

Efficiency of the CATIE and BACS neuropsychological batteries in assessing cognitive effects of antipsychotic treatments in schizophrenia

S. KRISTIAN HILL,¹ JOHN A. SWEENEY,¹ ROBERT M. HAMER,^{2,3} RICHARD S.E. KEEFE,⁴
DIANA O. PERKINS,² HONGBIN GU,² JOSEPH P. MCEVOY,⁴ AND JEFFREY A. LIEBERMAN⁵

¹Center for Cognitive Medicine, Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois

²Department of Psychiatry, University of North Carolina School of Medicine, Chapel Hill, North Carolina

³Department of Biostatistics, University of North Carolina School of Public Health, Chapel Hill, North Carolina

⁴Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina

⁵Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York State Psychiatric Institute, New York, New York

(RECEIVED January 22, 2007; FINAL REVISION September 12, 2007; ACCEPTED September 17, 2007)

Abstract

Efficient and reliable assessments of cognitive treatment effects are essential for the comparative evaluation of procognitive effects of pharmacologic therapies. Yet, no studies have addressed the sensitivity and efficiency with which neurocognitive batteries evaluate cognitive abilities before and after treatment. Participants were primarily first episode schizophrenia patients who completed baseline ($n = 367$) and 12-week ($n = 219$) assessments with the BACS (Brief Assessment of Cognition in Schizophrenia) and CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) neuropsychological batteries in a clinical trial comparing olanzapine, quetiapine, and risperidone. Exploratory factor analysis revealed that performance on both batteries was characterized by a single factor of generalized cognitive deficit for both baseline performance and cognitive change after treatment. Both batteries estimated similar levels of change following treatment, although the BACS battery required half the administration time. Because a unitary factor characterized baseline cognitive abilities in early psychosis as well as cognitive change after treatment with atypical antipsychotic medications, short batteries such as the BACS may efficiently provide sufficient assessment of procognitive treatment effects with antipsychotic medications. Assessment of cognitive effects of adjunctive therapies targeting specific cognitive domains or impairments may require more extensive testing of the domains targeted to maximize sensitivity for detecting specific predicted cognitive outcomes. (*JINS*, 2008, *14*, 209–221.)

Keywords: Schizophrenia, Neuropsychology, Antipsychotics, Cognition, Clinical trials methodology, CATIE, BACS

INTRODUCTION

Generalized cognitive deficits have been well documented in schizophrenia and are considered core features of the disorder (Bilder et al., 2000; Blanchard & Neale, 1994; Gold et al., 1992; Hill et al., 2001; Saykin et al., 1991). Moderate to marked deficits are typically seen across a wide

range of cognitive abilities, are present during the first episode of psychosis, and endure after pharmacological treatment (Hill et al., 2004a; Hoff et al., 1999). Moreover, cognitive deficits have been linked with long-term functional disability (Green, 1996). As a result, cognitive enhancement has been recognized as an important treatment target in schizophrenia.

Evaluation of the procognitive effects of atypical relative to typical antipsychotics has indicated somewhat greater benefit with atypical compared with typical antipsychotics (Bilder et al., 2002; Green et al., 2002; Harvey et al., 2000; Keefe et al., 1999; Purdon et al., 2000). In general, modest cognitive benefits from antipsychotic medications are characterized by reduction of generalized cognitive deficits across

Correspondence and reprint requests to: Dr. S. Kristian Hill, University of Illinois at Chicago, Center for Cognitive Medicine, Department of Psychiatry (M/C 913), 912 South Wood Street, Suite 235, Chicago, IL 60612. E-mail: shill@psych.uic.edu

Presented in part at the annual International Neuropsychology Society meeting in Portland, OR, February 2007; and the 2007 International Congress for Schizophrenia Research in Colorado Springs, CO.

a wide range of abilities, rather than particular effects on a specific neuropsychological domain (Buchanan et al., 1994; Cassens et al., 1990; Hill et al., 2004a; Rollnik et al., 2002).

Evaluating the cognitive benefits of treatments requires reliable, valid, and efficient assessment procedures. The cost of testing in large clinical trials and limited cooperation of schizophrenia patients are both motivators for developing brief efficient batteries, yet the degree to which shorter batteries may have reduced sensitivity to treatment effects is an opposing concern. In recent years, several brief test batteries have been developed for assessing cognition in clinical trials of antipsychotic medications in schizophrenia. One of these, the Brief Assessment of Cognition in Schizophrenia (BACS; Keefe et al., 2004), requires less than 35 min to administer and has an excellent completion rate and high reliability (Keefe et al., 2004). A second battery was developed and used in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project (Keefe et al., 2006). This was the first large community trial designed to evaluate the comparative effectiveness of different antipsychotic treatments for schizophrenia (Keefe et al., 2003). Tests for the CATIE battery were selected, in part, based on their sensitivity to known deficits in the disorder as well as the relation of measured deficits to outcome variables such as community function (Keefe et al., 2003). No prior studies have compared these or other batteries in terms of sensitivity and efficiency for estimating both cognitive abilities and change in cognitive performance during clinical trials.

Consistent with the wide use of atypical or “second generation” antipsychotics in first episode patients, the Comparison of Atypicals in First Episode schizophrenia (CAFE) study compared three atypical agents (olanzapine, quetiapine, and risperidone) in the treatment of first episode and early course schizophrenia with cognitive change as a secondary outcome as measured by the CATIE and BACS batteries. This provided a rare opportunity to directly compare the psychometrics, utility, and efficiency of two neuropsychological batteries in a clinical trial setting.

METHODS

The CAFE study compared the effectiveness of atypical antipsychotics in a randomized double-blind clinical trial across 26 sites. Details of the study design and direct comparison of tolerability and efficacy of olanzapine, quetiapine, and risperidone have been presented elsewhere (Keefe et al., 2003, 2007). The protocol was approved by the local internal review boards, and each participant provided written informed consent.

Participants

Patients were recruited who had recently experienced an episode of acute psychosis that required treatment initiation and met Diagnostic and Statistical Manual for Mental Disorders-IV (DSM-IV; American Psychiatric Association, 1987/1994) criteria for schizophreniform, schizophrenia,

or schizoaffective disorder based on Structured Clinical Interview for DSM-IV (SCID; First et al., 1995). Patients were excluded if they had been ill for more than 5 years or had prior lifetime antipsychotic treatment for 16 cumulative weeks. Other exclusion criteria included non-English speaking, mental retardation, unstable medical conditions, pregnancy or nursing, serious head injury, neurologic disease, substance abuse (past 3 months), past substance dependence, and systemic disorders known to affect brain function. At baseline, 400 patients were randomized to treatment with olanzapine (2.5–20 mg/day), quetiapine (100–800 mg/day), or risperidone (0.5–4 mg/day). Before study enrollment, 76% of participants were exposed to antipsychotic treatments for a median of 4 weeks (range, <1–52). Any previous antipsychotic therapy was tapered and discontinued during the first 2-weeks of double-blind treatment, and no subsequent use of an additional antipsychotic was allowed. Treatment with adjunctive antidepressants or mood stabilizers were not allowed during the first 8-weeks of treatment. Anticholinergic medications were permitted for a total of 2 weeks and low doses were encouraged. This strategy resulted in limited (<5%) use of benzodiazepines, antidepressants, mood stabilizers, and anticholinergics. Of the 219 participants who completed the 12-week cognitive assessments 35.6% ($n = 78$) were assigned olanzapine, 31.5% ($n = 69$) to quetiapine, and 32.9% ($n = 72$) to risperidone. Because there were no significant group differences on global cognitive performance for the atypical antipsychotics at baseline or the 12-week follow-up (Keefe et al., 2007), data were pooled across treatment conditions for the statistical analysis.

Clinical Assessments

The Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) and Clinical Global Impression scale (CGI; Guy, 1976) were used to assess psychopathology. All patients had ratings of ≥ 4 on at least one PANSS psychosis item at the point of maximum severity of illness to date. Social and occupational function were evaluated with the Heinrichs-Carpenter Quality of Life Scale (QLS), and the impact of insight on treatment adherence was evaluated using the Insight into Treatment and Attitudes Questionnaire (ITAQ). As detailed in a separate report, symptom reduction was substantial in each treatment group, while improvements in social and vocational function were small (<0.2 *SDs*) after 12 weeks of treatment (McEvoy et al., 2007).

Neurocognitive Measures

Baseline cognitive assessments were conducted before initiation of study treatment. The CATIE battery was the primary cognitive measure. This battery requires approximately 90 min to administer 10 tests that characterize six neuropsychological domains (see Keefe et al., 2003, 2006). The BACS can be administered in 35 min and consists of six tests covering four domains (Keefe et al., 2004). The CATIE

battery was always administered before the BACS. In the CATIE battery, alternate forms were available for the Hopkins Verbal Learning Test. Alternate forms for the BACS List Learning test were the same as the final versions described in the BACS validation report (Keefe et al., 2004). Alternate forms were also used for the Tower Test. The BACS validation study showed that alternate Verbal Fluency forms were not needed (Keefe et al., 2004); thus, subsequent BACS versions did not include alternate Verbal Fluency forms. The BACS Category Instance Generation (CIG) test was administered at baseline; however, the alternate version was redundant with the CATIE battery at follow-up, and the BACS CIG was consequently excluded from all data analysis. Each tester held a doctoral degree or was supervised by a PhD-level psychologist, had previous testing experience, and demonstrated testing competence during training.

Neurocognitive assessments were completed at baseline, 12-weeks, and 52-weeks/termination. Comparison of the cognitive batteries was restricted to baseline and 12-week data because of greater attrition at 52-weeks. Of the 400

patients enrolled in the study, 4.25% were not administered cognitive tests. Two patients were excluded due to extremely deficient baseline scores. Of the 16 CATIE/BACS tests 9.4% patients had missing data for one test, 3.4% for two tests, 1.0% for three tests, and an additional 2.6% for four or more tests. Missing data on a maximum of two tests was selected as the criterion for inclusion in the analyses, and baseline exploratory factor analyses were thus limited to 367 patients. Missing data points were imputed by means of linear regression using available neurocognitive data to predict missing values. At the 12-week follow-up, 222 patients were administered both cognitive batteries. Consistent with the criterion of 2 or fewer missing tests for inclusion, another 3 patients were excluded from the follow-up analyses. Computation of within-subject effect size of change, exploratory factor analysis of change scores, and regression analysis was restricted to a sample of 219 patients. As can be seen in Tables 1 and 2, the follow-up sample was well matched to the baseline sample in terms of demographics and neuropsychological performance. Change scores were computed as the difference between performance at

Table 1. Demographic characteristics for participants included in baseline analyses only ($n = 367$) and those who also completed the 12-week assessment

Demographic characteristic	Baseline group		Longitudinal group			
	$n = 367$		$n = 219$			
Sex						
Male	72.8%				71.7%	
Female	27.2%				28.3%	
Race						
White	52.0%				55.3%	
Black	42.0%				38.8%	
Asian/Latino/Other	6.0%				5.9%	
Less than High School Education						
Patients	29.7%				28.3%	
Parent	10.2%				10.8%	
Antipsychotic naive	23.8%				20.1%	
	Baseline data for all participants		Baseline data for longitudinal group		12-week data for longitudinal group	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	24.47	5.60	24.45	5.61		
Illness Duration (months)	12.61	14.59	13.53	17.91		
Duration Untreated Psychosis (weeks)	18.42	5.03	18.52	5.17		
WRAT-III Reading	93.46	16.61	94.68	16.92		
PANSS Total	73.34	15.46	72.58	15.74	57.24	15.52
PANSS Positive	18.43	4.95	18.31	5.08	12.47	4.71
PANSS Negative	19.40	6.12	19.28	6.01	16.77	5.94
CGI-Severity	4.24	0.75	4.18	0.76	3.17	0.99
Heinrichs-Carpenter QLS						
Vocational Subscale	9.02	6.99	9.55	7.23	10.45	7.04
Social Subscale	20.77	10.29	22.14	10.79	23.21	10.76
ITAQ Total	14.51	5.90	14.47	5.92	15.59	6.00

Note. PANSS = Positive and Negative Syndrome Scale; CGI = Clinical Global Impression scale; QLS = Quality of Life Scale; ITAQ = Insight into Treatment and Attitudes Questionnaire; WRAT-III = Wide Range Achievement Test-Third Edition.

Table 2. Neuropsychological performance for patients included in baseline analysis ($n = 367$) as well as those who also completed the 12-week neuropsychological assessment ($n = 219$) and provided data for longitudinal analysis of reliability, within-subject effect size, and exploratory factor analysis of domain change

	Admin time (minutes)	Baseline group ($n = 367$) Mean \pm <i>SD</i>	Longitudinal group ($n = 219$)		Domain ICC	Effect size of change
			Baseline Mean \pm <i>SD</i>	12-weeks Mean \pm <i>SD</i>		
CATIE battery						
Processing Speed					.89	.37
Controlled Oral Word Association:						
F,A,S words	3	30.28 \pm 10.80	30.43 \pm 10.82	31.09 \pm 11.86		.07
Category Instances	3	35.19 \pm 10.53	35.87 \pm 11.26	35.60 \pm 11.12		.04
Grooved Pegboard: Two Trial Mean	3	14.90 \pm 3.37	14.92 \pm 3.38	15.76 \pm 3.10		.34
Digit Symbol Test	3	46.49 \pm 13.09	47.51 \pm 13.42	51.21 \pm 14.34		.42
Reasoning and Problem Solving					.61	.23
WCST:	15					
Perseverative Errors		10.15 \pm 7.27	10.49 \pm 7.87	9.34 \pm 7.42		.24
Categories		2.79 \pm 1.56	2.77 \pm 1.56	3.06 \pm 1.63		.19
WISC-III: Mazes	11	20.25 \pm 4.52	20.53 \pm 4.40	21.09 \pm 4.25		.14
Verbal Memory					.72	.19
Hopkins Verbal Learning Test	5	20.75 \pm 5.56	21.19 \pm 5.60	22.17 \pm 5.35		
Working Memory					.80	.34
Comp. Visuospatial Working Memory	14	22.65 \pm 13.20	22.44 \pm 13.72	20.64 \pm 11.81		.17
Test ^a : Mean Error (5, 15 seconds) minus						
No Delay						
Letter–Number Sequencing Test	6	12.20 \pm 3.78	12.36 \pm 3.78	13.41 \pm 3.76		.37
Vigilance					.79	.32
CPT: Identical Pairs: Mean d-prime 3 trials	13	1.93 \pm 0.78	1.93 \pm 0.78	2.11 \pm 0.89		
Social Cognition					.68	.20
Penn Emotion Discrimination Test	10	25.21 \pm 3.14	25.12 \pm 3.32	25.72 \pm 2.84		
CATIE Composite	86				.88	.50
	Admin time (minutes)	Baseline only ($n = 367$)	Longitudinal group ($n = 219$)		Domain ICC	Effect size of change
BACS Battery			Baseline	12-weeks		
Processing Speed					.81	.33
Verbal Fluency: F and S words	2	22.22 \pm 7.49	22.18 \pm 7.45	23.62 \pm 8.75		.19
Token Motor: Total Tokens	2	62.24 \pm 17.10	63.05 \pm 17.12	66.84 \pm 18.05		.23
Symbol Coding: Total Score	3	46.91 \pm 13.38	47.46 \pm 13.24	50.25 \pm 13.46		.28
Reasoning and Problem Solving					.61	.39
Tower of London: Total Score	8	12.77 \pm 4.85	12.78 \pm 4.84	14.34 \pm 4.65		
Verbal Memory					.72	.25
List Learning: Total Words Recalled	8	37.74 \pm 12.10	38.84 \pm 12.33	40.96 \pm 12.19		
Working Memory					.74	.11
Digit Sequence: Correct Responses	8	17.64 \pm 4.96	17.80 \pm 5.10	18.20 \pm 4.99		
BACS Composite	31				.89	.44

Note. CATIE = Clinical Antipsychotic Trials of Intervention Effectiveness; ICC = intraclass correlations; BACS = Brief Assessment of Cognition in Schizophrenia; WCST = Wisconsin Card Sorting Test; WISC-III = Wechsler Intelligence Scale for Children–Third Edition; CPT = Continuous Performance Test.

^aLyons-Warren, Lillie, and Hershey (2004).

baseline and 12-weeks for each of the 10 domain scores provided by the two test batteries.

In the absence of a matched control group, it is difficult to make inferences regarding the level of cognitive deficit in this sample. However, to provide an approximate characterization of the current sample in terms of overall cognitive level, we compared BACS performance to previously published data on a healthy comparison sample (co-normed data are not available for the CATIE battery). Composite scores for the BACS were calculated as the mean of Z scores, separately computed for each measure relative to the mean and standard deviation of the healthy comparison sample used in the BACS validation study (Keefe et al., 2004). Consistent with previous reports characterizing first episode samples relative to healthy comparison samples with other batteries (Bilder et al., 2000; Hill et al., 2004a), the BACS composite indicated moderate overall cognitive impairment for participants who completed baseline ($Z = -1.54 \pm 0.93$) and 12-week ($= -1.49 \pm 0.92$) assessments.

Data Processing and Plans for Analysis

To provide a standard metric for combining test scores into domains and comparing performance over time, test scores were standardized (converted to Z scores) relative to the baseline sample. When necessary, skewed or kurtic distributions were normalized using log [Wisconsin Card Sorting Test (WCST): perseverative errors; Computerized Visuospatial Working Memory Test: mean delay minus no delay error] or cube (Penn Emotion Discrimination Test) transformations before computing Z scores. Scores for each domain were computed as the mean of Z scores within that neurocognitive domain (Saykin et al., 1991).

Decision-Making Processes in Exploratory Factor Analysis

Exploratory rather than confirmatory factor analysis was used because no previous study has evaluated the factor structure of cognitive change scores after antipsychotic treatment. Additionally, exploratory factor analysis is an empirically driven technique that places fewer constraints on the data and maximizes the likelihood of detecting differences in the factor structure of the CATIE and BACS batteries, should differences exist. Exploratory factor analysis has been used for a variety of applications in the social sciences. Based on well-established guidelines (Gorsuch, 1983; Loehlin, 1992) and recent “best practices” recommendations (Costello & Osborne, 2005), a conservative approach to exploratory factor analysis was used to obtain results that are likely to generalize to other samples. The following is a detailed rationale of our decision making with regard to the four major steps in exploratory factor analysis.

Power and Sample Size

Exploratory factor analysis is a large sample procedure typically appropriate for samples greater than 100. Conventional

guidelines recommend samples with subject to variable ratios of 10:1 or greater, while more liberal guidelines state that a ratio of 5:1 may be sufficient in some cases. The most replicable results are obtained with a subject to variable ratio of 20:1 or greater (Costello & Osborne, 2005), and all factor analyses presented in this report exceeded a 20:1 ratio.

Extraction Method

Principal components analysis (PCA), maximum likelihood, and principal axis factoring (PAF) are widely researched (Gorsuch, 1990; Loehlin, 1990) extraction methods. PCA can best be classified as a data reduction technique whose computations are applied without regard to underlying structure caused by latent variables (Loehlin, 1990). Specifically, components are calculated using all variance rather than separating shared and unique variance. Thus, PCA may produce inflated values of explained variance relative to true factor analysis methods (Gorsuch, 1997).

Assumptions regarding normality of multivariate distributions also influenced selection of extraction method, because a small number of CATIE tests required algebraic transformations to normalize distributions (yet all domain scores were normally distributed). Although most factor extraction techniques (i.e., maximum likelihood) are generally robust to non-normally distributed data (Fabrigar et al., 1999), we reported results of principal axis factoring, because it is robust to violations of multivariate normality (Costello & Osborne, 2005). However, to evaluate possible bias resulting from the selected extraction technique, we compared PCA, maximum likelihood, and PAF methods and all three extraction methods yielded similar results.

Number of Factors to Retain

After extraction, one must determine the number of factors to retain. The default in most software packages is the Kaiser criterion, which recommends that all factors with eigenvalues greater than 1.0 be retained. However, this is merely the first step in selecting the number of factors to retain because strict adherence to this guideline is “among the least accurate methods” for selecting a factor solution (Velicer & Jackson, 1990). The scree test better estimates the degree to which keeping/adding factors accounts for variance in the data (Costello & Osborne, 2005). Thus, we used the Kaiser criterion to indicate the maximum number of factors and scree plots to determine whether fewer factors were appropriate. Scree plots are provided (Figures 1 and 2) to illustrate how clearly and consistently a single-factor solution was indicated.

Rotation

Factor rotation was designed to simplify and clarify the data structure when multiple factors exist. Because all exploratory factor analyses indicated a single-factor solution, no rotation was needed.

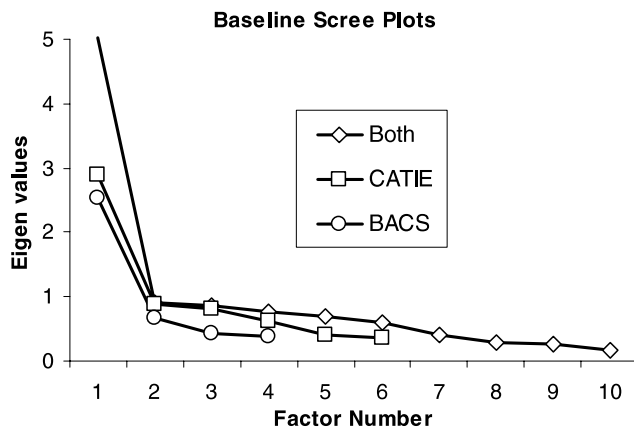


Fig. 1. Scree plots illustrating single factor solutions for baseline cognitive data of CATIE and BACS domains separately and combined

RESULTS

Psychometric Properties of the Two Batteries

Data structure

Psychometric aspects of the data were examined using exploratory factor analysis separately on baseline data for each battery. A single-factor solution was indicated regardless of extraction method. Scree plots of principal axis factoring are presented in Figure 1 for CATIE and BACS domains. Furthermore, according to guidelines regarding the proportion of explained variance (Gorsuch, 1983), a single-factor solution is the only appropriate solution when any factor accounts for more than 40% of the total variance, regardless of the size of additional factors. As can be seen in Table 3, all baseline factor analyses met this criterion as a single factor explained 48%, 63%, and 50% of total vari-

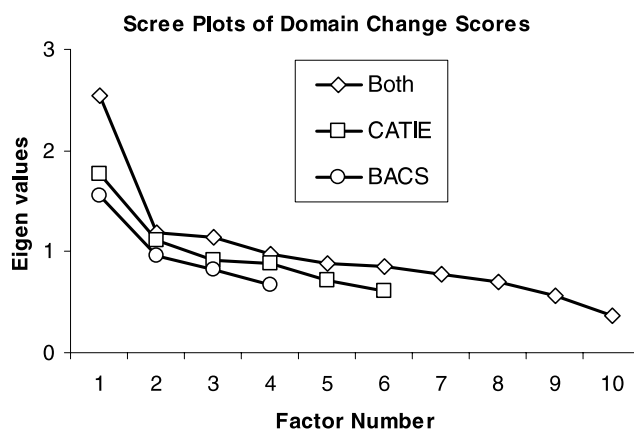


Fig. 2. Scree plots of domain change scores illustrating single factor solutions for treatment-related change in each battery separately and when domain change scores were combined across batteries for factor analysis

ance in baseline CATIE, BACS, and combined data from both batteries, respectively.

Given that data from both batteries were characterized by a single-factor solution, domain scores from both batteries were combined for factor analysis to examine whether a unitary dimension encompassed both batteries, and whether unitary dimensions underlying each battery were relatively independent. Again, a single-factor solution was indicated by scree plots (Figure 1) and percent of variance explained (Table 3), regardless of extraction method. Test scores from both batteries, rather than domain scores, were also submitted to factor analysis to examine whether a single-factor solution was applicable at the test level. Consistent with domain level solutions, and regardless of extraction method, a single-factor solution was indicated. Factor loadings (Table 4) showed that several tests from both batteries had high to medium loadings on the generalized factor, whereas only tests from the CATIE showed low or nonsignificant loadings on the generalized cognitive factor (without independently emerging as unique factors).

The presence of a single factor underlying the baseline cognitive data in both batteries may indicate that a single-composite index is the most appropriate starting point for evaluating cognitive change in treatment studies. However, it is unclear whether the factor structure of change after treatment is comparable to the generalized cognitive factor characterizing baseline performance, or whether the factor structure of cognitive change is multifactorial. This was empirically evaluated by means of exploratory factor analysis of domain change indices, and the findings again indicated single-factor solutions (see Figure 2), regardless of extraction method. When domain change indices from both batteries were considered together, factor analysis again indicated a unitary factor structure. Although the explained variance (25.52–38.64%) was below 40% (Table 5), this may be attributed to the range restriction associated with difference scores and the increased proportion of error variance in the data. Regardless, scree plots show a clear drop in eigen values after the initial factor was extracted from each battery. When the two batteries were combined and factor analyzed, a single-factor solution became more evident, even with three factors exceeding 1.0 eighens. These findings using the CATIE and BACS batteries indicated that both baseline neuropsychological performance and neuropsychological change after treatment with atypical antipsychotic medications were simple in factor structure in the current sample of patients in the early course of schizophrenia, and that both baseline abilities and change following treatment using these batteries are best represented by a general neurocognitive ability factor.

Test-retest reliability

When test-retest reliability is evaluated with intraclass correlations (ICCs) in the context of a treatment study, reliability of measurement can be lowered both by intrinsic unreliability in the measures and also by treatment effects.

Table 3. Principal axis factor loadings for baseline cognitive data of CATIE and BACS domains separately and when domains from both batteries were combined

	Factor loadings		
	CATIE domains	BACS domains	All domains
CATIE			
Processing	.78		.82
Reasoning	.40		.39
Verbal Memory	.62		.65
Working Memory	.80		.80
Vigilance	.70		.68
Social Cognition	.32		.31
BACS			
Processing		.79	.81
Reasoning		.55	.55
Verbal Memory		.75	.77
Working Memory		.76	.77
Statistics			
Explained Variance	48.07%	63.09%	50.21%
Determinant	.20	.28	.01
KMO	.82	.78	.89
Bartlett's Test of Sphericity	583.36, $df = 15$, $p < .001$	468.66, $df = 6$, $p < .001$	1800.38, $df = 45$, $p < .001$

Note. CATIE = Clinical Antipsychotic Trials of Intervention Effectiveness; BACS = Brief Assessment of Cognition in Schizophrenia; KMO = Kaiser-Meyer-Olkin.

Table 4. Principal axis factor loadings of baseline performance for individual tests from both the CATIE and BACS batteries

Test variable	Factor loadings
Digit Symbol Test (C)	.77
List Learning (B)	.76
Letter-Number Sequencing (C)	.76
Digit Sequencing (B)	.76
Symbol Coding (B)	.76
CATIE Verbal Fluency (C)	.72
CPT-Identical Pairs (C)	.69
Hopkins List Learning Test (C)	.64
BACS Verbal Fluency (B)	.58
Grooved Pegboard (C)	.55
Tower of London (B)	.55
Spatial Working Memory (C)	.53
Token Motor Task (B)	.51
WISC-III Mazes (C)	.41
Emotion Discrimination (C)	.31
Wisconsin Card Sorting Test (C)	-.01

Note. 41.09% of Total Variance Explained. Determinant = .001. Kaiser-Meyer-Olkin = .90. Bartlett's Test of Sphericity: 2678.76, $df = 120$, $p < .001$. (C): Test from CATIE battery; (B): test from BACS battery. CATIE = Clinical Antipsychotic Trials of Intervention Effectiveness; CPT = Continuous Performance Test; BACS = Brief Assessment of Cognition in Schizophrenia; WISC-III = Wechsler Intelligence Scale for Children-Third Edition.

However, ICCs still provide a useful estimate of the consistency of performance in composite and domain scores. As can be seen in Table 2, intraclass correlations ranged from .61 to .89, and were generally strong within each battery and across similar domains. These findings are consistent with previously reported BACS test-retest coefficients in schizophrenia patients who had not undergone a change in drug treatment status between evaluations (Keefe et al., 2004).

Sensitivity to change

Effect size estimates were used to assess the sensitivity of each battery to change in test performance after treatment at the composite, domain, and individual test level. Specifically, Cohen's d (Cohen, 1988) was computed by comparing normalized Z scores (not raw score data) for baseline and follow-up. Effect sizes of measured change ranged from small to medium for both tests and domain scores in each battery (see Table 2). This was consistent with modest effect sizes for neuropsychological change reported in meta-analytic studies and larger multisite studies (Harvey et al., 2000; Johnson-Selfridge and Zalewski, 2001; Keefe et al., 1999; Woodward et al., 2005).

Efficiency of detecting change

In terms of the amount of testing needed to derive a meaningful estimate of cognitive abilities in schizophrenia and to detect cognitive change following treatment, one impli-

Table 5. Principal axis factor loadings of domain change scores at the 12-week follow-up for CATIE and BACS domains separately and when domain change scores from both batteries were combined

	Factor loadings		
	CATIE domains	BACS domains	All domains
CATIE			
Processing	.49		.69
Reasoning	.20		.16
Verbal Memory	.31		.27
Working Memory	.54		.42
Vigilance	.57		.50
Social Cognition	.18		.16
BACS			
Processing		.64	.69
Reasoning		.22	.13
Verbal Memory		.51	.47
Working Memory		.33	.37
Statistics			
Explained Variance	29.57%	38.64%	25.52%
Determinant	.69	.82	.29
KMO	.64	.61	.70
Bartlett's Test of Sphericity	80.09; $df = 15$; $p < .001$	42.95; $df = 6$; $p < .001$	266.36; $df = 45$; $p < .001$

Note. CATIE = Clinical Antipsychotic Trials of Intervention Effectiveness; BACS = Brief Assessment of Cognition in Schizophrenia; KMO = Kaiser-Meyer-Olkin.

cation of a generalized deficit model is that a brief battery may be sufficient if it can reliably estimate global composite abilities. To directly compare how efficiently components of each battery predicted global change from baseline to follow-up testing, separate regressions were completed for each battery using a weighted global neuropsychological change index. To reduce the potential for measures with low factor loadings to bias the results, this weighted global composite was empirically guided by the exploratory factor analysis. Specifically, each domain was weighted according to its corresponding single-factor loading before domains from both batteries were combined into a single index of global change and used as the criterion variable. Four domains/predictors were entered for regression analysis of the BACS battery, and six domains/predictors were entered in a separate analysis of the CATIE battery. Predictors were entered one at a time in order of baseline factor loadings. Both batteries explained similar levels of cognitive change in aggregate (CATIE: $R^2 = .74$, $F = 104.02$, $df = 6212$, $p < .001$; BACS: $R^2 = .76$, $F = 166.58$, $df = 4214$, $p < .001$), yet the BACS achieved this in a much shorter period of test administration time. Thus, the extra 50–60 min of administration time for the CATIE battery failed to enhance the prediction of global cognitive change (Figure 3).

DISCUSSION

This study was designed to evaluate, in a large sample of primarily first episode schizophrenia patients, the psychometric characteristics of two prominent neuropsychological

batteries used in the evaluation of cognitive change following antipsychotic treatment. Exploratory factor analysis indicated that a single dimension of generalized cognitive performance underlies pretreatment neuropsychological abilities in early schizophrenia. Cognitive change, as assessed

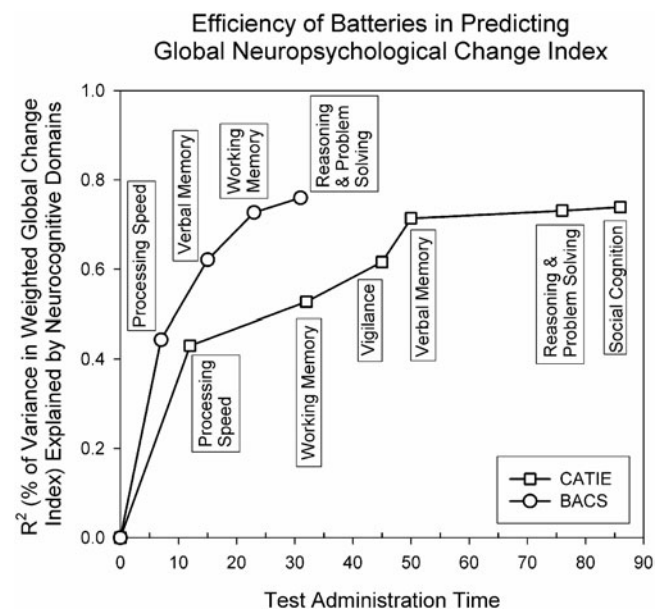


Fig. 3. Efficiency of the CATIE and BACS batteries in accounting for variance in the global neuropsychological change index as a function of administration time

by both batteries, was also characterized by a unitary generalized factor of the neuropsychological tests administered.

The finding of a generalized factor representing baseline deficits and change after treatment suggests that relatively brief neuropsychological assessment batteries may be sufficient to reliably assess global cognitive abilities and change in this generalized dimension after treatment. Indeed, regression analysis evaluating the efficiency for estimating global cognitive change revealed that, while both accounted for a similar portion of total variance in global neuropsychological change, the BACS battery did so in a fraction of the administration time (BACS: 31 min; CATIE: 86 min).

Factor Structure of Neuropsychological Batteries

This is the first study to evaluate neuropsychological constructs underlying change following treatment, and the findings also indicated a unitary factor underlying baseline performance and change in performance after treatment in the BACS and CATIE batteries. The generalized neuropsychological factor observed may reflect the complexity and multidimensional characteristics of many neuropsychological tests, which often evaluate multiple discrete cognitive processes simultaneously. Should the integrity of one component be compromised, impaired performance can occur in multiple tests and the net result is a sensitive but not necessarily specific measure.

With more specific measures, perhaps more directly linked to neurophysiological processes, additional variance in treatment response could be explained and separable factors defined. However, it is unclear whether other neuropsychological batteries, especially large ones, would produce similar findings. However, the present findings distinctly show that no separable group of deficits underlie performance on the CATIE and BACS batteries in the early course of schizophrenia, and data are not yet available to indicate that other approaches will provide a more complex factor structure for cognitive response to antipsychotic drugs. Thus, when evaluating the effect of atypical antipsychotics on the neuropsychological measures widely accepted as reliable and valid indicators of cognitive dysfunction in schizophrenia (Buchanan et al., 2005), a brief battery may be sufficient for estimating the broad cognitive factor underlying cognitive change following treatment.

The current findings are exploratory, by definition, and replication is needed in independent samples using theory driven confirmatory factor analysis. Indeed, before definitive conclusions can be drawn regarding a simple factor structure for neuropsychological abilities in schizophrenia and the impact of antipsychotic treatments, support is needed from both chronic and first episode samples using a broader range of tests. Currently, the literature regarding the factor structure underlying neuropsychological abilities in schizophrenia has produced mixed findings, perhaps related to variation in the use of factor analytic approaches. Multi-

factor models have been supported in schizophrenia using confirmatory factor analysis of the Wechsler Adult Intelligence Scale Revised (WAIS-R, Wechsler, 1981; Allen et al., 1998) and exploratory factor analysis of brief (Keefe et al., 2004) and extended neuropsychological batteries with and without measures of intelligence and memory (Gladysjo et al., 2004; Green et al., 2002; Hobart et al., 1999). However, several studies reporting multifactor solutions in schizophrenia have extracted factors with eigen values less than 1.0 or failed to use the scree test in determining the number of factors to retain (Green et al., 2002; Hobart et al., 1999; Keefe et al., 2004).

There is a preponderance of evidence supporting a unitary dimension underlying a wide range of neuropsychological measures in chronic and early course schizophrenia samples using both exploratory and confirmatory factor analysis. For example, Strauss and Summerfelt (2003) reported that a single factor sufficiently accounted for neuropsychological test performance in schizophrenia patients. When comparing WAIS-III/Wechsler Memory Scale-III (WMS-III) performance in outpatients with schizophrenia and healthy individuals, a single common factor accounted for the majority of patient deficits, and data from specific domains accounted for very little unique between-group variance (Dickinson et al., 2004). Additionally, despite extracting a three-factor solution from a lengthy neuropsychological battery, Green and colleagues argued that a large reliable general factor (accounting for 45% of total variance) justified combining all variables into a single composite to evaluate pharmacological treatment effects (Green et al., 2002). In a confirmatory factor analysis of cognitive data from the CATIE study (1332 schizophrenia spectrum patients), unitary and multifactor models were directly compared and a single-factor model provided a better fit than a five-factor model (Keefe et al., 2006). A principal components analysis of these data also supported a unitary factor, with just one component exceeding 1.0 eigens (Keefe et al., 2006). Similarly, a hierarchical model representing a broad cognitive dimension, rather than a multifactor model of separate latent cognitive factors, was a better fit for performance on individual tests in chronic schizophrenia (Dickinson et al., 2006). Thus, to the extent that neuropsychological measures can inform the nature of neurocognitive dysfunction in schizophrenia, available evidence supports a unitary structure underlying a wide range of neuropsychological measures using a variety of methodologies in well-designed/executed studies with diverse schizophrenia samples. Thus, consistent with findings from the present study, the multifactor models of neuropsychological performance in healthy individuals (Tulsky & Price, 2003) have not generalized to schizophrenia samples. One explanation for this difference may be that disturbances associated with the disorder are similar across higher cognitive abilities, and that the magnitude of this generalized "disease" effect overwhelms the more modest normal independence of various neuropsychological abilities.

Psychometrics and Efficiency

Traditionally, clinical neuropsychology has emphasized comprehensiveness, but brevity becomes important in large clinical trials due to cost and differential attrition (more severely ill patients are less likely to complete long batteries or consistently perform at their ability level). The current findings suggest that shorter batteries may be sufficient to reliably estimate broad neuropsychological ability and cognitive change after treatment with atypical antipsychotic medication in early schizophrenia. The nature of a generalized, unitary factor of neuropsychological abilities, at least as assessed by these two batteries, may partially account for why the shorter BACS battery compared so favorably to the CATIE battery. Direct comparison of reliability, albeit in the context of a clinical trial, revealed good overall reliability for each battery and comparable ICCs among common domains. The benefits of brevity, of course, are meant to apply to research studies where an assumption of generalized deficits seems to adequately characterize data at the group level. This may not be the case for individual patients in a clinical context.

From a psychometric perspective, although the BACS and CATIE batteries had similar reliability and sensitivity to change after treatment, each battery demonstrated some relative strengths and weaknesses. The most salient weakness in the CATIE battery was the WCST. Not only did the WCST composite fail to load robustly on the single-factor solution, but WCST scores failed to emerge as an independent cognitive factor (see Table 4) and none of the WCST variables were particularly sensitive to change after treatment (the collective effect of antipsychotic treatment and practice; Table 2). As a whole, the CATIE: Reasoning and Problem Solving domain (which contains the WCST) produced the lowest domain reliability in the battery, low factor loadings, and poor sensitivity to cognitive change while taking nearly one-third of the battery administration time. Although reasoning and problem solving was the least reliable BACS domain (perhaps illustrating the effects of repeat exposure to problem-solving paradigms in which a single exposure may fundamentally alter subsequent performance, despite alternate forms), the Tower of London loaded moderately on the general cognitive factor and was more sensitive to change effects (effect size = .39) than other reasoning and problem-solving tests.

Recently, social cognitive processing has garnered increased interest in schizophrenia outcome studies (Corcoran, 2001; Kee et al., 2003; Lancaster et al., 2003). To our knowledge, this is the first study to include social cognition in factor analysis of a schizophrenia spectrum sample. Factor loadings and change effect sizes for social cognition were relatively small, suggesting that this measure of social cognition was less sensitive to changes after treatment than was seen with other domains. This finding was consistent with previous reports indicating no significant improvement in emotion perception following antipsychotic treatment of first episode psychosis (Herbener et al., 2005). From an effi-

ciency perspective, evaluation of social cognition added little to the assessment of the global cognition factor underlying the BACS or CATIE or sensitivity to antipsychotic treatment effects of cognition. However, this was the only measure of its kind, and the notion that social cognition might load with a strictly neurocognitive factor may be premature. Indeed, improved sampling of social cognition components is needed to more accurately evaluate the independence of a social cognition factor.

The most salient weakness of the BACS was the poor sensitivity of Digit Sequence to change. Although this is a brief measure with strong reliability and a strong correlation with the general neuropsychological factor, it was relatively insensitive to cognitive change after antipsychotic treatment. In contrast, Letter–Number Sequencing of the CATIE matched Digit Sequencing in terms of factor loadings, administration time, and domain reliability, but provided greater sensitivity to the effects of atypical antipsychotic treatment and practice.

Verbal learning and memory deficits have become a hallmark of schizophrenia research, and an efficient means for evaluating the overall level of cognitive dysfunction in the disorder (Hill et al., 2004b; Paulsen et al., 1995; Saykin et al., 1991). Despite good reliability and moderate to large factor loadings, verbal list learning was only modestly sensitive to change following atypical antipsychotic treatment in both batteries. This was consistent with previous findings of stable verbal memory deficits over time (Hawkins & Wexler, 1999; Hill et al., 2004a; Hoff et al., 1999). Thus, aside from documenting expected deficits in the disorder, evaluation of verbal learning may have limited utility in detecting cognitive effects of atypical antipsychotics.

Limitations

Each instance of factor analysis indicated a single-factor solution, but a large amount of variance (typically ~50%) remained unexplained. Although this is not unusual for a single-factor solution, there remains the possibility that adding tests or domains might help define additional factors. The unitary dimension of neuropsychological performance and change observed in the present study may be limited merely to the domains assessed, the measures used to assess these domains, or the manner in which domains were assessed. Furthermore, it is possible that measures sensitive to multiple independent factors were not adequately assessed by either battery used in the present study, and a more extensive battery with multiple tests of each domain might uncover a more differentiated factor structure. Both batteries used in the present study have a limited number of tests within each domain, and a minimum of three tests per domain is recommended for adequate coverage of multiple latent variables (Kenny et al., 1988). However, one benefit of evaluating factor structure at the test level, using combined data from both batteries, was broader coverage of several domains. These findings also indicated a single dimension (Table 4).

There are potential limitations to the generalizability of the current findings. All patients in the present study were early in the course of illness, thereby reducing the potential effects of chronicity on treatment responsiveness. It is unclear whether the present findings would generalize to chronic patients with more persistent dysfunction, individuals in the prodrome phase, or other diagnostic groups. Also, although effects of the three atypical antipsychotics were similar, generalizability to other treatments cannot be assumed. Finally, prior medication exposure may attenuate treatment effects. That is, Table 1 noted that, however brief, the majority of patients had prior exposure to antipsychotic medication. Although previous medications were tapered, a washout sufficient to clear all antipsychotic drugs before baseline testing was not ethically viable. Thus, prior antipsychotic treatment may have reduced the extent or simplified the factor structure of change after treatment, to a degree. Last, because a placebo control group cannot be used ethically with acutely psychotic first episode patients, change measures at follow-up include influences of both drug and practice effects. This too may have led to an underestimate of multifactorial change in cognitive abilities after treatment.

Practical Implications

The BACS battery demonstrated a distinct advantage in efficiency of assessing global cognitive treatment outcome over the CATIE battery. It accounted for a similar proportion of global change in generalized cognitive performance after treatment in approximately one-third the administration time with a minimal cost in sensitivity to aggregate antipsychotic effect on measured cognition (BACS ES = .44, CATIE ES = .50). Rather than inherent flaws with the CATIE battery, this finding may simply reflect the relative ease of reliably and validly assessing a generalized dimension with fewer tests, and the limit in incremental knowledge provided by additional test data in this population.

Shorter batteries such as the BACS may provide an adequate estimate of generalized cognitive deficits in studies of antipsychotic treatment on functionally important neuropsychological deficits. The multifactorial approach of the MATRICS consensus battery (Neuchterlein et al., 2004), which aims to independently assess six cognitive domains, may not be necessary for evaluating antipsychotic effects on cognition. Further studies are needed to fully demonstrate the utility of such larger test batteries in assessing cognitive outcomes relative to the brief batteries. Multifactor approaches may prove to be a crucial strategy for drug evaluation, particularly if potential procognitive adjunctive treatments are predicted to have effects on specific receptor systems and functional circuits as well as the cognitive abilities they support. For example, if a new nicotinic agent primarily improves attention, then multiple tests of attention may be more useful in assessing change than tests of general cognitive ability. Change in attention may be more pronounced than change in other domains, and then latent

factor structure of change after treatment may be different than with antipsychotic treatments. Adequate assessment of domains targeted by new treatments will be crucial components to cognitive batteries evaluating possible differential effects of adjunctive treatments in the context of clinical trials.

ACKNOWLEDGMENTS

The CAFE study was funded by AstraZeneca Pharmaceuticals LP and coordinated by the University of North Carolina. The University of Illinois at Chicago participated in the study as a performance site, with Dr. Sweeney as the site PI. Drs. Hill and Sweeney have no conflicts of interest. Although Dr. Keefe has a conflict of interest with respect to royalties from the BACS battery, his contributions to the study design and coordination warrant authorship. Dr. Keefe was included as an author, but he was not involved in data analysis or interpretation.

REFERENCES

- Allen, D.N., Huegel, S.G., Seaton, B.E., Goldstein, G., Gurklis, J.A., & van Kammen, D.P. (1998). Confirmatory factor analysis of the WAIS-R in patients with schizophrenia. *Schizophrenia Research*, *34*, 87–94.
- American Psychiatric Association. (1987/1994). *Diagnostic and statistical manual* (3rd ed., revised/4th ed.). Washington, DC: APA.
- Bilder, R.M., Goldman, R.S., Robinson, D., Reiter, G., Bell, L., Bates, J.A., Pappadopulos, E., Willson, D.F., Alvir, J.M.J., Woerner, M.G., Geisler, S., Kane, J.M., & Lieberman, J.A. (2000). Neuropsychology of first-episode schizophrenia: Initial characterization and clinical correlates. *American Journal of Psychiatry*, *157*, 549–559.
- Bilder, R.M., Goldman, R.S., Volavka, J., Czobor, P., Hoptman, M., Sheitman, B., Lindenmayer, J., Citrome, L., McEvoy, J., Kunz, M., Chakos, M., Cooper, T.B., Horowitz, T.L., & Lieberman, J.A. (2002). Neurocognitive effects of clozapine, olanzapine, risperidone, and haloperidol in patients with chronic schizophrenia or schizoaffective disorder. *American Journal of Psychiatry*, *159*, 1018–1028.
- Blanchard, J.J. & Neale, J.M. (1994). The neuropsychological signature of schizophrenia: Generalized or differential deficit? *American Journal of Psychiatry*, *151*, 40–48.
- Buchanan, R.W., Davis, M., Goff, D., Green, M.F., Keefe, R.S.E., Leon, A.C., Neuchterlein, K.H., Laughren, T., Levin, R., Stover, E., Fenton, W., & Marder, S.R. (2005). A summary of the FDA-NIMH-MATRICES workshop on clinical trial design for neurocognitive drugs for schizophrenia. *Schizophrenia Bulletin*, *31*, 5–19.
- Buchanan, R.W., Koepl, P., & Breier, A. (1994). Stability of neurological signs with clozapine treatment. *Biological Psychiatry*, *36*, 198–200.
- Cassens, G., Inglis, A.K., Appelbaum, P.S., & Gutheil, T.G. (1990). Neuroleptics: Effects on neuropsychological function in chronic schizophrenia. *Schizophrenia Bulletin*, *16*, 477–499.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum.
- Corcoran, R. (2001). Theory of mind and schizophrenia. In P.W. Corrigan & D.L. Penn (Eds.), *Social cognition and schizophre-*

- nia (1st ed., pp. 149–174). Washington, DC: American Psychological Association.
- Costello, A.B. & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, & Evaluation, 10*, 1–9.
- Dickinson, D., Iannone, V.N., Wilk, C.M., & Gold, J.M. (2004). General and specific cognitive deficits in schizophrenia. *Biological Psychiatry, 55*, 826–833.
- Dickinson, D., Ragland, J.D., Calkins, M.E., Gold, J.M., & Gur, R.E. (2006). A comparison of cognitive structure in schizophrenia patients and healthy controls using confirmatory factor analysis. *Schizophrenia Research, 85*, 20–29.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272–299.
- First, M.D., Gibbon, G.E., Spitzer, R.L., & Williams, J.B.W. (1995). *Structured clinical interview for DSM-IV Axis I disorders—patient edition*. New York: Biometrics Research Department, NYSPI.
- Gladysjo, J.A., McAdams, L.A., Palmer, B.W., Moore, D.J., Jeste, D.V., & Heaton, R.K. (2004). A six-factor model of cognition in schizophrenia and related psychotic disorders: Relationships with clinical symptoms and functional capacity. *Schizophrenia Bulletin, 30*, 739–754.
- Gold, J.M., Randolph, C., Carpenter, C.J., Goldberg, T.E., & Weinberger, D.R. (1992). Forms of memory failure in schizophrenia. *Journal of Abnormal Psychology, 101*, 487–494.
- Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Gorsuch, R.L. (1990). Common factor analysis versus component analysis—some well and little known facts. *Multivariate Behavioral Research, 25*, 33–39.
- Gorsuch, R.L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68*, 532–560.
- Green, M.F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry, 153*, 321–330.
- Green, M.F., Marder, S.R., Glynn, S.M., McGurk, S.R., Wirshing, W.C., Wirshing, D.A., Liberman, R.P., & Mintz, J. (2002). The neurocognitive effects of low-dose haloperidol: A two-year comparison with risperidone. *Biological Psychiatry, 51*, 972–978.
- Guy, W. (1976). *Early Clinical Drug Evaluation (ECDEU), assessment manual*. Rockville, MD: National Institute of Mental Health.
- Harvey, P.D., Moriarty, P.J., Serper, M.R., Schnur, E., & Lieber, D. (2000). Practice-related improvement in information processing with novel antipsychotic treatment. *Schizophrenia Bulletin, 46*, 139–148.
- Hawkins, K.A. & Wexler, B.E. (1999). California Verbal Learning Test practice effects in a schizophrenia sample. *Schizophrenia Research, 23*, 73–78.
- Herbener, E.S., Hill, S.K., Marvin, R.W., & Sweeney, J.A. (2005). Effects of antipsychotic treatment on emotion perception deficits in first episode schizophrenia. *American Journal of Psychiatry, 162*, 1746–1748.
- Hill, S.K., Beers, S.R., Kmiec, J.A., Keshavan, M.S., & Sweeney, J.A. (2004b). Impairment of verbal memory and learning in antipsychotic-naïve patients with first-episode schizophrenia. *Schizophrenia Research, 68*, 127–136.
- Hill, S.K., Ragland, J.D., Gur, R.C., & Gur, R.E. (2001). Neuropsychological differences among empirically derived clinical subtypes of schizophrenia. *Neuropsychology, 15*, 492–501.
- Hill, S.K., Schuepbach, D., Herbener, E.S., Keshavan, M.S., & Sweeney, J.A. (2004a). Pretreatment and Longitudinal studies of neuropsychological deficits in antipsychotic-naïve patients with schizophrenia. *Schizophrenia Research, 68*, 49–63.
- Hobart, M.P., Goldberg, R., Bartko, J.J., & Gold, J.M. (1999). Repeatable battery for the assessment of neuropsychological status as a screening test in schizophrenia: II. Convergent/discriminant validity and diagnostic group comparisons. *American Journal of Psychiatry, 156*, 1951–1957.
- Hoff, A.L., Sakuma, M., Wieneke, M., Horon, R., Kushner, M., & DeLisi, L.E. (1999). Longitudinal neuropsychological follow-up study of patients with first-episode schizophrenia. *American Journal of Psychiatry, 156*, 1336–1341.
- Johnson-Selfridge, M. & Zalewski, C. (2001). Moderator variables of executive functioning in schizophrenia: Meta-analytic findings. *Schizophrenia Bulletin, 27*, 305–316.
- Kay, S.R., Fiszbein, A., & Opler, L.A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*, 261–276.
- Kee, K.S., Green, M.F., Mintz, J., & Brekke, J.S. (2003). Is emotional processing a predictor of functional outcome in schizophrenia? *Schizophrenia Bulletin, 23*, 487–497.
- Keefe, R.S.E., Bilder, R.M., Harvey, P.D., Davis, S.M., Palmer, B.W., Gold, J.M., Meltzer, H.Y., Green, M.F., Miller, D.D., Canive, J.M., Adler, L.W., Marder, T.C., Swartz, M., Rosenheck, R., Perkins, D.O., Walker, T.M., Stroup, T.S., McEvoy, J.P., & Lieberman, J.A. (2006). Baseline neurocognitive deficits in the CATIE schizophrenia trial. *Neuropsychopharmacology, 31*, 2033–2046.
- Keefe, R.S.E., Goldberg, T.E., Harvey, P.D., Gold, J.M., Poe, M., & Coughenour, L. (2004). The brief assessment of cognition in schizophrenia: Reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research, 68*, 283–297.
- Keefe, R.S.E., Mohs, R.C., Bilder, R.M., Harvey, P.D., Green, M.F., Meltzer, H.Y., Gold, J.M., & Sano, M. (2003). Neurocognitive assessment in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project schizophrenia trial: Development, methodology and rationale. *Schizophrenia Bulletin, 29*, 45–55.
- Keefe, R.S.E., Perkins, S., Silva, S.M., & Lieberman, J.A. (1999). The effect of atypical antipsychotic drugs on neurocognitive function in schizophrenia. *Schizophrenia Bulletin, 25*, 201–222.
- Keefe, R.S.E., Sweeney, J.A., Gu, H., Hamer, R.M., Perkins, D.O., McEvoy, J.P., & Lieberman, J.A. (2007). A comparison of the effects of olanzapine, quetiapine, and risperidone on neurocognitive function in first-episode psychosis: A randomized, double-blind clinical trial. *American Journal of Psychiatry, 164*, 1061–1071.
- Kenny, D.A., Kashy, D.A., & Bolger, N. (1988). Data analysis in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 233–265). Boston: McGraw-Hill.
- Kern, R., Green, M., Neuchterlein, K., & Deng, B. (2004). NIMH-MATRICES survey on assessment of neurocognition in schizophrenia. *Schizophrenia Research, 72*, 11–19.
- Lancaster, R.S., Evans, J.D., Bond, G.R., & Lysaker, P.H. (2003). Social cognition and neurocognitive deficits in schizophrenia. *Journal of Nervous and Mental Disease, 191*, 295–299.
- Loehlin, J.C. (1990). Component analysis versus common factor analysis—a case of disputed authorship. *Multivariate Behavioral Research, 25*, 29–31.

- Loehlin, J.C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis* (2nd ed.) New Jersey: Lawrence Erlbaum.
- Lyons-Warren, A., Lillie, R., & Hershey, T. (2004). Short and long-term spatial delayed response performance across the lifespan. *Developmental Neuropsychology*, 26, 661–678.
- McEvoy, J.P., Lieberman, J.A., Perkins, D.O., Hamer, R.M., Gu, H., Lazarus, A., Sweitzer, D., Olexy, C., Weiden, P., & Strakowski, S.M. (2007). Efficacy and tolerability of olanzapine, quetiapine, and risperidone in the treatment of early psychosis: A randomized, double-blind, 52-week comparison. *American Journal of Psychiatry*, 164, 1050–1060.
- Neuchterlein, 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.
- Paulsen, J.S., Heaton, R.K., Sadek, J.R., Perry, W., Delis, D.C., Braff, D., Kuck, J., Zisook, S., & Jeste, D.V. (1995). The nature of learning and memory impairments in schizophrenia. *Journal of the International Neuropsychological Society*, 1, 88–99.
- Purdon, S.E., Jones, B.D., Stip, E., Labelle, A., Addington, D., David, S.R., Breier, A., & Tollefson, G.D. (2000). Neuropsychological change in early phase schizophrenia during 12 months of treatment with olanzapine, risperidone, or haloperidol. *Archives of General Psychiatry*, 57, 249–258.
- Rollnik, J.D., Borsutzky, M., Huber, T.J., Mogk, H., Seifert, J., Emrich, H.M., & Schneider, U. (2002). Short-term cognitive improvement in schizophrenics treated with typical and atypical neuroleptics. *Neuropsychobiology*, 45, 74–80.
- Saykin, A.J., Gur, R.C., Gur, R.E., Mozley, P.D., Mozley, L.H., Resnick, S.M., Kester, B., & Stafiniak, P. (1991). Neuropsychological function in schizophrenia: Selective impairment in memory and learning. *Archives of General Psychiatry*, 48, 618–624.
- Strauss, M.E. & Summerfelt, A. (2003). The neuropsychological study of schizophrenia: A methodological perspective. In M.F. Lenzenweger & J.M. Hooley (Eds.), *Principles of experimental psychopathology: Essays in honor of Brendan A. Maher* (pp. 119–134). Washington, DC: American Psychological Association.
- Tulsky, D.S. & Price, L.R. (2003). The joint WAIS-III and WMS-III factor structure: Development and cross-validation of a six-factor model of cognitive functioning. *Psychological Assessment*, 15, 149–162.
- Velicer, W.F. & Jackson, D.N. (1990). Component analysis versus common factor-analysis—some further observations. *Multivariate Behavioral Research*, 25, 97–14.
- Woodward, N.D., Purdon, S.E., Meltzer, H.Y., & Zald, D.H. (2005). A meta-analysis of neuropsychological change to clozapine, olanzapine, quetiapine, and risperidone in schizophrenia. *International Journal of Neuropsychopharmacology*, 8, 1–16.
- Wechsler, D.A. (1981). *Wechsler Adult Intelligence Scale Revised (WAIS-R), manual*. Cleveland, OH: The Psychological Corporation.