Published in final edited form as:
J Int Neuropsychol Soc. 2013 July ; 19(6): . doi:10.1017/S1355617713000179.

The Factor Structure of Neurocognition and Functional Capacity in Schizophrenia: A Multidimensional Examination of Temporal Stability

Philip D. Harvey[1,2], **Tenko Raykov**[3], **Elizabeth W. Twamley**[4,5], **Lea Vella**[6], **Robert K. Heaton**[4], and **Thomas L. Patterson**[4]

[1]Department of Psychiatry, University of Miami Miller School of Medicine, Miami, Florida

[2]Research Service, Miami VA Medical Center, Miami, Florida

[3]Department of Education and Measurement, Michigan State University, East Lansing, Michigan

[4]Department of Psychiatry, UCSD Medical Center, LaJolla, California

[5]Psychiatry Service, San Diego VA Medical Center, LaJolla, California

[6]Department of Psychology, UCSD-SDSU Joint Doctoral Program, San Diego, California

Abstract

Although neurocognition is commonly described in terms of different functional domains, some factor analytic studies have suggested a simpler dimensional structure for neuropsychological (NP) tests in patients with schizophrenia. Standardized tasks of everyday functioning, or tests of "functional capacity" (FC), are viewed differently from traditional NP tests, and are hence used as a co-primary measure in treatment studies. However, FC and NP tests have been found to be highly correlated. In fact, a recent study of ours suggested that performances on these different types of tasks constituted a single latent trait in a cross-sectional analysis. The current study examined the longitudinal factor structure of a combined set of NP and FC tests. Patients with schizophrenia (n=195) were examined at two assessment occasions separated by periods ranging from 6 weeks to 6 months. Participants were assessed with the MATRICS Consensus Cognitive Battery (MCCB) and two performance-based assessments of FC. A single latent trait was extracted using full information maximum likelihood procedures, and its temporal stability was examined in terms of: stability of the latent trait scores, the intercorrelations of the three indicators of the latent trait, and the stability of loadings for the FC and NP items underlying the latent trait at the two measurement occasions. All indices of temporal stability were confirmed, with stability not related to follow-up duration. Variation in clinical symptoms and treatments across the measurement occasions was negligible. These findings raise the question of whether cognitive abilities measured by NP tests and FC instruments are tapping a single ability construct, which might have shared causal influences as well.

Correspondence to: Philip D. Harvey, PhD, Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, 1120 NW 14th Street, Suite 1450, Miami, FL 33136, 305-243-1619 (Fax), 305-243-4094 (Voice), philipdharvey1@cs.com.

Conflict Of Interest Statement.

Dr. Harvey has received consulting fees from Abbott Labs, Amgen, Genentech, Johnson and Johnson, Pharma Neuroboost, Roche Pharma, Sunovion Pharma, and Takeda Pharma during the past year. Dr. Patterson has served as a consultant to Abbott Labs and Amgen. None of the other authors have any commercial interests to report

Keywords

Neuropsychology; Disability; Latent Traits; Longitudinal Studies; Functional capacity; Linear models

Cognitive impairments are present in nearly all patients with schizophrenia and are considered a central illness feature, although they are not included in the diagnostic criteria. Several pieces of evidence support this idea, including (1) the minimal relationship between cognitive impairment and severity of psychosis (Addington, et al., 1991), (2) stability of cognitive impairment across changes in clinical state (Harvey et al., 1990; Heaton et al., 2001), (3) suggestions that cognitive impairments are not completely caused by treatments for the illness (Keefe et al., 2007), differences in the environment (Harvey et al., 2009), deficits in motivation (Leifker et al., 2009), or global intellectual deficiencies (Palmer et al., 1997), and (4) findings of similar levels of impairment early and later in the illness (Saykin et al., 1994). Many studies have shown that cognitive impairments are stable over time, even over fairly long-term follow-ups (Heaton et al., 2001). Some evidence suggests that there are subsets of patients whose performance worsens over time (Harvey et al., 2010), but the results of those studies demonstrate rank-order similarity of performance in the context of worsening of the overall level of performance in these patients.

Also addressed in multiple studies is the factorial structure of cognitive impairments in schizophrenia. Although cognitive impairments in schizophrenia are typically described in terms of different functional domains (e.g., attention, memory, processing speed; Nuechterlein et al., 2004) measured by neuropsychological (NP) tests, several studies have found that much simpler factor structures often describe these impairments adequately. For instance, in the large CATIE study, a sample of 1331 schizophrenia cases was examined for the factor structure of their performance on 9 neurocognitive tests (Keefe et al., 2006). A principal components analysis suggested a single principal component (Eigenvalue=4.07, variance accounted for=45%). When followed with a confirmatory factor analysis, the best fitting model was a hierarchical model that posited 5 different cognitive domains contributing to a single overall factor structure. Thus, in this very large study, cognitive domains were found to have empirical validity, but the best fitting model suggested that these domains were best conceptualized as indicators of a global, unifactorial structure. While this is a common way to conceptualize cognitive impairment in schizophrenia (Chapman & Chapman, 1973), it does raise some questions about whether current measurement strategies are suitable or sensitive for detecting “specific” changes in “separable” (Nuechterlein et al., 2004) cognitive ability domains (e.g., working memory; processing speed; episodic memory) that might be specifically induced by treatments (pharmacological or cognitive remediation) targeting these specific functions.

A recent development in the study of cognitive impairment and functional disability in schizophrenia has been the study of functional capacity (FC): the ability to perform the cognitively demanding skills required for everyday residential, social, and vocational functioning (Harvey et al., 2007). Performance on FC measures has been found to be quite highly and consistently correlated with performance on traditional cognitive (NP) assessments (see Leifker et al., 2011 for a review of these relationships). The correlation between performance on composite scores of FC and global NP performance across different NP assessment strategies has been found to be quite high, with Pearson correlations routinely higher than r=.60. This level of correlation is essentially the same as seen between cognitive functioning composite scores and performance on individual cognitive tests, raising the question as to whether scores on FC measures may combine with NP

performance to constitute a single “ability” domain. This may be due to the lack of specificity of either NP or FC measures or redundancy of FC and NP measures.

This is more than a conceptual or terminological discussion, because of major pragmatic concerns. The US Food and Drug Administration (FDA) has adopted the perspective that treatments aimed at cognitive enhancement, either pharmacological or cognitive remediation, need a “separate” confirmation of their efficacy with a “co-primary” measure. Based on the results of a systematic comparative study (Green et al., 2011), the FDA has agreed that the UCSD Performance-Based Skills Assessment (UPSA) meets the criteria as a separate co-primary measure. A longitudinal finding, consistent with our previous cross-sectional results, that performance on the UPSA was statistically indistinguishable from traditional NP performance would raise questions about whether these assessments should both be required to demonstrate improvement in a treatment study. If all are measures of the same latent trait, why would separate tests of statistical significance be required for approval of a treatment? If they are indistinguishable statistically, then other evidence, such as differential sensitivity to treatment response, would be required to differentiate them.

There are additional reasons to examine the relationship between these performance-based domains. Although we have argued that FC is a central feature of the illness, like cognitive impairment, and that performance-based FC measures may be suitable as endophenotypic indicators for the illness (Harvey et al., 2012), more data are clearly needed. FC performance, like NP performance, appears to be minimally affected by psychotic symptoms (Bowie et al., 2006), to not be fully accounted for by poor motivation or educational disadvantage (McIntosh et al., 2011), and to be similarly performed across different phases of the illness (Harvey et al., 2010; 2011), and marked differences in real-world environments (Harvey et al., 2009). The temporal stability and factor structure of FC measures, however, are not well established. Retest studies of FC measures (Leifker et al., 2010; Keefe et al., 2011) have suggested very similar rank-order correlations and consistently small practice effects, similar to NP test performance in the same studies. No studies yet have examined the factor structure of FC performance over time in schizophrenia, particularly in reference to the relationships with NP performance.

In our analyses of the baseline data in Phase I of the Validation of Everyday Real-World Outcomes (VALERO) study, we found that the best fitting model of the structure of performance on NP tests and FC measures was a single latent trait (i.e., a unifactorial statistical solution; Harvey et al., 2011). A structural equation model was developed and fit to the available data with the latent variable modeling software Mplus. A single latent trait reflecting the shared variance of the three performance-based “ability” variables was developed using hierarchical linear modeling and three performance-based variables (NP performance measured with the modified MATRICS Consensus Cognitive Battery (MCCB; Nuechterlein et al., 2008); the UCSD Performance-Based Skills Assessment, Brief Version (UPSA-B; Mausbach et al., 2007); and the Advanced Finances Subtest of the Everyday Functioning Battery (EFB; Heaton et al., 2004) were significantly related to the ability latent trait, with the unifactorial solution having the best statistical properties. This latent trait, derived from three performance-based measures spanning cognition and functional capacity, was significantly related, as a group, to real-world functioning, independently rated across 6 rating scales using a comprehensive procedure of interviews with patients and informants. Only one of the six rating scales examined, the Specific Levels of Functioning (SLOF, Schneider et al., 1983) individually manifested a statistically significant relationship with the underlying latent trait.

Here we report the results of a longitudinal examination of the characteristics of FC and NP performance. We repeated our initial performance-based assessment at a non-randomized

follow-up interval that ranged from 6 weeks to 6 months (in order to yield variability in follow-up period). We examined all standard indices of temporal stability, including group-mean and rank-order (Pearson correlations; Intra Class Correlations), as well as the stability of the factor structure of the relationship between NP and FC performance. In so doing, we comprehensively examined whether the factor structure was similar at both assessments, whether the scores on the latent trait were the same at both assessments, and whether the correlational structure met criteria for "temporal invariance." This was defined in terms of cross-temporal stability of the overall latent trait scores, the intercorrelations of the three indicators of the latent trait, and whether the loadings for the FC and NP items underlying the latent trait were similar at the two measurement occasions. We also examined clinical symptoms at each assessment occasion. Thus, we examined all aspects of short-term temporal stability of functional capacity performance in a sample that was not involved in any interventions aimed at improvement of NP or FC performance. Based on our previous cross-sectional findings, we anticipated substantial stability of these ability domains and a possible single-factor solution over the course of the follow-up period.

Methods

Subjects

This is an additional set of analyses from Phase I of the VALERO study. The study participants were patients with schizophrenia or schizoaffective disorder who were receiving treatment at one of three different outpatient service delivery systems, two in Atlanta and one in San Diego. In addition, informants were interviewed concerning the everyday functioning of each of the patients, with these informants either being a high-contact clinician (case manager, psychiatrist, therapist, or residential facility manager; 20% of cases) or a friend or relative (80% of cases). All of these research participants provided signed, informed consent, and this research study was approved by local IRBs. The recruitment and assessment procedures for this study and demographic characteristics of the sample have been presented previously (Leifker et al., 2011; Harvey et al., 2011). Demographic and clinical characteristics of the sample are presented in Table 1.

Procedure

All patients were examined with a performance-based assessment of neurocognitive abilities and functional capacity. All available patients were re-examined between 6 weeks and 6 months after their initial assessment. Each aspect of the baseline assessment was performed again at the follow-up. The follow-up duration was determined by the availability of informants and patients, with contact with the participants re-initiated after 6 weeks and the assessment scheduled when convenient for the participants and the assessment sites. The mean duration of the follow-up interval was 126 days (SD=83).

Performance-based assessment

Neurocognition. We examined cognitive performance with a modified version of the MCCB. The MCCB cognitive domains (other than social cognition) include processing speed, verbal memory, visual memory, working memory, attention and concentration, and reasoning and problem solving. These domains are measured by one or more tests per domain for a total of 10 different tests (Nuechterlein et al., 2008). For this study, we did not include the social cognition measure from the MCCB, the Mayer–Salovey–Caruso Emotional Intelligence Test-Managing Emotions (MSCEIT), because recent meta-analyses have shown minimal correlations between neurocognitive and social cognitive measures (Fett et al., 2011; Ventura et al., in press). See Table 2 for each individual test and the t-scores for the sample as a whole in this study. This minor modification of the MCCB would make the results similar to previous work, such as our own (Bowie et al., 2006; 2008) that

did not include social cognition measures. We calculated a composite score, an average of 9 age-corrected T-scores based on the MCCB normative program, as our primary dependent variable.

Functional Capacity—We administered two different performance-based functional capacity measures. Participants' functional abilities were assessed using the UPSA-B, a measure of functional capacity in which patients are asked to perform everyday tasks related to communication and finances. During the *Communication* subtest, participants role-play exercises using an unplugged telephone (e.g., emergency call; dialing a number from memory; calling to reschedule a doctor's appointment). For the *Finance* subtest, participants count change, read a utility bill, and write a check for the bill. The UPSA-B requires approximately 10–15 minutes, and raw scores are converted into a total score ranging from 0–100, with higher scores indicating better functional capacity. We also administered the *Advanced Finances* subscale of the Everyday Functioning Battery (EFB; Heaton et al., 2004), designed to examine financial management in higher functioning individuals. The Advanced Finances test requires individuals to prepare bank deposits and write checks to pay bills, maintain a checkbook balance, and organize payments such that a pre-specified amount of money is left available at the end of the task. This instrument was selected because it measures abilities considered important for independent living and, at the time the study was planned, we were concerned that younger individuals with schizophrenia might evidence ceiling effects on the UPSA-B. Total scores on the Advances Finances subtest range from 0–13.

Data Analysis

Our goals in these analyses were to examine the temporal stability of the performance-based ability variables. This included basic analyses of group-mean and rank order stability, as well as similarity of the factor structure at the two assessments. In addition, the stability of the scores on the latent trait and the intercorrelations between the performance-based indicators and covariance between the performance-based indicators and the latent trait at the two assessments were examined.

All analyses performed used the Full Information Maximum Likelihood method of model fitting and parameter estimation (e.g., Raykov & Marcoulides, 2008). This method fits the model to all subjects and all available data in an incomplete data set, in that each subject contributes all available data to the model fitting and estimation process (under the widely made assumption of data missing at random; e.g., Little & Roubin, 2002). Thus, the fact that some subjects were missing some of the indicators at one or another of the assessments did not lead to dropping any subject in the case of this data set (i.e., no listwise deletion was carried out), with complete data available at both assessments for over 95% of the cases.

Results

We present the scores for the three performance-based variables that were the indicators for the latent trait in Table 3. Paired t-tests revealed that there was no significant difference over time in performance on any of the three variables (all $p>.11$) and that the Pearson correlations between performance across the two time points were statistically significant ($p<.001$) and substantial (all $r> .58$); these results were maintained even after application of a Bonferroni correction for multiple statistical tests. Clinical symptoms were equivalently stable, as evidenced by the scores presented in Table 3. There were no hospital admissions, other serious adverse events, and no changes in medication status between assessments 1 and 2.

When we compared the statistical significance of the differences in stability between the MCCB and UPSA-B performance, the MCCB correlation was significantly larger, z=6.53, p<.001. Next we repeated the analyses with Intra-Class Correlations (Two-way mixed, fixed raters absolute agreement: SPSS 20), finding that all of these correlations were also statistically significant (p<.001). When the analyses for the stability of the trait indicators were repeated as repeated-measures analysis of covariance, with inter-assessment time as the covariate, there were no significant covariate effects involving time between reassessments (all F<.24, all p>.63). Thus, test-retest interval did not affect performance differences between baseline and the reassessment.

We began our factor structure stability analyses by examining temporal stability in the covariance matrix of the three indicators of the latent trait over time (e.g., Jöreskog & Sörbom, 1996). These indicators were the mean t-score on the modified MCCB and the total scores on the UPSA-B and the EFB. In our first model, referred to as Model 1 below, we tested the hypothesis that the variances and covariances of these three indicators remained the same across the two assessment occasions, hence suggesting identical factor structure for the latent trait. We evaluated fit using the popular root mean square error of approximation (RMSEA). According to a widely adopted 'rule-of-thumb', the lower endpoint of the 90%-confidence interval (CI) for RMSEA is particularly informative when assessing model fit. Specifically, a finding of this endpoint being considerably below .05 is indicative of a plausible model (e.g., Raykov & Marcoulides, 2006).

Since Model 1 is obtained from the saturated model by imposing the constraints of the 3 measure variances being the same over time as well as the three measure covariances being also temporally identical, the chi-square associated with Model 1 represents a test statistic for the null hypothesis that these 6 variance and covariance parameters for the latent trait indicators are time invariant. Due to the fact that there were some mild deviations from normality on the measures at each assessment, we used the robust maximum likelihood (RML) method to fit the models (e.g., Muthen & Muthen, 2010). The resulting fit statistics for Model 1 indicated its tenability: chi-square value (X^2) = 13.08, for degrees of freedom (df) = 6, with associated p-value (p) = .04 and RMSEA = *.08* with a 90%-CI = (.01, .14). We interpret these fit results as indicative of Model 1 being an acceptable approximation to the analyzed data and thus conclude that the hypothesis of time invariance is plausible for the variances (standard deviations) and covariances of the three latent trait indicators. Since the correlations of these 3 measures are the ratios of covariances to products of standard deviations, it also follows that the hypothesis of stability in the degree of (linear) interrelationships between the three latent trait indicators is plausible. This supports the hypothesis that the correlations between the three latent trait indicators are also time-invariant across the two assessments.

With this finding from Model 1 of time-invariance in the interrelationship indexes of the latent trait indicators and stability in variance, we moved on to specific modeling of the temporal change of the values in the latent trait across the two assessment occasions. To this end, we fitted to the data of these measures a correspondingly extended model, referred to as Model 2 below. This model postulated at each assessment a latent trait measured by the above set of performance-based indicators. In addition, we postulated time invariance in the factor loadings and intercepts of its three indicators, and parameterized the change in the performance-based latent trait as a third latent variable (e.g., Raykov & Marcoulides, 2006). Hence, in addition to its fit, of particular interest in Model 2 are the mean change in the performance based latent trait (which is parameterized as the mean of that third latent variable).

Model 2 was found to be associated with tenable fit indexes as well: X^2 = 13.14, df = 17, p = .72 and RMSEA = *.02* with a 90%-CI = (.0, .05). In the model, the estimate of mean trait change (true change) was estimated at .49, with a standard error (SE) = .28 and associated p-value (p) = .08. This finding of nonsignificance of mean change in the latent trait (i) suggests that the underlying level of performance of the latent trait as a whole was also stable over time, excluding the possibility that practice and exposure affected the scores on the latent trait, and (ii) is consistent with the findings above of the temporal stability for the three performance-based indicators of the latent trait. Further, in this model the correlation of true change with baseline performance (starting position on the latent trait) was estimated at −.022 (SE=.11), p = .85. This finding was interpreted as suggesting that the performance at baseline on the trait is not related (linearly) to any temporal changes, thus indicating that baseline performance predicts reassessment *performance*, but not *changes* in performance over time.

Finally, Model 2 allows us to address also the question whether the degree of relationship between the trait indicators (indexed by their “loadings”), on the one hand, and the underlying factor on the other, is time invariant. To this end, we fitted a relaxed version of Model 2, referred to below as Model 3, where the three trait loadings were not constrained for time invariance, in order to evaluate the possibility of lack of time invariance. Model 3 was similarly associated with plausible fit indexes: X^2 = 11.93, df = 15, p = .69 and RMSEA = *.03* with a 90%-CI = (.0, .05). To test the hypothesis of time-invariance of interest here, owing to the fact that we used the robust maximum likelihood method for model testing and parameter estimation, as mentioned above, we applied the Satorra correction difference (change, denoted as “$_c$”) in the chi-square test statistic (e.g., Satorra, 2000): X^2_c = 1.452, for df = 2, p = .484. This nonsignificant finding suggests that the degree of (linear) relationship between latent trait and its indicators remains stable over time. These loadings are presented in Table 4.

Discussion

Several findings regarding the temporal stability of performance-based ability variables in schizophrenia were revealed in these analyses. First, the three performance-based ability variables themselves were stable over time, with no evidence of change at the group-mean or rank order level. These data suggest the possible interpretation that there is a single underlying ability trait, with the NP tests having slightly higher test-retest reliability than the FC measures, as indicators of this single latent trait. Although NP performance was significantly more stable, the statistically significant stability of FC measures was again demonstrated in a completely independent sample from the previous studies (Leifker et al., 2010; Keefe et al., 2011). A model where all three measures, examining performance on NP tests and two different FC measures, constituted a single latent trait, was confirmed at baseline and replicated at the reassessment. The intercorrelations between the variables were stable over time and the loadings of the three performance-based variables on the latent trait were also stable over time. There was no significant change for the group as a whole on the ability-based latent trait, and baseline performance did not influence change scores, although it strongly predicted performance at the reassessment (because of the high stability of the indicators). Potential confounds such as clinical symptoms were equivalently temporally stable. Finally, previous studies of cognitive assessments with tasks similar to or included in the MCCB (McClure et al., 2007) have found that individual cognitive measures have relatively equivalent (and considerably smaller) correlations with FC measures than composite scores.

There are several limitations of this research design and study. The follow-up interval was designed for assessment of test-retest stability and not long-term changes, so it is necessarily

short. We did not examine all the cognitive tests in the MCCB separately as latent trait indicators; cognitive performance with similar tests has been examined with factor analyses in much larger samples (Keefe et al., 2006), and we have addressed the issue of the best fitting model in this database using all 9 tests independently in a separate paper (submitted). The MCCB is designed to be used as an outcome measure by calculating a total score such as that employed in the present study (Nuechterlein et al., 2008; Keefe et al., 2011), thus suggesting that treating the MCCB as a single outcome measure based on a sampling of multiple NP domains is consistent with the intentions of its developers.

These results suggest that the ability of people with schizophrenia to perform cognitively challenging tests, whether they are designated as “neurocognitive” or “functional capacity,” seems statistically unidimensional. This raises the question as to whether clinical treatment studies of cognition in schizophrenia ought to consider cognition and functional capacity as separate outcomes domains. Some studies have found that functional capacity measures were slightly stronger correlates of real-world functioning than functional capacity (Bowie et al., 2006; 2008; Leifker et al., 2009), and other have found the opposite result (Heinrichs, et al., 2010); in fact, both FC and NP performance predict RW functioning, probably because of their high intercorrelations. Scores on both types of measures in the current study manifest all typical characteristics of single-factor temporal stability. Similar to other conceptions of the global nature of cognitive deficits measured with performance-based assessments in people with schizophrenia (for review see Dickinson & Harvey, 2009), these findings implicate a single ability factor, already found to be related to everyday outcomes as evidenced by the baseline assessment of this sample.

These findings may have implications as well for the genetic determinants of these abilities. The elements of the current set of NP tests in the MCCB has been studied for heritability in families, with a number of significant results (Gur et al., 2008; Greenwood, et al., 2008), and performance on variants of these tests has been shown to be related to genetic variation (Greenwood et al., 2011). If these different ability domains, NP and FC, are not really separable, it may be worthwhile to determine whether the same genomic factors that impact on NP performance influence FC.

In conclusion our results implicate a single ability factor that is stably related to both NP and FC test performance. All aspects of this ability trait measured by these tests manifested temporal stability. The differences in correlations between the ability latent trait and NP as compared to FC measures may be due largely to differences in test-retest reliability or the restricted ranges of the FC measures. Since these measures are targeted at disability and not general abilities, nondisabled research participants would be expected to obtain close to perfect scores. This range truncation would explain the differences in test-retest reliability and are consistent with the idea that the differences measures may still have a common ability origin. These data raise questions about whether both NP and FC measures should be required to substantiate improvements in clinical treatment studies.

Acknowledgments

Grant Support: This research was supported by Grants MH078775 to Dr. Harvey and MH078737 to Dr. Patterson from the National Institute of Mental Health.

References

Addington J, Addington D, Maticka-Tyndale E. Cognitive functioning and positive and negative symptoms in schizophrenia. Schizophrenia Research. 1991; 4:123–134. [PubMed: 1931805]

Bowie CR, Leung WW, Reichenberg A, McClure MM, Patterson TL, Heaton RK, Harvey PD. Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. Biological Psychiatry. 2008; 63:505–511. [PubMed: 17662256]

Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of Real World Functional Performance in Schizophrenia: Correlations with Cognition, Functional Capacity, and Symptoms. American Journal of Psychiatry. 2006; 163:418–425. [PubMed: 16513862]

Chapman, LJ.; Chapman, JM. Disordered thought in schizophrenia. New York: Appleton-Century-Crofts; 1973.

Dickinson D, Harvey PD. Systemic hypotheses for generalized cognitive deficits in schizophrenia: A new take on an old problem. Schizophrenia Bulletin. 2009; 35:403–414. [PubMed: 18689868]

Fett AK, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: meta-analysis. Neuroscience and Biobehavioral Review. 2011; 35:573–588.

Greenwood TA, Braff DL, Light GA, Cadenhead KS, Caulkins ME, Dorcas DJ, Schork NJ. Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. Archives of General Psychiatry. 2007; 64:1242–1250. [PubMed: 17984393]

Greenwood TA, Lazzeroni LC, Murray SS, Cadenhead KS, Caulkins ME, Dorcas DJ, Braff DL. Analysis of 94 candidate genes and twelve endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. American Journal of Psychiatry. 2011; 168:930–946. [PubMed: 21498463]

Gur RE, Calkins ME, Gur RC, Horan WP, Nuechterlein KH, Seidman LJ, Stone WS. The Consortium on the Genetics of Schizophrenia: Neurocognitive endophenotypes. Schizophrenia Bulletin. 2007; 33:49–68. [PubMed: 17101692]

Harvey PD, Docherty NM, Serper MR, Rasmussen M. Cognitive deficits and thought disorder: II. An eight-month followup study. Schizophrenia Bulletin. 1990; 16:147–156. [PubMed: 2333475]

Harvey PD, Helldin L, Bowie CR, Heaton RK, Olsson AK, Hjärthag F, Patterson TL. Performance-based measurement of functional disability in schizophrenia: a cross-national study in the United States and Sweden. American Journal of Psychiatry. 2009; 166:821–827. [PubMed: 19487393]

Harvey PD, McClure MM, Patterson TL, McGrath JA, Pulver AE, Bowie CR, Siever LJ. Impairment in Functional Capacity as an Endophenotype Candidate in Severe Mental Illness. Schizophrenia Bulletin. 2012; 38:1318–1326. [PubMed: 21562142]

Harvey PD, Raykov T, Twamley EW, Vella L, Heaton RK, Patterson TL. Validating the measurement of real-world functional outcome: Phase I results of the VALERO study. American Journal of Psychiatry. 2011; 168:1195–1201. [PubMed: 21572166]

Harvey PD, Reichenberg A, Bowie CR, Patterson TL, Heaton RK. The course of neuropsychological performance and functional capacity in older patients with schizophrenia: Influences of previous history of long-term institutional stay. Biological Psychiatry. 2010; 67:933–939. [PubMed: 20202624]

Harvey PD, Velligan DI, Bellack AS. Performance-based measures of functional skills: usefulness in clinical treatment studies. Schizophrenia Bulletin. 2007; 33:1138–1148. [PubMed: 17493956]

Heaton RK, Gladsjo JA, Palmer BW, et al. Stability and course of neuropsychological deficits in schizophrenia. Archives of General Psychiatry. 2001; 58:24–32. [PubMed: 11146755]

Heaton RK, Marcotte TD, Mindt MR, Sadek J, Moore DJ, Bentley H, Grant I. The impact of HIV-associated neuropsychological impairment on everyday functioning. Journal of the International Neuropsychological Society. 2004; 10:317–331. [PubMed: 15147590]

Heinrichs RW, Ammari N, Miles AA, McDermid, Vaz S. Cognitive performance and functional competence as predictors of community independence in schizophrenia. Schizophrenia Bulletin. 2010; 36:381–387. [PubMed: 18667392]

Jöreskog, KG.; Sörbom, D. Lisrel 8 user's reference guide. Lincolnwood, IL: Scientific Software International; 1996.

Keefe RS, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, Lieberman JA. Neurocognitive Effects of Antipsychotic Medications in Patients with Chronic Schizophrenia in the CATIE Schizophrenia Trial. Archives of General Psychiatry. 2007; 64:633–647. [PubMed: 17548746]

Keefe RS, Bilder RM, Harvey PD, Davis SM, Palmer BW, Gold JM, Lieberman JA. Baseline neurocognitive deficits in the CATIE Schizophrenia Trial. Neuropsychopharmacology. 2006; 31:2033–2046. [PubMed: 16641947]

Keefe RS, Fox KH, Harvey PD, Cucchiaro J, Siu C, Loebel A. Characteristics of the MATRICS Consensus Cognitive Battery in a 29-site antipsychotic schizophrenia clinical trial. Schizophrenia Research. 2011; 125:161–168. [PubMed: 21075600]

Leifker FR, Bowie CR, Harvey PD. The determinants of everyday outcomes in schizophrenia: Influences of cognitive impairment, clinical symptoms, and functional capacity. Schizophrenia Research. 2009; 115:82–87. [PubMed: 19775869]

Leifker FR, Patterson TL, Heaton RK, Harvey PD. Validating measures of real-world outcome: the results of the VALERO Expert Survey and RAND Panel. Schizophrenia Bulletin. 2011; 37:334–343. [PubMed: 19525354]

Leifker FR, Patterson TL, Bowie CR, Mausbach BT, Harvey PD. Psychometric properties of performance-based measurements of functional capacity: test-retest reliability, practice effects, and potential sensitivity to change. Schizophrenia Research. 2010; 119:246–252. [PubMed: 20399613]

Little, RJ.; Rubin, DB. Statistical analysis with missing data. New York: Wiley; 2002.

Mausbach BT, Harvey PD, Goldman SR, Jeste DV, Patterson TL. Development of a brief scale of everyday functioning in persons with serious mental illness. Schizophrenia Bulletin. 2007; 33:1364–1372. [PubMed: 17341468]

McClure MM, Bowie CR, Patterson TL, Heaton RK, Weaver C, Anderson H, Harvey PD. Correlations of Functional Capacity and Neuropsychological Performance in older Patients with Schizophrenia: Evidence for Specificity of Relationships? Schizophrenia Research. 2007; 89:330–338. [PubMed: 16982175]

McIntosh BJ, Zhang XY, Kosten T, Tan SP, Xiu MH, Rakofsky J, Harvey PD. Performance based assessment of functional skills in severe mental illness: Results of a large-scale study in China. Journal of Psychiatric Research. 2011; 45:1089–1094. [PubMed: 21300378]

Muthen, LK.; Muthen, B. Mplus user's guide. Los Angeles, CA: Muthen & Muthen; 2010.

Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. Schizophrenia Research. 2004; 72:29–39. [PubMed: 15531405]

Neuchterlein KH, Green MF, Kern RS, Baade LE, Barch D, Cohen J, Stover E. The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. American Journal of Psychiatry. 2008; 165:203–213. [PubMed: 18172019]

Raykov, T.; Marcoulides, GA. A first course in structural equation modeling. Mahwah, NJ: Lawrence Erlbaum and Associates; 2006.

Raykov, T.; Marcoulides, GA. An introduction to applied multivariate analysis. New York: Taylor & Francis; 2008.

Rosen A, Hadzi-Pavlovic D, Parker G. The Life Skills Profile: A measure assessing function and disability in schizophrenia. Schizophrenia Bulletin. 1989; 15:325–337. [PubMed: 2749191]

Satorra, A. Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In: Heijmans, RDH.; Pollock, DSG.; Satorra, A., editors. Innovations in multivariate statistical analysis. A festschrift for Heinz Neudecker. London: Kluwer Academic Publishers; 2000. p. 233-247.

Saykin AJ, Shtasel DL, Gur RE, Kester DB, Mozley LH, Stafiniak P. Neuropsychological deficits in neuroleptic naive patients with first episode schizophrenia. Archives of General Psychiatry. 1994; 51:124–131. [PubMed: 7905258]

Schneider LC, Struening EL. SLOF: a behavioral rating scale for assessing the mentally ill. Social Work Research and Abstracts. 1983; 19:9–21. [PubMed: 10264257]

Ventura J, Wood RC, Hellemann GS. Symptom domains and neurocognitive functioning can help differentiate social cognitive processes in schizophrenia. Schizophrenia Bulletin. In press.

J Int Neuropsychol Soc. Author manuscript; available in PMC 2014 July 01.

Table 1

Demographic information and performance-based scores for the three sites

	Skyland Trail		Atlanta VA		San Diego				
	n	**%**	**n**	**%**	**n**	**%**	**χ^2**	**p-value**	
Sex							1.08	0.582	
(% Male)	38	69%	30	75%	66	66%			
Racial characteristics							41.23	<0.001	
Caucasian	43	78%	9	23%	54	54%			
African American	10	18%	30	77%	34	34%			
Other	2	4%	0	0%	12	12%			
Ethnicity							15.57	<0.001	
Hispanic	2	4%	0	0%	21	21%			
Residential Status							78.99	<0.001	
Independent, financially responsible	7	13%	24	60%	57	57%			
Independent, not financially responsible	11	20%	5	13%	17	17%			
Unsupervised residential facility	4	7%	5	13%	23	23%			
Supervised residential facility	33	60%	6	15%	3	3%			
Employment Status							10.12	0.120	
Employed part-time	10	18%	3	8%	7	7%			
Employed full-time	3	5%	2	5%	1	1%			
Unemployed	40	73%	35	88%	89	89%			
Retired	2	4%	0	0%	3	3%			
	N	M(SD)	n	M(SD)	n	M(SD)	F	df	p-value
Age	55	35.78 (14.13)	40	47.3 (8.58)	100	47.25 (8.89)	23.27	2, 192	<0.001
Years of Education	54	14.13 (2.81)	27	13.07 (1.64)	100	12.32 (2.34)	9.95	2, 178	<0.001
Baseline Scores									
PANSS Total Score (range: 30–210)	55	62.11 (13.55)	40	64.55 (15.86)	100	62.21 (14.47)	0.43	2, 192	0.653

Note: PANSS: Positive and Negative Syndrome scale; UPSA-B UCSD Performance-Based skills Assessment-Brief; EFB: Everyday Functioning Battery; MCCB: MATRICS Consensus Cognitive Battery

Table 2

Standard Scores on Individual MCCB Tests (N=195)

	M	SD
Processing Speed		
Trail-Making Test Part A	36.79	12.37
BACS Symbol Coding	33.33	10.20
Animal Naming	42.49	9.21
Working Memory		
Maryland-Letter Number Span	37.36	10.14
WMS-III Spatial Span	39.22	10.74
Verbal Memory		
HVLT-R Learning Trials	37.22	8.48
Spatial Memory		
BVMT-R Learning Trials	39.96	13.03
Attention		
IP-CPT d′	34.18	10.40
Reasoning and Problem Solving		
NAB Mazes	41.42	10.00

Notes. T scores. M=50; SD-10;

BACS: Brief Assessment of Cognition in Schizophrenia; WMS: Wechsler Memory Scale; HVLT-R: Hopkins verbal Learning Test-Revised; BVMT-R: Brief Visual Memory Test, revised; IP-CPT: Identical Pairs Continuous Performance Test; NAB: Neuropsychological Assessment Battery

Table 3

Scores on the Three Performance-Based Ability Variables and Clinical Symptoms at the Baseline and Follow-up Assessments

	Baseline		Retest						
Variable	**M**	**SD**	**M**	**SD**	**t**	**p**	**Pearson r 1–2**	**IntraClass Correlation**	**p**
Modified MCCB t-Score[a]	37.57	6.96	38.03	7.72	1.59	.13	.88	.91	.001
Total Score on the UPSA-B[b]	76.55	12.98	77.25	14.03	0.74	.46	.61	.76	.001
EFB Advanced Finances Subscale[c]	8.78	3.67	8.95	8.81	0.58	.56	.59	.74	.001
PANSS Total Score	57.47	13.07	58.95	13.86	0.45	.33	.69	.78	.001

[a] Scores range from 20 to 80 with a mean of 50 and SD of 10

[b] Score range from 0–100

[c] Scores range from 0–13

Table 4

Factor Loadings for the Ability Variables onto the Latent Trait

	Performance-based ability measures					
	Time 1			Time 2		
	Standardized factor loading	t	p	Standardized factor loading	t	p
MATRICS Consensus Cognitive Battery, Modified	0.78	16.54	0.001	0.86	19.52	0.001
UCSD Performance-Based Skills Assessment, Brief Version	0.71	15.22	0.001	0.71	15.11	0.001
Everyday Functioning Battery	0.62	12.02	0.001	0.68	13.23	0.001