercise is required to obtain the benefits of steroids used to increase muscle mass, so too is "systematic cognitive training to 'exercise' any newly found cognitive potential that [schizophrenia patients] may have acquired from drug treatment" (p.e1). Further, novel approaches to existing behavioural interventions are being considered and implemented in order to explore and maximize the concurrent community independence change potential. For example, Bowie and colleagues (2012) demonstrated that a cognitive remediation intervention resulted in real world outcome improvements only when combined with functional skills training.

Limitations

One limitation from the present study relates to the instruments used to assess both cognitive and functional status. Data collection for this project began prior to the publication of the MATRICS Cognitive Consensus Battery (MCCB; Nuechterlein & Green, 2006), which is now the most widely used cognitive battery in schizophrenia research. Accordingly, while these data provide important benchmarks against which future intervention studies can assess meaningful change, in some cases they may only allow for ability factor comparisons rather than test specific comparisons. Further, the measures used in the present study do not tap into all eight separable ability factors that were identified by the MATRICS initiative (Nuechterlein, Barch, Gold, Goldberg, Green, & Heaton, 2004). These unrepresented ability domains include visual memory and

learning, attention, and social cognition. Accordingly, it is possible that significant cognitive change in these three domains may have associated with corresponding functional change. Interestingly, verbal comprehension was identified as a separable ability factor but not included in the MCCB as it was deemed extremely insensitive to change. In the present study, word knowledge was included as an estimate of crystallized verbal ability and was in fact found to have similar stability and change prevalence rates compared to other cognitive domains. Indeed, future cognitive-outcome change studies may benefit from including an estimate of verbal comprehension in their battery.

Limitations of real world outcome measures have been discussed in detail by Miles and colleagues (2011). However, a brief discussion is warranted here. Real world outcome measures represent a significant improvement over instruments that are performance based and administered in a laboratory setting in that they better assess a patient's true community independence status rather than the practical skills required to complete activities of daily living. However, one potential disadvantage is that these outcome measures, including the MSIF, rely on self-report data. Indeed, these types of instruments have been criticized for low convergence between informant (e.g., case manager, employer, relatives) and self-report data (Bowie et al., 2007) with certain subsets of patients tending to overestimate their functional status (Sabbag et al., 2012). These criticisms are in part due to poor correlations between patient self-reported functioning and performance based estimates of functional capacity (e.g., Sabbag et al., 2011). Yet, in light of previous reports that functional capacity measures may be more closely linked to cognition rather than real world outcome (Heinrichs et al., 2010), these

weak correlations may not be as meaningful as previously claimed. Further, Sabbag and colleagues (2011) argue that perhaps certain areas of functioning are more likely to be self-reported with accuracy than others. Indeed, many indicators of functional status on the MSIF are based on factual information that does not involve a subjective rating of one's quality of performance. For example, in the role position domains, questions involve the number of hours worked per week, rate of pay, living situation (alone, with family members, in a group home), and list of chores for which one is responsible. It is possible, however, that the performance domain might be more susceptible to underestimating actual deficits. For example, questions include: "How has the quality of your work been this month?" (rated as good, fair, or poor), "How often have you missed work this month?" (rated as none, some, a lot), or "Has anyone expressed concern about your work?" (rated as none, some, a lot). In spite of these and other limitations, measures of real world outcome continue to provide important information about true community functioning not otherwise attainable from functional capacity measures. However, the assessment of community independence is complex and continued efforts from national collaborations (e.g., Validation of Everyday Real World Outcomes in Schizophrenia [VALERO]; Leifker et al., 2011) are required to systematically evaluate existing instruments and if needed, to develop better measures.

Generalizability of real world outcome measures may be further limited as they are subject to the jurisdiction of its raters. Thus, the research assistants, study participants, and informant raters may have different expectations of community independence among schizophrenia patients compared with other raters. These

differences might be evidenced within particular regions of a city, urban versus suburban settings, or on provincial or national levels. Further, raters may be influenced by the particular government sponsored social benefits, or lack thereof, afforded to patients in a particular region. For example, all patients in this study were fortunate to be affiliated with outpatient rehabilitation services and a significant majority was receiving Ontario Disability Support Payments (ODSP) from the provincial government. Many also had access to public trustees to manage their finances, resided in subsidized and/or supportive housing, and had the cost of their medications covered. In contrast, the social security disability benefits program in the United States does not yet deem patients with schizophrenia automatically eligible for benefits based on their diagnosis alone (Harvey et al., 2012b). Taken together then, community adjustment is a multifaceted construct that is influenced by available familial and social support, access to paid medication and treatment facilities, access to funds, and a host of other variables. It is unlikely that results from any one study will generalize perfectly given the varied circumstances experienced by schizophrenia patients.

With respect to the participants included in this study, a selection bias may bear on the present results. It is likely that patients who are more clinically stable are at an increased likelihood to volunteer at baseline and that those who remain stable over the testing year are more likely to participate in the follow-up assessments. Further, there may be important differences between those patients who continued their participation at follow-up compared with those who dropped out of the study. From the initial sample of 156 patients with schizophrenia, 82% (n=128) were also tested one year later. These

groups of patients (i.e., those who did and did not complete follow-up testing) were statistically indistinguishable in terms of all demographic, symptom, cognitive, and community outcome variables.

In addition to methodological challenges, studies investigating change in features of the schizophrenia illness continue to face statistical limitations due to their reliance on correlational analyses. In particular, analyses based exclusively on Pearson product-moment correlations speak only to the relative position or rank order of individuals in a study and not to the stability of scores over time (Duff, 2012). This investigation represents an improvement over some existing studies then in terms of its reliance on intraclass correlations (ICC). However, ICCs are also imperfect estimates of reliability as the magnitude of the correlation is determined by between-subject variability. Thus, if subjects differ very little from one another, then ICC values will be small, even when trial-to-trial variability is small (Weir, 2005). Similarly, if the patients differ substantially from each other, the resultant ICCs can be large even if trial-to-trial variability is substantial. Another limitation of ICCs is the lack of agreement on a minimum reliability requirement. Thus, it is difficult to comment on the strength of the reliability coefficients in the present study. A cut off of 0.7 is often cited as the minimum level of agreement. However, whether 0.7 is considered too strict or lenient may depend on the questions being asked or the delay in follow-up testing among other variables. Yet some researchers take a stronger position and argue that it is "not theoretically defensible to set a universal standard for test score reliability" (Charter & Feldt, 2001, p.536). More recent considerations of reliability calculations have concluded that this concept continues to

require further investigation (Hinton-Bayre, 2011). Thus, as the search continues for valid and reliable outcome measures for use in intervention studies, researchers must be cautious in making conclusions about the psychometric properties of instruments based only on reliability statistics.

Reliable change analyses have provided more sophisticated methods for evaluating meaningful change. However, these calculations are plagued by similar issues as the correlational analyses mentioned earlier. This is because RCI equations take into account the reliability of a given measure by using correlational test statistics. Accordingly, the same issues of rank ordering and between subject variance apply. Moreover, even assuming accurate correlation coefficients, the RCI measurement is limited by the reliability of the test. That is, when tests have low reliability across test-retest conditions, error estimates are increased and confidence interval ranges become larger. As a result, significant change is more difficult to detect (Hinton-Bayre, 2010). Further, when making dichotomous decisions from dimensional data (e.g., depressed versus not-depressed from questionnaire data), the cutoff point is always arbitrary even when supported by empirical evidence. Accordingly, although the present study used the convention of 90% confidence intervals or a ±1.654 cut off to determine when significant change had occurred, improvements of 1.59 or declines of -1.48 still provide meaningful information that would not be included in the analyses.

The assessment of significant neuropsychological change is always plagued with the issue of practice effects. Newer models of RCI measurement now account for this factor by including a mean sample change score in the calculation. However, this

improvement in change score calculations does not entirely eliminate these effects. For example, in a comprehensive review of practice effects on neuropsychological tests, a bias was still noted up to 7 years after initial testing (Calamia et al., 2012). Numerous methods have been suggested to help account for these types of lasting effects. For example, the authors suggested that the largest influence from repeated exposure to the same material is between the first and second administration. Thus, one recommendation suggested the inclusion of two baseline assessments within a short interval. In this case, scores from the second assessment would be utilized as the baseline estimate for future comparisons. Placebo groups are frequently incorporated into study designs to provide an independent assessment of practice effects. However, this method assumes that practice effects are equal across groups, which is not always true. Further, with certain patient populations (e.g., schizophrenia), the scientific information gained from use of a placebo where one patient group receives no medication may not be worth the risks (Harvey et al., 2005). Another common approach is the use of alternate test forms. However, the form difficulty is often not equated and due to the novel nature of some tests, a practice effect can still be detected. Although costly and time consuming, a combination of these and other efforts have been recommended as research continues on the effects of serial neuropsychological testing (Calamia et al., 2012).

Final conclusion

This study was the first to examine the temporal stability and comparative change relationships between symptoms and real world outcome as well as cognition and outcome. The data presented here represent significant improvements over previous

studies in that comprehensive evaluations of each feature of the illness were obtained and compared. Further, these data, which were collected in a rather large outpatient sample of chronic schizophrenia patients, provide a benchmark against which future investigations can be compared.

The data largely demonstrate that symptoms, cognition, and community independence are similarly stable over one year, with approximately 90% of the sample demonstrating statistically stable performance in each of the domains. Among those that improved or deteriorated in terms of their symptom or cognitive profile, at most 3% demonstrated concurrent community outcome change. Thus, real world outcome change is largely independent of symptom and cognitive change predictors. These data raise questions about the validity and evidential basis for improving cognition as a vehicle for increased community independence.

There is an immediate need for additional studies to specifically evaluate the cognition-outcome change relationship in typical schizophrenia patients receiving treatment as usual. These studies would benefit from larger sample sizes from multisite national collaborations, which would maximize the potential number of change-change patients. This may allow for more sophisticated analyses of this understudied relationship. It is hoped that specific cognitive domains will be identified that are most consistently associated with functional change and that these domains will then be the target of future pharmacological and behavioural interventions. It will be important for future intervention-based studies to control for treatment adherence as this variable will likely have a significant impact on the stability of clinical, cognitive, and community

outcome status. In the meantime, until the nature of this relationship is better understood, it is recommended that novel approaches to cognitive remediation and functional skills training be rigorously explored before additional resources and efforts are directed at evaluating pharmacological interventions.

References

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Press.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision). Washington, DC: American Psychiatric Press.

Ammari, N., Heinrichs, R.W., & Miles, A.A. (2010). An investigation of 3 neurocognitive subtypes in schizophrenia. *Schizophrenia Research 121(1-3),* 32-38.

Arndt, S., Andreasen, N. C., Flaum, M., Miller, D., & Nopoulos, P. (1995). A longitudinal study of symptom dimensions in schizophrenia: Predictions and patterns of change. *Archives of General Psychiatry, 52*(5), 352-360.

Barnard-Thormpson, K., & Leichner, P. (1997). The future of psychiatric hospitals in Ontario. *Sante Mente Que, 22*(2), 53-70.

Bell, M. D., Bryson, G., & Kaplan, E. (1999). Work rehabilitation in schizophrenia: Cognitive predictors of best and worst performance [Abstract]. *Schizophrenia Research, 36,* 322.

Bell, M. D., Zito, W., Greig, T., & Wexler, B. E. (2008). Neurocognitive enhancement therapy with vocational services: Work outcomes at two-year follow-up. *Schizophrenia Research, 105,* 18-29.

Bellack, A. S., Green, M. F., Cook, J. A., Fenton, W., Harvey, P. D., Heaton, R. K., ... Wykes, T. (2007). Assessment of community functioning in people with schizophrenia and other severe mental illnesses: A white paper based on an

NIMH sponsored workshop. *Schizophrenia Bulletin, 33*(3), 805-822.

Benton, A. L., Hamsher, K. deS., & Sivan, A. B. (1994). *Multilingual Aphasia Examination* (3rd ed.). San Antonio, Texas: Psychological Corporation.

Bowie, C. R., McGurk, S. R., Mausbach, B., Patterson, T. L., & Harvey, P. D. (2012). Combined cognitive remediation and functional skills training for schizophrenia: Effects on cognition, functional competence, and real world behaviour. *American Journal of Psychiatry.* Advance online publication. doi: 10.1176/appi.ajp.2012.11091337

Bowie, C. R., Twamley, E. W., Anderson, H., Halpern, B., Patterson, T. L., & Harvey, P. D. (2007). Self-assessment of functional status in schizophrenia. *Journal of Psychiatric Research, 41*, 1012-1018.

Bozikas, V. P., & Andreou, C. (2011). Longitudinal studies of cognition in first episode psychosis: A systematic review of the literature. *The Australian and New Zealand Journal of Psychiatry, 45*(2), 93-108.

Buchanan, R. W., Hostein, C., & Breier, A. (1994). The comparitive efficacy and long-term effect of clozapine treatment on neuropsychological test performance. *Biological Psychiatry, 36,* 717- 725.

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist, 26*(4), 543-570.

Carter, S. C., Barch, D. M., Buchanan, R. W., Bullmore, E., Krystal, J. H., Cohen, J., ... Heinssen, R. (2008). Identifying cognitive mechanisms targeted for treatment

development in schizophrenia: An overview of the first meeting of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Initiative. *Biological Psychiatry, 64,* 4-10.

Charter, R. A., & Feldt, L. S. (2001). Meaning of reliability in terms of correct and incorrect clinical decisions: The art of decision making is still alive. *Journal of Clinical and Experimental Neuropsychology, 23*, 530-537.

Chelune, G. J., Naugle, R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology, 7,* 41-52.

Cohen, B. (2001). *Explaining psychological statistics* (2nd ed.). New York, NY: John Wiley and Sons, Inc.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The functional significance of social cognition in schizophrenia: A review. *Schizophrenia Bulletin, 32*(S1), S44-S63.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test Second Edition (CVLT-II)*. San Antonio, TS: The Psychological Corporation.

Dickinson, D., Bellack, A. S., & Gold, J. M. (2007). Social/communication skills, cognition and vocational functioning in schizophrenia. *Schizophrenia Bulletin, 33*(5), 1213-1220.

Dickinson, D., Ramsey, M. E., & Gold, J. M. (2007). Overlooking the obvious: A meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Archives of General Psychiatry, 64,* 532-542.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of expanded Halstead-Reitan neuropsychological test battery. *Journal of the International Neuropsychological Society, 5*, 346-356.

Dollfus, S., & Petit, M. (1995). Principal component analyses of PANSS and SANS-SAPS in schizophrenia: Their stability in an acute phase. *European Psychiatry, 10*(2), 97-106.

Dominguez, M. D., Viechtbauer, W., Simons, C. J. P., van Os, J., & Krabbendam, L. (2009). Are psychotic psychopathology and neurocognition orthogonal? A systematic review of their associations. *Psychological Bulletin, 135*(1), 157-171.

Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology, 27*, 248-261.

Eack, S.M., & Newhill, C.E. (2007). Psychiatric symptoms and quality of life in schizophrenia: A meta-analysis. *Schizophrenia Bulletin, 33*, 1225-1237.

Eack, S. M., Pogue-Geile, M. F., Greenwald., Hogarty, S. S., & Keshavan, M. S. (2010). Mechanisms of functional improvement in a 2-year trial of cognitive enhancement therapy for early schizophrenia. *Psychological Medicine, 41*(6), 1253-1261.

Elvevag, B., & Goldberg, T. E. (2000). Cognitive impairment in schizophrenia is the core of the disorder. *Critical Reviews in Neurobiology, 14*(1), 1-21.

Erickson, M., Jaafari, N., & Lysaker, P. (2011). Insight and negative symptoms as predictors of functioning in a work setting in patients with schizophrenia. *Psychiatry Research, 189*, 161-165.

Evans, J. D., Bond, G. R., Meyer, P. S., Kim, H. W., Lysaker, P. H., Gibson, P. J., & Tunis, S. (2004). Cognitive and clinical predictors of success in vocational rehabilitation in schizophrenia. *Schizophrenia Research, 70*, 331-342.

Fennig, S., Bromet, E. J., Galambos, N., & Putnam, K. (1996). Diagnosis and six-month stability of negative symptoms in psychotic disorders. *European Archives of Psychiatry and Clinical Neuroscience, 246*(2), 63-70.

Fett, A. K., Viechtbauer, W., Dominguez, M. D., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 35*(3), 573-588.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute.

Fiszdon, J. M., Choi, J., Goulet, J., & Bell, M. D. (2008). Temporal relationship between change in cognition and change in functioning in schizophrenia. *Schizophrenia Research, 105*, 105-113.

Friedman, J. I., Adler, D. N., Howanitz, E., Harvey, P. D., Brenner, G., Temporini, H., ... Davis, K. L. (2002). A double blind placebo controlled trial of donepezil

adjunctive, treatment to risperidone for the cognitive impairment of schizophrenia. *Biological Psychiatry, 51*,m349-357.

Galletly, C. A., Clark, C. R., McFarlane, A. C., & Weber, D. L. (1997). Relationships between changes in symptom ratings, neurophysiological test performance and quality of life in schizophrenic patients treated with clozapine. *Psychiatry Research, 72,* 161-166.

Goeree, R., Farahati, F., Burke, N., Blackhouse, G., O'Reilly, D., Pyne, J., & Tarride, J. E. (2005). The economic burden of schizophrenia in Canada in 2004. *Current Medical Research and Opinion, 21*(12), 2017-2028.

Goldman, R. S., Tandon, R., Liberzon, I., Goodson, J., & Greden, J. F. (1991). Stability of positive and negative symptom constructs during neuroleptic treatment in schizophrenia. *Psychopathology, 24*(4), 247-252.

Government of Canada (2006). *The Human Face of Mental Health and Mental Illness in Canada 2006.* Ottawa, ON: Minister of Public Works and Government Services Canada.

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry, 153,* 321-330.

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the right stuff? *Schizophrenia Bulletin, 26*(1), 119-136.

Green, M. F., Kern, R. S., & Heaton, R. K. (2004). Longitudinal studies of cognition and functional outcome in schizophrenia: implications for MATRICS. *Schizophrenia*

Research, 72, 41-51.

Gupta, M., Bassett, E., Iftene, F., & Bowie, C. R. (2012). Functional outcomes in schizophrenia: Understanding the competence-performance discrepancy. *Journal of Psychiatric Research, 46,* 205-211.

Harvey, P. D., Bertisch, H., Friedman, J. I., Marcus, S., Parrella., M., White, L., & Davis, K. L. (2003). The course of functional decline in geriatric patients with schizophrenia: Cognitive-functional and clinical symptoms as determinants of change. *American Journal of Geriatric Psychiatry, 11,* 610-619.

Harvey, P. D., Green, M. F., Keefe, R. S. E., & Velligan, D. I. (2004). Cognitive functioning in schizophrenia: A consensus statement on its role in the definition and evaluation of effective treatments for the illness. *Journal of Clinical Psychiatry, 65,* 361-372.

Harvey, P. D., Heaton, R. K., Carpenter, W. T. jr., Green, M. F., Gold, J. M., & Schoebaum, M. (2012a). Diagnosis of schizophrenia: Consistency across information sources and stability of the condition. *Schizophrenia Research.* Advance online publication. dx.doi.org/10.1016/j.schres.2012.03.026

Harvey, P. D., Heaton, R. K., Carpenter, W. T. jr., Green, M. F., Gold, J. M., & Schoebaum, M. (2012b). Functional impairment in people with schizophrenia: Focus on employability and eligibility for disability compensation. *Schizophrenia Research.* Advance online publication. dx.doi.org/10.1016/j.schres.2012.03.025

Harvey, P. D., Palmer, B., Heaton, R., Mohomed, S., Kennedy, J., & Brickman, A. (2005). Stability of cognitive performance in older patients with schizophrenia:

An 8-week test-retest study. *American Journal of Psychiatry, 162,* 110-117.

Harvey, P. D., Parrella, M., White, L., Mohs, R. C., Davidson, M., & Davis, K. L. (1999). Convergence of cognitive and adaptive decline in late-life schizophrenia. *Schizophrenia Research, 35,* 77-84.

Heaton, R. K., Gladsjo, J. A., Palmer, B. W., Kuck, J., Marcotte, T. D., & Jeste, D. V. (2001). Stability and course of neuropsychological deficits in schizophrenia. *Archives of General Psychiatry, 58,* 24-32.

Heilbronner, R. L., Sweet, J. J., Attix, D. K., Krull, K. R., Henry, G. K., & Hart, R. P. (2010). Official position of the American Academy of Clinical Neuropsychology on serial neuropsychological assessments: The utility and challenges of repeat test administrations in clinical and forensic contexts. *The Clinical Neuropsychologist, 24,* 1267-1278.

Heinrichs, D. W., Hanlon, T. E., & Carpenter, W. T. Jr. (1984). The Quality of Life Scale: An instrument for rating the schizophrenia deficit syndrome. *Schizophrenia Bulletin, 10,* 388-398.

Heinrichs, R. W., Ammari, N., Miles, A. A., & McDermid Vaz, S. (2010). Cognitive performance and functional competence as predictors of community independence in schizophrenia. *Schizophrenia Bulletin, 36,* 381-387.

Heinrichs, R. W., Miles, A. A., Smith, D., Zargarian, T, Goldberg, J.O, McDermid Vaz, S., & Ammari, N. (2008). Cognitive, clinical and functional characteristics of verbally superior schizophrenia patients. *Neuropsychology, 22*(3), 321-328.

Heinrichs, R. W., Ruttan, L., Zakzanis, K. K., & Case, D. (1997). Parsing schizophrenia

with neurocognitive tests: Evidence of stability and validity. *Brain and Cognition, 35,* 207-224.

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12,* 426-445.

Helgeland, M. I., & Torgersen, S. (2005). Stability and prediction of schizophrenia from adolescence to adulthood. *European Child and Adolescent Psychiatry, 14*(2), 83-94.

Hinton-Bayre, A. D. (2010). Deriving change statistics from test-retest normative data: Comparison of models and mathematical expressions. *Archives of Clinical Neuropsychology, 25*(3), 244-256.

Hinton-Bayre, A. D. (2011). Specificity of reliable change models and review of the within-subjects standard deviation as an error term. *Archives of Clinical Neuropsychology, 26*(1), 67-75.

Hogarty, G. E., Flesher, S., Ulrich, R., Carter, M., Greenwald, D., Pogue-Geile, M., ... Zoretich, R. (2004). Cognitive enhancement therapy for schizophrenia: Effects of a 2-year randomized trial on cognition and behavior. *Archives of General Psychiatry, 61,* 866-876.

Irani, F., Kalkstein, S., Moberg, E. A., & Moberg, P. J. (2011). Neuropsychological performance in older patients with schizophrenia: A meta-analysis of cross-sectional and longitudinal studies. *Schizophrenia Bulletin, 37*(6), 1318-1326.

Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology, 16*(2), 183-191.

Iverson, G. L., Lovell., M. R., & Collins, M. W. (2003). Interpreting change on impact following sport concussion. *The Clinical Neuropsychologist, 17,* 460-467.

Jacobson, N. S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Jaeger, J., Berns, S. M., & Czobor, P. (2003). The multidimensional scale of independent functioning: A new instrument for measuring functional disability in psychiatric populations. *Schizophrenia Bulletin, 29*(1), 153-167.

Johnstone, E. C., Owens, D. G., Frith, C. D., & Crow, T. J. (1987). The relative stability of positive and negative features in chronic schizophrenia. *British Journal of Psychiatry, 150,* 60-64.

Kay, S. R., & Lindenmayer, J-P. (1991). Stability of psychopathology dimensions in chronic schizophrenia: Response to clozapine treatment. *Comprehensive Psychiatry, 32*(1), 28-35.

Kay, S. R., Opler, L. A., & Fiszbein, A. (2000). *Positive and negative syndrome scale: Manual.* New York, NY: Multi-Health Systems, Inc.

Kay, S. R., Opler, L. A., & Fiszbein, A. (2005). *Positive and negative syndrome scale: Manual.* Toronto, ON: Multi-Health Systems, Inc.

Keefe, R. S., & Fenton, W. S. (2007). How should DSM-V criteria for schizophrenia include cognitive impairment? *Schizophrenia Bulletin, 33*(4), 912-920.

Keefe, R. S. E., Poe, M., Walker, T. M., & Harvey, P. D. (2006). The relationship of the Brief Assessment of Cognition in Schizophrenia to functional capacity and real-

world functional outcome. *Journal of Clinical and Experimental Neuropsychology, 28,* 260-269.

Keefe, R. S. E., Vinogradov, S., Medalia, A., Buckley, P. F., Caroff, S. N., D'Souza, D. C., ... Stroup, T. S. (2012). Feasibility and pilot efficacy results from the multisite Cognitive Remediation in the Schizophrenia Trials Network (CRSTN) randomized control trial. *Journal of Clinical Psychiatry.* Advance online publication. doi:10.4088/JCP.11m07100

Kurtz, M. M. (2005). Neurocognitive impairment across the lifespan in schizophrenia: An update. *Schizophrenia Research, 74,* 15-26.

Kurtz, M. M., Donato, J., & Rose, J. (2011). Crystallized verbal skills in schizophrenia: Relationship to neurocognition, symptoms, and functional status. *Neuropsychology, 25*(6), 784-791.

Kurtz, M. M., Seltzer, J. C., Ferrand, J. L., & Wexler, B. E. (2005). Neurocognitive function in schizophrenia at a 10-year follow-up: a preliminary investigation. *CNS Spectrums, 10,* 277-280.

Kurtz, M. M., Wexler, B. E., Fujimoto, M., Shagan, D. S., & Seltzer, J. C. (2008). Symptoms versus neurocognition as predictors of change in life skills in schizophrenia. *Schizophrenia Research, 102*(1-3), 303-311.

Lacy, M. A., Gore, P. A. jr., Pliskin, N. H., Henry, G. K., Heilbronner, R. L., & Hamer, D. P. (1996). Verbal fluency task equivalence. *The Clinical Neuropsychologist, 10*(3), 305-308.

Lamb, H. R., & Bachrach, L. L. (2001). Some perspectives on deinstitutionalization.

Psychiatric Services, 52, 1039-1045.

Lançon, Auquier, Nayt, G. & Reine, G. (2000). Stability of the five-factor structure of the Positive and Negative Syndrome Scale (PANSS). *Schizophrenia Research, 42,* 231-239.

Leeson, V. C., Robbins, T. W., Matheson, E., Hutton, S. B., Ron, M. A., Barnes, T. R. E., & Joyce, E. M. (2009). Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: Stability over six years and specific associations with medication type and disorganization syndrome. *Biological Psychiatry, 66*(6), 586-593.

Leeson, V. C., Sharma, P., Harrison, M., Ron, M. A., Barnes, T. R. E., & Joyce, E. (2011). IQ trajectory, cognitive reserve, and clinical outcome following a first episode of psychosis: A 3-year longitudinal study. *Schizophrenia Bulletin, 37*(4), 768-777.

Leifker, F. R., Patterson, T. L., Bowie, C. R., Mausbach, B. T., & Harvey, P. D. (2010). Psychometric properties of performance-based measurements of functional capacity: Test–retest reliability, practice effects, and potential sensitivity to change. *Schizophrenia Research, 119*(1-3), 246-252.

Leifker, F. R., Patterson, T. L., Heaton, R., & Harvey, P. D. (2011). Validating measures of real-world outcome: The results of the VALERO expert survey and RAND panel. *Schizophrenia Bulletin, 37*(2), 334-343.

Leung, W. W., Bowie, C. R., & Harvey, P. D. (2008). Functional implications of neuropsychological normality and symptom remission in older outpatients

diagnosed with schizophrenia: A cross-sectional study. *Journal of the International Neuropsychological Society, 14,* 479-488.

Lindenmayer, J. P., Kay, S. R., & Friedman, C. (1986). Negative and positive syndromes after acute phase: A prospective follow up. *Comprehensive Psychiatry, 27,* 276-286.

Lineweaver, T. T., & Chelune, G. J. (2003). Reliable change. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, A. Prifitera, & M. F. Ledbetter (Eds.), *Clinical interpretation of the WAIS-III and WMS-III.* San Diego, California: Academic Press.

Maassen, G. H. (2004). The standard error in the Jacobson and Truax reliable change index: The classical approach to the assessment of reliable change. *Journal of the International Neuropsychological Society, 10,* 888-893.

Maassen, G. H., Bossema, E. R., & Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology, 31,* 339-352.

MacCabe, J. H., Brebion, G., Reichenberg, A., Ganguly, T., McKenna, P. J., Murray, R. M., & David, A. S. (2012). Superior intellectual ability in schizophrenia: Neuropsychological characteristics. *Neuropsychology, 26*(2), 181-190.

Malla, A. K., Norman, R. M., & Williamson, P. (1993). Stability of positive and negative symptoms in schizophrenia. *The Canadian Journal of Psychiatry, 38*(9), 617-621.

Marder, S. R. (2011). Lessons from MATRICS. *Schizophrenia Bulletin, 37*(2), 233-234.

Marder, S. R., & Fenton, W. (2004). Measurement and treatment research to improve

cognition in schizophrenia: NIMH MATRICS initiative to support the development of agents for improving cognition in schizophrenia. *Schizophrenia Research, 72,* 5-9.

Marwaha, S., & Johnson, S. (2004). Schizophrenia and employment: A review. *Social Psychiatry and Psychiatric Epidemiology, 39,* 337-349.

Matza, L. S., Buchanan, R., Purdon, S., Brewster-Jordan, J., Zhao, Y., & Revicki, D. A. (2006). Measuring changes in functional status among patients with schizophrenia: The link with cognitive impairment. *Schizophrenia Bulletin, 32*(4), 666-678.

Mausbach, B. T., Bowie, C. R., Harvey, P. D., Twamley, E. W., Goldman, S. R., Jeste, D. J., & Patterson, T. L. (2008). Usefulness of the UCSD performance-based skills assessment. *Journal of Psychiatric Research, 42*(4), 320-327.

McCaffrey, R. J., Duff, K., & Westervelt, H. J. (2000). *Practitioner's guide to evaluating change with neuropsychological assessment instruments*. New York, NY: Kluwer Academic, Plenum Publishers.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46.

McGurk, S. R., & Mueser, K. T. (2004). Cognitive functioning, symptoms, and work in supported employment: A review and heuristic model. *Schizophrenia Research, 70,* 147-173.

McGurk, S. R., Twamley, E. W., Sitzer, D. I., McHugo, G. J., & Mueser, K. T. (2007). A meta-analysis of cognitive remediation in schizophrenia. *American Journal of*

Psychiatry, 164, 1791-1802.

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology, 23*(3), 315-336.

Miles, A. A. (2008). *Verbally gifted schizophrenia patients and healthy adults: Differences in functional outcome, but not in cognition?* (Master's thesis). Available from ProQuest Dissertations and Theses Database. (MR45958)

Miles, A. A., Heinrichs, R. W., & Ammari, N. (2011). "Real world" functioning in schizophrenia patients and healthy adults: Assessing validity of the Multidimensional Scale of Independent Functioning. *Psychiatry Research, 186,* 23-27.

Miles, J., & Shelvin, M. (2001). *Applying regression & correlation: A guide for students and researchers.* London, England: Sage Publications.

Niekawa, N., Sakuraba, Y., Uto, H., Kumazawa, Y., & Matsuda, O. (2007). Relationship between financial competence and cognitive function in patients with schizophrenia. *Psychiatry and Clinical Neuroscience, 61,* 455-461.

Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., & Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research, 72,* 29-39.

Nuechterlein, K. H., & Green, M. F. (2006). *MATRICS Cognitive Consensus Battery Manual.* Los Angeles, CA: MATRICS Assessment, Inc.

O'Leary, D. S., Flaum, M., Kesler, M. L., Flashman, L. A., Arndt, S., & Andreasen, N.

(2000). Cognitive correlates of the negative, disorganized, and psychotic symptom dimensions of schizophrenia. *Journal of Neuropsychiatry & Clinical Neurosciences, 12*(1), 4-15.

Opler, L. A., Kay, S. R., Lindenmayer, J. P., & Fiszbein, A. (1999). *Structured Clinical Interview: The Positive and Negative Syndrome Scale (SCI-PANSS).* North Tonawanda, NY: Multi-Health Systems Inc.

Park, D. Y., Choi, K. S., Lee, D., & Hong, K. S. (2004). Stability of the diagnosis of deficit syndrome in schizophrenia: A 5-year follow-up study. *Journal of the Korean Neuropsychiatric Association, 43*(3), 296-302.

Patterson, T. L., Goldman, S., McKibbon, C. L., Hughs T., & Jeste, D. V. (2001). UCSD performance based skills assessment: Development of a new measure of everyday functioning for severely mentally ill adults. *Schizophrenia Bulletin, 27*, 235-245.

Psychological Corporation. (1997). *WAIS-III-WMS-III technical manual.* San Antonio, Texas: The Psychological Corporation, Harcourt Brace & Company.

Rabinowitz, J., Levine, S. Z., Garibaldi, G., Bugarski-Kirola, D., Berardo, C. G., & Kapur, S. (2012). Negative symptoms have greater impact on functioning than positive symptoms in schizophrenia: Analysis of CATIE data. *Schizophrenia Research, 137*, 147-150.

Reichenberg, A., Rieckmann, N., & Harvey, P. D. (2005). Stability in schizophrenia symptoms over time: Findings from the Mount Sinai Pilgrim Psychiatric Center Longitudinal Study. *Journal of Abnormal Psychology, 114*(3), 363-372.

Rey, E-R., Bailer, J., Brauer, W., Handel, M., Laubenstein, D., & Stein, A. (1994).

Stability trends and longitudinal correlations of negative and positive syndromes within a three-year follow-up of initially hospitalized schizophrenics. *Acta Psychiatrica Scandinavica, 90*, 405-412.

Ross, T. P. (2003). The reliability of cluster and switch scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology, 18*, 153-164.

Ross, T. P., Calhoun, E., Cox, T., Wenner, C., Kono, W., & Pleasant, M. (2007). The reliability and validity of qualitative scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology, 22*, 475-488.

Ruff, R. M., Light, R. H., & Parker, S. B. (1996). Benton Controlled Oral Word Association Test: Reliability and updated norms. *Archives of Clinical Neuropsychology, 11*(4), 329-338.

Sabbag, S., Twamley, E. W., Vella, L., Heaton, R. K., Patterson, T. L., & Harvey, P. D. (2011). Assessing everyday functioning in schizophrenia: Not all informants seem equally informative. *Schizophrenia Research, 131,* 250-255.

Sabbag, S., Twamley, E. W., Vella, L., Heaton, R. K., Patterson, T. L., & Harvey, P. D. (2012). Predictors of the accuracy of self assessment of everyday functioning in people with schizophrenia. *Schizophrenia Research, 1-3,* 190-195.

Sartorius, N. (2002). Iatrogenic stigma of mental illness: Begins with behaviour and attitudes of medical professionals, especially psychiatrists. *British Medical Journal, 22,* 1470-1471.

Schooler, N. R., Hogarty, G. E., & Weissman, M. M. (1979). Social adjustment scale II (SAS II). In W. A. Hargreaes, C. C. Attkisson, & J. E. Sorenson (Eds.), *Resource*

materials for community mental health evaluators (pp. 290-330). Washington, DC: U.S. Department of Health, Education, and Welfare.

Schutt, R. K., Seidman, L. J., Caplan, B., Martsinkiv, A., & Goldfinger, S. M. (2007). The role of neurocognition and social context in predicting community functioning among formerly homeless seriously mentally ill persons. *Schizophrenia Bulletin, 33*(6), 1388-1396.

Seidman, L. J., Giuliano, A. J., Meyer, E. C., Addington, J., Cadenhead, K. S., Cannon, T. D., ... Cornblatt, B. A. (2010). Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. *Archives of General Psychiatry, 67*(6), 578-588.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 36*, 420-428.

Spaulding, W. D., Fleming, S. K., Reed, D., Sullivan, M., Storzbach, D., & Lam, M. (1999). Cognitive functioning in schizophrenia: Implications for psychiatric rehabilitation. *Schizophrenia Bulletin, 25,* 275-289.

Sponheim, S. R., Jung, R. E., Seidman, L. J., Mesholam-Gately, R. I., Manoach, D. S., O'Leary, D. S., ... Schulz, S. C. (2010). Cognitive deficits in recent-onset and chronic schizophrenia. *Journal of Psychiatric Research, 44*(7), 421-428.

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use* (4th ed.). New York, NY: Oxford University Press.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MS: Pearson Education, Inc.

Tandon, R., Nasrallah, H. A., & Keshavan, M. S. (2010). Schizophrenia, "Just the Facts" 5. Treatment and prevention past, present, and future. *Schizophrenia Research, 122,* 1-23.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*(2), 167-177.

Tsang, H. W. H., Leung, A. Y., Chung, R. C. K., Bell, M., & Cheung, W. M. (2010). Review on vocational predictors: A systematic review of predictors of vocational outcomes among individuals with schizophrenia: an update since 1998. *The Australian and New Zealand Journal of Psychiatry, 44*, 495-504.

Vella, L., Burton, C., Harvey, P. D., Heaton, R., Patterson, T., & Twamley, E. W. (2011). The MATRICS Consensus Cognitive Battery: Relationships with performance-based measures of functional capacity. [Abstract]. *Schizophrenia Bulletin, 37*(supp 1), 255.

Velligan, D. I., Prihoda, T. J., Sui, D., Ritch, J. L., Maples, N., & Miller, A. L. (2003). The effectiveness of quetiapine versus conventional antipsychotics in improving cognitive and functional outcomes in standard treatment settings. *Journal of Clinical Psychiatry, 64,* 524-531.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition: WAIS-III.* San Antonio, TS: The Psychological Corporation.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research, 19*(1),

231-240.

Wexler, B. E., & Bell, M. D. (2005). Cognitive remediation and cognitive rehabilitation for schizophrenia. *Schizophrenia Bulletin, 31*(4), 931-941.

Wilk, C. M., Gold, J. M., McMahon, R. P., Humber, K., Iannone, V. N., & Buchanan, R. W. (2005). No, it is not possible to be schizophrenic yet neuropsychologically normal. *Neuropsychology, 19*, 778-786.

Wolter, A., Preuss, U. W., Krischke, N. R., Wong, W. M., & Zimmermann, J. (2010). Remission, prediction and stability of symptoms in schizophrenia: A naturalistic 12-month follow-up study. *International Journal of Psychiatry in Clinical Practice, 14*, 160-167.

Woods, S. P., Delis, D. C., Scott, J. C., Kramer, J. H., & Holdnack, J. A. (2006). The California Verbal Learning Test – second edition: Test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Archives of Clinical Neuropsychology, 21*, 413-420.

World Health Organization. (1993). *The ICD-10 classification of mental and behavioral disorders*. Geneva, Switzerland: World Health Organization.

World Psychiatric Association. (2001). *The WPA global programme to reduce the stigma and discrimination because of schizophrenia – an interim report 2001*. Geneva, Switzerland: World Psychiatric Association.

Wykes, T., Reeder, C., Corner, J., Williams, C., & Everitt, B. (1999). The effects of neurocognitive remediation on executive processing in patients with schizophrenia. *Schizophrenia Bulletin, 25*, 291-307.

Wykes, T., Reeder, C., Landau, S., Everitt, B., Knapp, M., Patel, A., & Romeo, R. (2007). Cognitive remediation therapy in schizophrenia: Randomised controlled trial. *British Journal of Psychiatry, 190*, 421-427.

Wykes, T., Reeder, C., Huddy, V., Taylor, R., Wood, H., Ghirasim, N., ... Landau, S. (2012). Developing models of how cognitive improvements change functioning: Mediation, moderation and moderated mediation. *Schizophrenia Research, 1*, 88-93.

Table 1

Demographic Characteristics of Schizophrenia Patients

Variable[a]	
Age, years[b]	41.45 (9.0)
High school graduate	91 (71%)
Sex (males)	82 (64%)
Employment status	
Full-time	9 (7%)
Part-time	58 (45%)
Volunteer	14 (11%)
Unemployed	44 (34%)
Student	3 (2%)
First language English	107 (84%)

Note. Sample size is 128 and data is based on information collected at baseline.
[a]data presented as raw number of patients (percentage of sample) unless otherwise stated. [b]data presented as mean, standard deviation.

Table 2

Clinical Characteristics of Schizophrenia Patients

Variable	Time 1	Time 2	SE_{diff}	ICC	t[b]	*p*[c]
PANSS[a]						
General	52.41 (8.75)	50.35 (7.75)	8.24	.50	2.82	.006
Positive	50.10 (9.46)	48.33 (7.76)	6.62	.69	2.96	.004
Negative	46.11 (9.32)	46.41 (8.01)	9.16	.44	-.36	n.s.
Total	49.81 (8.83)	49.63 (9.12)	10.34	.34	.21	n.s.
Anergia	50.90 (11.01)	50.13 (9.67)	10.37	.49	.84	n.s.
Thought Disturbance	48.50 (9.67)	47.35 (8.31)	6.09	.76	2.10	.038
Activation	51.45 (9.16)	51.03 (10.18)	10.22	.44	.46	n.s.
Paranoid Belligerence	48.20 (8.17)	48.02 (8.04)	8.96	.39	.23	n.s.
Depression	57.56 (9.92)	56.75 (9.66)	9.66	.51	.95	n.s.

Table 2 continued

Variable	Time 1	Time 2
Diagnosis		
Schizophrenia	78 (61%)	
Schizoaffective	50 (39%)	
Medication		
Typical Antipsychotic	18 (14%)	
Atypical Antipsychotic	108 (84%)	
None[d]	1 (1%)	

Note. Sample size is 128. PANSS = Positive and Negative Syndrome Scale; SE_{diff} = standard error of the difference; ICC = intraclass correlation coefficient.
[a]PANSS mean subscale scores are reported as *t*-scores (M=50, SD=10). [b]paired samples *t*-tests where df = 127. [c]n.s. indicates non-significant results where $p > .05$. [d]n=127.

Table 3

Cognitive Performance among Schizophrenia Patients

Variable	Time 1	Time 2	M_D[e]	SE_{diff}	ICC	t[f]	*p*[g]
WAIS-III[a]							
Vocabulary	9.79 (3.49)	9.77 (3.78)	-0.02	1.15	.95	0.14	n.s.
Matrix Reasoning	9.23 (3.34)	9.73 (3.23)	0.5	1.88	.83	-3.00	.003
Letter Number Sequencing	8.56 (3.16)	8.15 (2.65)	-0.41	2.16	.72	2.13	.035
Symbol Search	7.48 (3.00)	8.38 (2.93)	0.9	2.14	.74	-4.79	< .0001
CVLT-II							
Trials 1-5[b]	41.16 (12.08)	44.23 (12.43)	3.07	8.65	.75	-4.02	< .0001
Short Delay Free Recall[c]	-0.75 (1.08)	-0.77 (1.28)	-0.03	1.03	.63	-0.30	n.s.
Long Delay Free Recall[c]	-0.88 (1.30)	-0.88 (1.24)	0	1.04	.66	0	n.s.
Intrusions[c]	0.58 (1.52)	0.52 (1.54)	-0.06	1.39	.59	0.51	n.s.
COWAT[d]							
Phonemic Fluency	34.62 (11.58)	35.25 (11.95)	0.63	7.66	.79	-0.92	n.s.
Semantic Fluency	18.28 (5.46)	16.83 (4.87)	-1.45	5.17	.50	3.17	.002

Table 3 continued

Note. Sample size is 128. WAIS-III = Wechsler Adult Intelligence Scale 3rd edition; COWAT= Controlled Oral Word Association Test; CVLT-II = California Verbal Learning Test 2nd edition; Trials 1-5 = t score based on sum of trials one through five; SE_{diff} = standard error of the difference; ICC = intraclass correlation coefficient.
[a]scaled scores (M=10, SD=3). [b]*t*-score (M=50, SD=10). [c]*z*-scores (M=0, SD=1). [d]raw scores with no upper limit. [e]practice effect calculated as the mean difference scores. [e]paired samples *t*-tests where df = 127; [f]n.s. indicates non-significant results where $p > .05$.

Table 4

Real World Outcome Status among Schizophrenia Patients

Variable	Time 1	Time 2	SE_{diff}	ICC	t[a]	p[b]
MSIF[c]						
Work	4.97 (1.85)	4.90 (1.71)	1.77	.51	.45	n.s.
Residential	3.77 (1.38)	3.93 (1.22)	1.14	.61	-1.54	n.s.
Role Position	4.68 (1.24)	4.55 (1.27)	1.14	.59	1.24	n.s.
Support	3.78 (1.40)	3.61 (1.29)	1.15	.63	1.68	n.s.
Performance	2.91 (1.19)	2.84 (1.28)	1.45	.31	.55	n.s.
Global	4.38 (1.10)	4.37 (1.03)	1.00	.56	.18	n.s.

Note. Sample size is 128. SE_{diff} = standard error of the difference; ICC = intraclass correlation coefficient.
[a]paired samples *t*-tests where df = 127; [b] = n.s. indicates non-significant results where $p > .05$; [c]scores range from 1-7.

Table 5

Changes in Symptom Severity and Real World Outcome

Variable		Positive		Negative		Total		General	
n (%)		Pos[c]	Neg[d]	Pos	Neg	Pos	Neg	Pos	Neg
Work	Pos[a]	1 (0.8)	0 (0)	3 (2.3)	0 (0)	2 (1.6)	0 (0)	4 (3.1)	0 (0)
	Neg[b]	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
Residential	Pos	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	Neg	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Role Position	Pos	1 (0.8)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	0 (0)
	Neg	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
Support	Pos	3 (2.3)	0 (0)	1 (0.8)	0 (0)	1 (0.8)	0 (0)	3 (2.3)	1 (0.8)
	Neg	0 (0)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
Performance	Pos	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	Neg	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Global	Pos	1 (0.8)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	0 (0)
	Neg	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)

Table 5 continued

Variable		Anergia		Thought Dist.		Activation		Paranoid Bell.		Depression	
n (%)		Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
Work	Pos	3 (2.3)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	2 (1.6)	0 (0)	0 (0)	1 (0.8)
	Neg	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	2 (1.6)	1 (0.8)	2 (1.6)	1 (0.8)	0 (0)
Residential	Pos	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
	Neg	0 (0)	0 (0)	2 (1.6)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)
Role Position	Pos	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	2 (1.6)	0 (0)	0 (0)	1 (0.8)
	Neg	1 (0.8)	0 (0)	1 (0.8)	0 (0)	0 (0)	3 (2.3)	0 (0)	0 (0)	0 (0)	0 (0)
Support	Pos	2 (1.6)	0 (0)	2 (1.6)	0 (0)	3 (2.3)	2 (1.6)	1 (0.8)	0 (0)	0 (0)	0 (0)
	Neg	1 (0.8)	0 (0)	1 (0.8)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	1 (0.8)	1 (0.8)
Performance	Pos	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)
	Neg	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Global	Pos	1 (0.8)	1 (0.8)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	2 (1.6)	0 (0)	0 (0)	0 (0)
	Neg	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)

Note. Sample size is 128. Thought Dist. = Thought Disturbance; Paranoid Bell. = Paranoid Belligerence. Data presented as raw number of patients (percentage of sample).

[a]Decline or negative change on the MSIF scale corresponding with better functioning. [b]Increase or positive change on MSIF scale corresponding with worse functioning. [c]Decline or negative change in symptom severity corresponding with less severe symptoms. [d]Increase or positive change in symptom severity corresponding with greater symptom severity.

Table 6

Changes in Cognitive Performance and Real World Outcome

Variable		Vocabulary		Matrix Reasoning		Letter-Number Sequencing		Symbol Search		Trials 1-5	
n (%)		Pos[c]	Neg[d]	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
Work	Pos[a]	0 (0)	3 (2.3)	0 (0)	2 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.8)
	Neg[b]	0 (0)	1 (0.8)	0 (0)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)
Residential	Pos	0 (0)	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
	Neg	1 (0.8)	2 (1.6)	0 (0)	4 (3.1)	3 (2.3)	0 (0)	0 (0)	1 (0.8)	2 (1.6)	2 (1.6)
Role Position	Pos	2 (1.6)	1 (0.8)	0 (0)	3 (2.3)	2 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	Neg	0 (0)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	1 (0.8)	0 (0)	1 (0.8)	0 (0)	1 (0.8)
Support	Pos	1 (0.8)	0 (0)	0 (0)	2 (1.6)	1 (0.8)	1 (0.8)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)
	Neg	0 (0)	0 (0)	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	2 (1.6)
Performance	Pos	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
	Neg	0 (0)	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)
Global	Pos	0 (0)	0 (0)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)
	Neg	0 (0)	1 (0.8)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.8)

Table 6 continued

Variable		Short Delay Free Recall		Long Delay Free Recall		Intrusions		Phonemic Fluency		Semantic Fluency	
n (%)		Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
Work	Pos	1 (0.8)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	1 (0.8)	0 (0)
	Neg	0 (0)	2 (1.6)	0 (0)	0 (0)	0 (0)	2 (1.6)	1 (0.8)	0 (0)	1 (0.8)	0 (0)
Residential	Pos	1 (0.8)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)	1 (0.8)	0 (0)	0 (0)
	Neg	1 (0.8)	1 (0.8)	1 (0.8)	4 (3.1)	0 (0)	0 (0)	0 (0)	1 (0.8)	2 (1.6)	0 (0)
Role Position	Pos	0 (0)	0 (0)	0 (0)	2 (1.6)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)
	Neg	0 (0)	1 (0.8)	0 (0)	3 (2.3)	0 (0)	1 (0.8)	0 (0)	0 (0)	1 (0.8)	0 (0)
Support	Pos	1 (0.8)	1 (0.8)	1 (0.8)	2 (1.6)	0 (0)	1 (0.8)	0 (0)	0 (0)	1 (0.8)	0 (0)
	Neg	1 (0.8)	2 (1.6)	1 (0.8)	3 (2.3)	0 (0)	1 (0.8)	0 (0)	1 (0.8)	1 (0.8)	0 (0)
Performance	Pos	0 (0)	2 (1.6)	1 (0.8)	2 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	Neg	1 (0.8)	0 (0)	1 (0.8)	2 (1.6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Global	Pos	1 (0.8)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)	1 (0.8)	0 (0)	0 (0)	0 (0)
	Neg	0 (0)	0 (0)	0 (0)	2 (1.6)	0 (0)	2 (1.6)	1 (0.8)	0 (0)	1 (0.8)	0 (0)

Note. Sample size is 128. Vocabulary, Matrix Reasoning, Letter Number Sequencing, and Symbol Search subtests from the WAIS-III. Trials 1-5, Short Delay Free Recall, Long Delay Free Recall, and Intrusions are from the California Verbal Learning Test. Phonemic and Semantic Fluency are from the Controlled Oral Word Association Test. Data presented as raw number of patients (percentage of sample). [a]Decline or negative change on the MSIF scale corresponding with better functioning. [b]Increase or positive change on MSIF scale corresponding with worse functioning. [c]Decline or negative change in test score corresponding with worse performance. [d]Increase or positive change in test score corresponding with better performance.

Figure 1

Stability of symptom severity in schizophrenia (n=128)

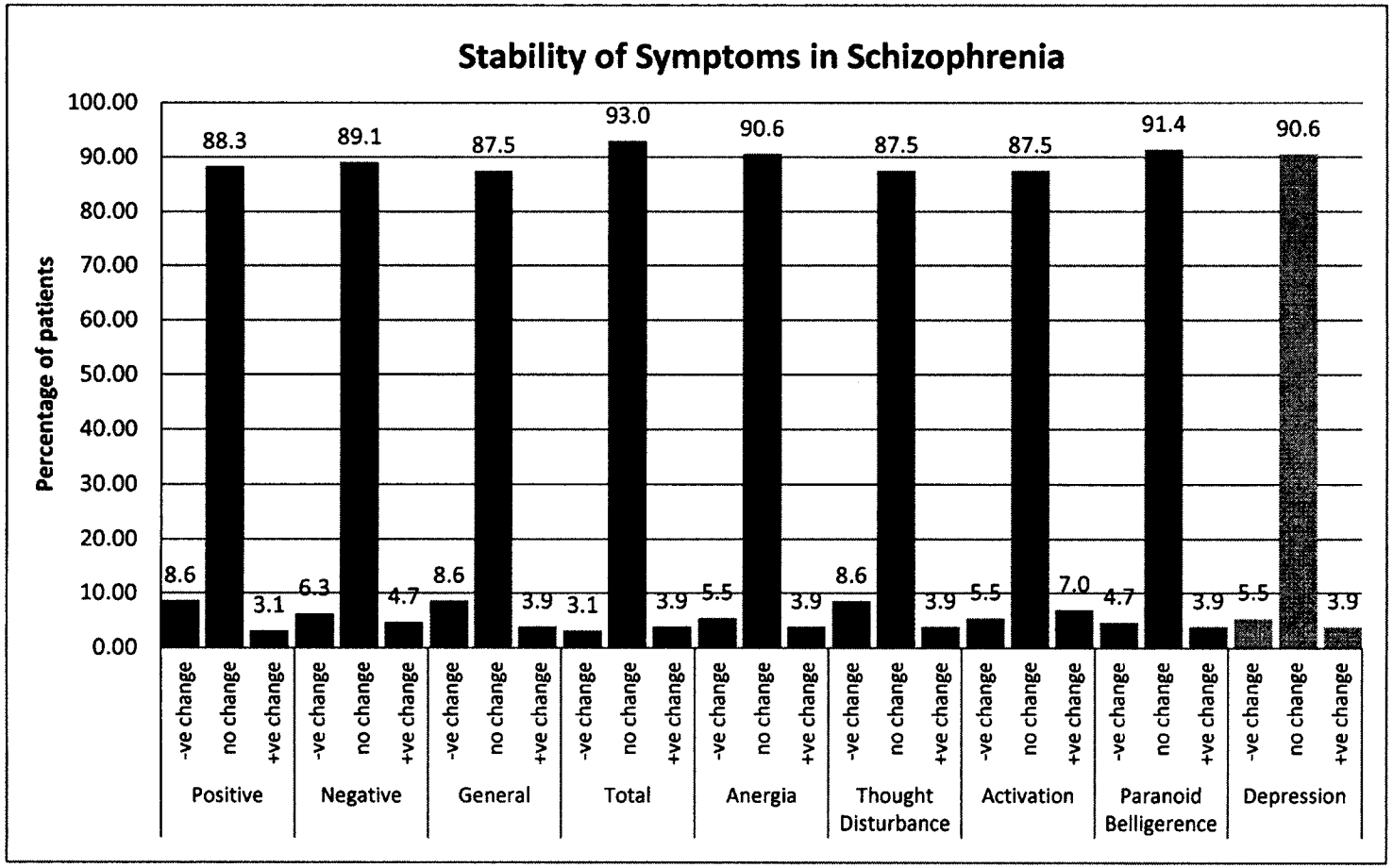

Figure 2

Stability of cognitive performance in schizophrenia (n=128)

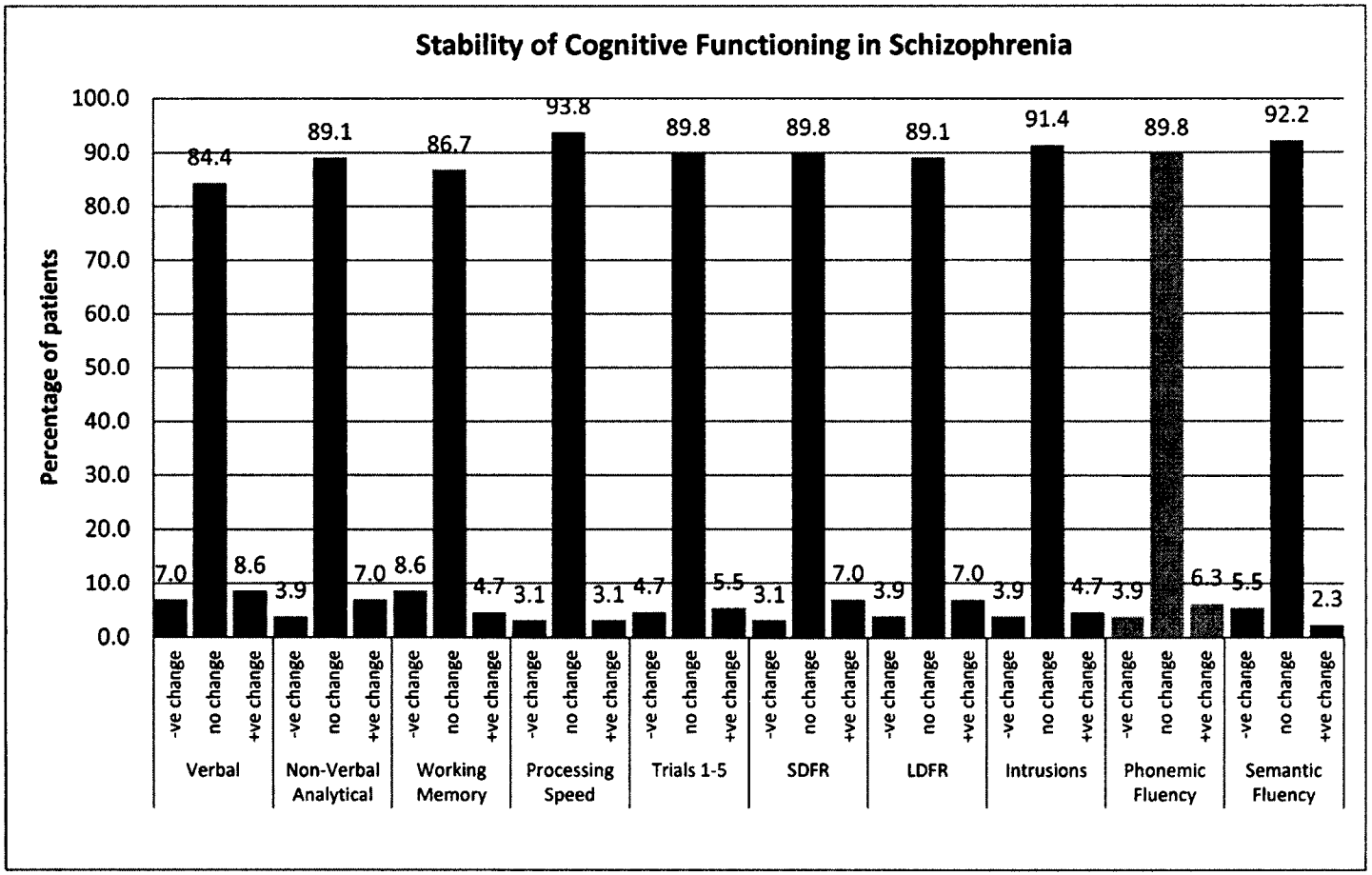

Figure 3

Stability of real world outcome in schizophrenia (n=128)