<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Refining the latent structure of neuropsychological performance in schizophrenia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Lam, M.; Collinson, S. L.; Eng, G. K.; Rapisarda, A.; Kraus, M.; Lee, J.; Chong, S. A.; Keefe, R. S. E.</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Lam, M., Collinson, S. L., Eng, G. K., Rapisarda, A., Kraus, M., Lee, J., et al. (2014). Refining the latent structure of neuropsychological performance in schizophrenia. Psychological Medicine, 44(16), 3557-3570.</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2014</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/20006">http://hdl.handle.net/10220/20006</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2014 Cambridge University Press. This paper was published in Psychological Medicine and is made available as an electronic reprint (preprint) with permission of Cambridge University Press. The paper can be found at the following official DOI:<a href="http://dx.doi.org/10.1017/S0033291714001020">http://dx.doi.org/10.1017/S0033291714001020</a>. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.</td>
</tr>
</tbody>
</table>
Refining the latent structure of neuropsychological performance in schizophrenia

M. Lam1, S. L. Collinson2, G. K. Eng1,7, A. Rapisarda1,3, M. Kraus4, J. Lee1,5,6, S. A. Chong1 and R. S. E. Keefe4,*

1Research Division, Institute of Mental Health, Singapore
2Department of Psychology, National University of Singapore, Singapore
3Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore
4Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA
5Department of General Psychiatry I, Institute of Mental Health, Singapore
6Office of Clinical Sciences, Duke-NUS Graduate Medical School, Singapore
7Division of Psychology, School of Humanities and Social Sciences, Nanyang Technological University, Singapore

Background. Elucidating the cognitive architecture of schizophrenia promises to advance understanding of the clinical and biological substrates of the illness. Traditional cross-sectional neuropsychological approaches differentiate impaired from normal cognitive abilities but are limited in their ability to determine latent substructure. The current study examined the latent architecture of abnormal cognition in schizophrenia via a systematic approach.

Method. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were carried out on a large neuropsychological dataset including the Brief Assessment of Cognition in Schizophrenia, Continuous Performance Test, Wisconsin Card Sorting Test, Benton Judgment of Line Orientation Test, and Wechsler Abbreviated Scale of Intelligence matrix reasoning derived from 1012 English-speaking ethnic Chinese healthy controls and 707 schizophrenia cases recruited from in- and out-patient clinics.

Results. An initial six-factor model fit cognitive data in healthy and schizophrenia subjects. Further modeling, which accounted for methodological variance between tests, resulted in a three-factor model of executive functioning, vigilance/speed of processing and memory that appeared to best discriminate schizophrenia cases from controls. Factor analytic-derived $g$ estimands and conventionally calculated $g$ showed similar case–control discrimination. However, agreement analysis suggested systematic differences between both $g$ indices.

Conclusions. Factor structures derived in the current study were broadly similar to those reported previously. However, factor structures between schizophrenia subjects and healthy controls were different. Roles of factor analytic-derived $g$ estimands and conventional composite score $g$ were further discussed. Cognitive structures underlying cognitive deficits in schizophrenia may prove useful for interrogating biological substrates and enriching effect sizes for subsequent work.

Received 18 September 2013; Revised 21 March 2014; Accepted 26 March 2014

Key words: Bland–Altman analysis, confirmatory factor analysis, endophenotypes, exploratory factor analysis, $g$.

Introduction

Cognitive deficits are well recognized in schizophrenia (Saykin et al. 1991; Harvey & Keefe, 1997; Heinrichs & Zakzanis, 1998; Aleman et al. 1999; Bokat & Goldberg, 2003; Harvey et al. 2003, 2004; Henry & Crawford, 2005; Lee & Park, 2005; Keele et al. 2006a, b; Keshavan et al. 2008; Szöke et al. 2008; Mesholam-Gately et al. 2009) and are hypothesized to be a more direct expression of underlying biological abnormalities than formal diagnosis (Harvey & Keefe, 1997; Heinrichs & Zakzanis, 1998). More refined neuropsychological measures could facilitate better understanding of the illness, with regards to how neurobiological susceptibility progresses into a disease phenotype (Gottesman & Gould, 2003; Heinrichs, 2005; Cannon & Keller, 2006; Keshavan et al. 2008; Prasad & Keshavan, 2008).

Factor analytic approaches (Holdnack et al. 2011) can be employed to determine the cognitive architecture of schizophrenia derived from neuropsychological tests (Wechsler, 1945, 1955; Keefe et al. 2006c; Kern et al. 2008; Nuechterlein et al. 2008). Speed of processing, attention/vigilance, working memory, verbal learning and memory, visual learning and memory and reasoning/problem solving have been identified via factorial approaches as areas of abnormality in schizophrenia (Nuechterlein et al. 2008). A review of several
confirmatory factor analysis (CFA) studies highlighted several intriguing trends (online Supplementary Table S1):

(1) Subtle differences in cognitive architecture have been reported; separation of extracted factors may be biased by test-specific features (e.g. tests that measure reaction time may form strong covariance with other tests that are reaction time based; see Podsakoff et al. 2003), leading to factor structures that may not entirely reflect cognitive substructure.

(2) The correlated and hierarchical factor models have been commonly tested. The former, a first-order CFA model, indicates associated but separate cognitive domains; the latter, a second-order CFA model, assumes that a single general cognitive factor g subserves all cognitive domains. The use of either model influences interpretation of the derived cognitive architecture.

(3) Most studies suggest that cognitive substructures in schizophrenia and control samples (mostly healthy participants) are qualitatively homologous (see online Supplementary Table S1). Only one study in the review (Dickinson et al. 2006) reported non-invariance in the cognitive structure of healthy controls and schizophrenia. Further investigation is necessary. Homology in cognitive structures appears to run contrary to the broad and profound cognitive impairments observed in schizophrenia.

The present study aims to: (i) define the latent cognitive architecture in schizophrenia cases and healthy controls; (ii) examine covariances across neuropsychological tests, to broadly compare latent factors against published evidence; (iii) refine and reduce the latent architecture to fit the data without methodological variance; and (iv) establish a final model that best discriminates cases and controls. The secondary objectives of the current study are: (i) establish agreement between factor analytic g and conventional methods of g calculation (or composite g) – averaging standardized test scores with pooled standard deviations; (ii) establish discriminability of factor analysis-derived g and composite g; and (iii) establish convergent validity between g estimands and education as a candidate reference.

Method

Participants

A total of 1012 healthy participants (controls) and 707 schizophrenia cases were recruited as part of the Singapore Translational and Clinical Research in Psychosis (project title: Elucidating the Genetic Architecture of Neurocognitive Endophenotypes in Schizophrenia; grant no. NMRC/TCR/003/2008). Healthy controls were recruited from the community while schizophrenia cases were recruited from rehabilitation centers, community care centers across the country, out-patient clinics and in-patient wards, under purview of the Institute of Mental Health, Singapore. Data collection was completed in approximately 3 years. Inclusion criteria were: Chinese ethnicity, to ensure a genetically homogeneous sample; and completion of a minimum of 6 years of primary school education. Additional exclusion criteria precluded all participants with significant history of substance abuse, clinically significant neurological disease or injury, color blindness, and healthy participants with first-degree relatives suffering from schizophrenia or other psychotic disorders. Schizophrenia cases fulfilled Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) diagnostic criteria for schizophrenia based on the Structured Clinical Interview for DSM-IV Axis I disorders (First et al. 2002). All participants consented to participate in research procedures prior to data collection. Consent procedures adhered to the guidelines specified by the National Healthcare Group Domain Specific Review Board’s (domain A, NHG DSRB) requirements for human subject research.

Procedures

Neuropsychological and clinical evaluations were carried out by psychometricians (trained by R.S.E.K., S.L.C., M.K. and A.R.), Wechsler Abbreviated Scale of Intelligence (WASI) matrix reasoning (Wechsler, 1999), Continuous Performance Tests-Identical Pairs (CPT-IP; Cornblatt et al. 1988), Wisconsin Card Sorting Test, 64-card version (WCST-64; Heaton, 1993), Benton Judgment of Line Orientation Test (Benton et al. 1994) and the Brief Assessment of Cognition in Schizophrenia (BACS; Keefe et al. 2004, 2008), consisting of verbal memory, digit sequencing, token motor task, semantic fluency, symbol coding and Tower of London, were administered to all participants. A total of 32 subtests were obtained from the current battery of 10 neuropsychological tests (e.g. within semantic fluency, there were three subtests – animals, fruits and vegetables; see Fig. 1 for list of all subtests). Data including age, gender, education, duration of illness, medications and Positive and Negative Syndrome Scale ratings were also collected.

Data analysis

Data preparation and analysis software

All neuropsychological subtests were corrected for age and gender, normalized via Blom inverse rank transformation (Blom, 1958) and standardized against
healthy controls. IBM SPSS version 20 (IBM, USA) and AMOS™ version 18 (Amos Development Corp., USA; Albright & Hun, 2009) were used for analyses (see online Supplementary Fig. S5 for analysis flowchart).

Examining covariance across test performances

Data-driven EFA procedures were employed to examine patterns of covariances across neuropsychological test performances. This was to investigate if neuropsychological tests were indeed approximating known cognitive domains. No assumptions were made regarding the covariance structure of neuropsychological tests. All subtests from the neuropsychological battery were entered in an EFA model, using principal components extraction and varimax rotation. Subtests with the highest factor loadings and communalities × factor loading indices were selected for further CFA.

Addressing methodological variances in factor analysis

Methodological variances have been flagged as a challenge in the schizophrenia literature to factor analysis of neuropsychological tests (Podsakoff et al. 2003; Genderson et al. 2007; Dickinson & Gold, 2008; Dickinson & Harvey, 2009). Within the 32 neuropsychological subtests, 12 were CPT-IP subtests, three were semantic fluency subtests, and five were WCST-64 subtests (see online Supplementary Table S2 for a list of subtests). These subtests from the same tests are expected to demonstrate strong covariance. In a separate analysis, each block of subtests from CPT-IP, WCST-64 and semantic fluency were reduced via principal component analysis (PCA). Derived regression factors scores were then used for subsequent EFAs and CFAs.

Evaluation of model types

Correlated and hierarchical models were tested in cases and controls, respectively. Subtests that were identified in earlier EFA were entered in the models. Both unrefined and refined models were tested.

Model fitting for derived factors

CFA model fit was evaluated. CFA fit indices included the normed fit index (NFI; Bentler & Bonnett, 1980), relative fit index (RFI; Bollen, 1986), incremental fit index (Bollen, 1989), non-normed fit index/Tucker–Lewis index (Bollen, 1989), comparative fit index (Bentler, 1990) and root mean square of approximation (RMSEA; Browne & Cudeck, 1993). A good model fit is indicated by a RMSEA <0.05 and >0.9 for the other indices (Bentler & Bonnett, 1980; Bollen, 1986, 1989; Bentler, 1990; Browne & Cudeck, 1993). CFA factor scores were generated via full information
maximum likelihood (FIML) imputations for subsequent investigations. CFA factor scores were generated via FIML imputations for subsequent investigations.

**Evaluation of model discrimination of case–control status**

True factor scores for latent models were entered as predictors in a logistic regression model. Case–control status was entered as a dependent variable. Discriminability of factor analytic estimation of $g$ and composite $g$ was also evaluated.

**Agreement of composite $g$ and factor-derived $g$**

Use of $g$ composites from test batteries as a cognitive phenotype has been previously discussed in the literature (Dickinson et al. 2013; Donohoe et al. 2013). The conventional method of deriving $g$ or general cognitive composite from a neuropsychological battery has been to average tests scores by pooled standard deviations. The sum scores approach is putatively desirable when computation is considered exploratory, wherein each item is weighted equally (DiStefano et al. 2009; Hair & Anderson, 2010), whereas factor analytic-derived $g$ takes into account factor weights and correlations specific to the sample with which the analysis is performed (DiStefano et al. 2009). However, it is not known if factor-derived $g$ and conventional $g$ calculation approaches are equivalent. Post-hoc Bland–Altman analysis (Bland & Altman, 1986) was performed to examine level of agreements between factor analysis-derived latent factors and composite $g$.

**Evaluation of convergent validity of factor-derived $g$**

Convergent validity of factor-derived $g$ and composite $g$ was examined alongside education attainment. Education is a candidate measure for establishing convergent validity for $g$ due to previously known associations with cognitive performances. Bivariate correlation was performed for education attainment and factor analysis-derived $g$ and composite score-derived $g$. Education attainment was estimated by adjusted years depending on the stage of education that the participants were in (for more in-depth discussion, see Lam et al. 2012).

**Results**

**Sample description**

Sample characteristics are reported in Table 1. Scaled and adjusted neuropsychological profiles of schizophrenia cases are presented in Fig. 1. Cognitive scores in schizophrenia cases ranged between 1 and 2 standard deviations from controls.

**Examining covariance across test performance**

BACS, CPT-IP and WCST-64 subtests, WASI matrix reasoning and Benton Judgment of Line Orientation Test, 32 subtests in total, were subjected to exploratory factor analysis (EFA). Subtests with factor communalities $<0.4$ and equal to 1 were removed. These may reflect low reliability or collinearity. This resulted in the exclusion of BACS token motor task, WCST-64 correct responses, BACS semantic fluency (total), CPT-IP
average hits, CPT-IP average $d'$, CPT-IP average random errors and CPT-IP average false alarms. EFA was repeated on the remaining 24 items. The CPT-IP two-digit subtest was excluded from subsequent analysis as it was deemed non-specific to factor solutions (see online Supplementary Table S2).

Six factors were extracted from EFA in both samples. Subtests were selected based on breakpoints on factor loadings, factor loadings × communalities indices and judgment of item relevance for each modeled factor for subsequent CFA model fitting (online Supplementary Table S2 and Fig. S1). Subtests selected were: (i) factor 1: CPT-IP three- and four-digit hits; (ii) factor 2: WASI matrix reasoning and BACS Tower of London; (iii) factor 3: WCST-64 perseverative errors, completed categories and first category scores; (iv) factor 4: BACS semantic fluency, animals, fruits and vegetables; (v) factor 5: CPT-IP three- and four-digit random errors; and (vi) factor 6: CPT-IP three- and four-digit false alarms.

Addressing methodological variances in factor analysis

BACS semantic fluency animals, fruits and vegetables; WCST-64 perseverative errors, completed categories, final category responses, first category responses, correct responses; and CPT-IP two-, three-, four-digits, and average hits, false alarms, random errors and $d'$ were subjected to reduction procedures for each set of subtests. PCA extraction was employed to establish component scores for the entire case-control sample. Four reduced components were obtained for CPT-IP, and one reduced component each for BACS semantic fluency and WCST-64 subtests. The first CPT-IP reduced component corresponded to CPT-IP hits and $d'$, the second CPT-IP reduced component corresponded to CPT-IP random errors, and the third and fourth CPT-IP reduced components corresponded to three- and four-digit false alarms and two-digit false alarms, respectively (online Supplementary Table S3).

Five subtests from the BACS battery (verbal memory, digit sequencing, token motor task, symbol coding, and Tower of London), the Benton Judgment of Line Orientation Test and WASI matrix reasoning, four reduced CPT-IP factors, and one reduced factor each from the BACS semantic fluency tests, and the WCST-64 were entered into separate case-control EFA models. Four- and three-factor solutions were obtained in healthy controls and schizophrenia cases, respectively (online Supplementary Table S4). Patterns of item loadings suggested that factor loadings differed in schizophrenia. Most cognitive subtests loaded on one single factor in schizophrenia cases, indicating high congruity of cognitive performances across tests. Due to interpretation challenges in attempting to further deconstruct a single-factor solution in schizophrenia, EFA solutions obtained from controls were referenced for test selection in subsequent CFA modeling in schizophrenia cases as well. Using similar methods as in earlier EFA procedures, the following items were selected for subsequent CFA: (i) solution 1: Benton Judgment of Line Orientation Test and WASI matrix reasoning; (ii) solution 2: BAC semantic fluency – reduced component and BACS verbal memory; (iii) solution 3: CPT-IP three- and four-digit hits – reduced component and CPT-IP three- and four-digit false alarms – reduced component; (iv) solution 4: BACS symbol coding and BACS token motor task (online Supplementary Table S4 and Fig. S2). The term ‘solution’ is used for disambiguating ‘factors’ from the six-factor EFA.

Evaluation of model types

Derived from previous EFA procedures, six, four and three correlated and hierarchical CFA factor models were built. Models were applied to both schizophrenia and healthy controls separately; results of CFA are reported in Tables 2 and 3.

CFA model fitting for derived factors

Six-factor correlated and hierarchical models appeared to fit both case and control data at a reasonable level. The four-factor correlated model was found to fit cognitive data in the healthy controls, but did not converge in schizophrenia. The structural model for schizophrenia was minimally re-specified; the CPT-IP false alarm reduced component was removed, while the CPT-IP hits reduced component was modeled with the BACS token motor task and BACS symbol coding, obtaining a three-factor correlated model for schizophrenia. After re-specification, the three correlated and hierarchical factor models were found to fit cognitive data in schizophrenia. To facilitate subsequent sample comparisons, both three-factor models were tested on the entire sample of cases and controls. Good model fit was found when the case models were applied to all the subjects (see model fit indices; Table 3). Future work is required to validate the executive function latent factor conceptualized by matrix reasoning and line orientation. These tasks are not typically utilized for assessment of executive function. However, the nomenclature for the latent variable was assigned as such due to the initial loading of these tasks with the WCST-64 and Tower of London, suggesting that variances in these tasks were not separable from executive function. Matrix reasoning and line orientation appeared to have higher loadings in the
Table 2. Model fitting of six-, four- and three-factor CFA models

<table>
<thead>
<tr>
<th>Six-factor models</th>
<th>HC-6 CFM</th>
<th>HC-6 HFM</th>
<th>SCZ-6 CFM</th>
<th>SCZ-6 HFM</th>
<th>Factor labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 -&gt; CPT-IP three-digit hits</td>
<td>0.868</td>
<td>0.902</td>
<td>0.844</td>
<td>0.947</td>
<td>CPT hits</td>
</tr>
<tr>
<td>F1 -&gt; CPT-IP four-digit hits</td>
<td>0.727</td>
<td>0.858</td>
<td>0.727</td>
<td>0.878</td>
<td></td>
</tr>
<tr>
<td>F2 -&gt; BACS Tower of London</td>
<td>0.552</td>
<td>0.672</td>
<td>0.580</td>
<td>0.670</td>
<td>Executive function/spatial</td>
</tr>
<tr>
<td>F2 -&gt; WASI matrix reasoning</td>
<td>0.699</td>
<td>0.793</td>
<td>0.665</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td>F3 -&gt; WCST-64 perseverative errors</td>
<td>-0.909</td>
<td>-1.005</td>
<td>-0.906</td>
<td>-1.006</td>
<td>WCST-64</td>
</tr>
<tr>
<td>F3 -&gt; WCST-64 completed categories</td>
<td>0.763</td>
<td>0.799</td>
<td>0.765</td>
<td>0.798</td>
<td></td>
</tr>
<tr>
<td>F4 -&gt; BACS semantic fluency (animals)</td>
<td>0.655</td>
<td>0.731</td>
<td>0.656</td>
<td>0.728</td>
<td></td>
</tr>
<tr>
<td>F4 -&gt; BACS semantic fluency (vegetables)</td>
<td>0.617</td>
<td>0.709</td>
<td>0.619</td>
<td>0.708</td>
<td>Semantic fluency</td>
</tr>
<tr>
<td>F5 -&gt; CPT-IP three-digit random errors</td>
<td>0.637</td>
<td>0.697</td>
<td>0.713</td>
<td>0.731</td>
<td>CPT-I random errors</td>
</tr>
<tr>
<td>F5 -&gt; CPT-IP four-digit random errors</td>
<td>0.733</td>
<td>0.739</td>
<td>0.655</td>
<td>0.705</td>
<td></td>
</tr>
<tr>
<td>F6 -&gt; CPT-IP three-digit false alarms</td>
<td>0.677</td>
<td>0.655</td>
<td>0.697</td>
<td>0.778</td>
<td>CPT-IP false alarms</td>
</tr>
<tr>
<td>F6 -&gt; CPT-IP four-digit false alarms</td>
<td>0.635</td>
<td>0.702</td>
<td>0.616</td>
<td>0.590</td>
<td></td>
</tr>
<tr>
<td>F1&lt;=F2</td>
<td>0.217</td>
<td>0.413</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1&lt;=F3</td>
<td>-0.231</td>
<td>-0.168</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1&lt;=F4</td>
<td>0.129</td>
<td>0.361</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1&lt;=F5</td>
<td>-0.385</td>
<td>-0.387</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1&lt;=F6</td>
<td>-0.215</td>
<td>0.135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2&lt;=F3</td>
<td>-0.536</td>
<td>-0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2&lt;=F4</td>
<td>0.443</td>
<td>0.523</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2&lt;=F5</td>
<td>-0.245</td>
<td>-0.452</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2&lt;=F6</td>
<td>-0.344</td>
<td>-0.298</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3&lt;=F4</td>
<td>-0.207</td>
<td>-0.272</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3&lt;=F5</td>
<td>0.235</td>
<td>0.287</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3&lt;=F6</td>
<td>0.210</td>
<td>0.204</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4&lt;=F5</td>
<td>-0.104</td>
<td>-0.295</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4&lt;=F6</td>
<td>-0.040</td>
<td>-0.094</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5&lt;=F6</td>
<td>0.386</td>
<td>0.573</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g -&gt; F1</td>
<td></td>
<td></td>
<td>-0.395</td>
<td>0.455</td>
<td></td>
</tr>
<tr>
<td>g -&gt; F2</td>
<td></td>
<td></td>
<td>-0.830</td>
<td>0.896</td>
<td></td>
</tr>
<tr>
<td>g -&gt; F3</td>
<td></td>
<td></td>
<td>-0.388</td>
<td>0.572</td>
<td></td>
</tr>
<tr>
<td>g -&gt; F4</td>
<td></td>
<td></td>
<td>0.606</td>
<td>-0.534</td>
<td></td>
</tr>
<tr>
<td>g -&gt; F5</td>
<td></td>
<td></td>
<td>0.455</td>
<td>-0.589</td>
<td></td>
</tr>
<tr>
<td>g -&gt; F6</td>
<td></td>
<td></td>
<td>0.421</td>
<td>-0.345</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refined four-factor models</th>
<th>HC-4 CFM</th>
<th>HC-4 HFM</th>
<th>No convergence for SCZ models</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 -&gt; Benton Judgment of Line Orientation Test</td>
<td>0.549</td>
<td>0.554</td>
<td>Executive function*</td>
</tr>
<tr>
<td>S1 -&gt; WASI matrix reasoning</td>
<td>0.692</td>
<td>0.685</td>
<td></td>
</tr>
<tr>
<td>S2 -&gt; BACS semantic fluency H01</td>
<td>0.517</td>
<td>0.509</td>
<td>Fluency/memory</td>
</tr>
<tr>
<td>S2 -&gt; Verbal memory</td>
<td>0.654</td>
<td>0.664</td>
<td></td>
</tr>
<tr>
<td>S3 -&gt; CPT-IP H01</td>
<td>0.651</td>
<td>0.604</td>
<td>Vigilance/attention</td>
</tr>
<tr>
<td>S3 -&gt; CPT-IP H03</td>
<td>-0.302</td>
<td>-0.326</td>
<td></td>
</tr>
<tr>
<td>S4 -&gt; BACS token motor task</td>
<td>0.291</td>
<td>0.283</td>
<td>Speed</td>
</tr>
<tr>
<td>S4 -&gt; BACS symbol coding task</td>
<td>0.680</td>
<td>0.699</td>
<td></td>
</tr>
<tr>
<td>S1&lt;=S2</td>
<td>0.712</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1&lt;=S3</td>
<td>0.279</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1&lt;=S4</td>
<td>0.520</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2&lt;=S3</td>
<td>0.282</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2&lt;=S4</td>
<td>0.633</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3&lt;=S4</td>
<td>0.622</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g -&gt; S1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g -&gt; S2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g -&gt; S3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g -&gt; S4</td>
<td></td>
<td>0.758</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 (cont.)

<table>
<thead>
<tr>
<th>Refined three-factor models</th>
<th>SCZ-3 CFM</th>
<th>SCZ-3 HFM</th>
<th>HC+SCZ-3 HFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 -&gt; Benton Judgment of Line Orientation Test</td>
<td>0.639</td>
<td>0.549</td>
<td>0.608</td>
</tr>
<tr>
<td>S1 -&gt; WASI matrix reasoning</td>
<td>0.847</td>
<td>0.692</td>
<td>0.845</td>
</tr>
<tr>
<td>S2 -&gt; BACS semantic fluency H01</td>
<td>0.642</td>
<td>0.516</td>
<td>0.728</td>
</tr>
<tr>
<td>S2 -&gt; Verbal memory</td>
<td>0.705</td>
<td>0.655</td>
<td>0.752</td>
</tr>
<tr>
<td>S4 -&gt; CPT-IP H01</td>
<td>0.508</td>
<td>0.370</td>
<td>0.611</td>
</tr>
<tr>
<td>S4 -&gt; BACS token motor task</td>
<td>0.474</td>
<td>0.284</td>
<td>0.626</td>
</tr>
<tr>
<td>S4 -&gt; BACS symbol coding task</td>
<td>0.866</td>
<td>0.735</td>
<td>0.877</td>
</tr>
<tr>
<td>S1&lt;-&gt;S2</td>
<td>0.716</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>S1&lt;-&gt;S4</td>
<td>0.717</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>S2&lt;-&gt;S4</td>
<td>0.780</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>g -&gt; S1</td>
<td>–</td>
<td>0.758</td>
<td>0.812</td>
</tr>
<tr>
<td>g -&gt; S2</td>
<td>–</td>
<td>0.938</td>
<td>0.967</td>
</tr>
<tr>
<td>g -&gt; S4</td>
<td>–</td>
<td>0.625</td>
<td>0.883</td>
</tr>
</tbody>
</table>

CFA, Confirmatory factor analysis; HC, healthy controls; CFM, correlated factor model; SCZ, schizophrenia cases; HFM, hierarchical factor model; CPT-IP, Continuous Performance Tests-Identical Pairs; BACS, Brief Assessment of Cognition in Schizophrenia; WASI, Wechsler Abbreviated Scale of Intelligence; WCST-64, Wisconsin Card Sorting Test, 64-card version.

a Earlier factor analysis indicated Tower of London and WCST-64 loaded on the same factor. These were removed in favour of selecting higher loading tests. Nevertheless, this suggested that that was significant overlapping variances. Hence the label was retained as executive function.

initial EFA and thus were selected to be indicators of the latent variable executive function.

**Evaluation of model discrimination of case-control status**

Negative predictive values (NPV) and positive predictive values (PPV) were reported for each model tested. Three-factor models (three correlated factors: NPV: 89.1%; PPV: 80.6%; three hierarchical factors: NPV: 89.1%; PPV: 80.6%) appeared to discriminate cases and controls equally well. However six-factor models were poorer in classifying cases and controls equally well. However six-factor models were poorer in classifying cases and controls (six correlated factors: NPV: 88.4%; PPV: 73.6%; six hierarchical factors: NPV: 86.6%; PPV: 70.9%). The g estimate from the six-factor model appeared to be the poorest classifier (six hierarchical factor g: NPV: 83.3%; PPV: 64.5%; composite g: NPV: 87.2%; PPV: 72.6%; three hierarchical factor g: NPV: 88.7%; PPV: 76.5%). To further evaluate three-factor models, forward stepwise logistic regression was conducted. Executive function [odds ratio (OR) = 3.13, 95% confidence interval (CI) = 1.89–5.18] and speed/vigilance (OR = 0.002, 95% CI = 0.001–0.004) remained significant predictors of case-control status (online Supplementary Fig. S3).

**Agreement of composite g and factor-derived g**

To clarify the utility of g estimands, post-hoc Bland–Altman (Bland & Altman, 1986) evaluation of g estimands from the six- and three-factor hierarchical models and composite g were tested (online Supplementary Fig. S4). Bland–Altman plots revealed that though composite scores showed tighter associations with the three-factor g and greater agreement compared with the six-factor g, there appeared to be systematic differences between the composite score-and factor-derived g.

**Evaluation of convergent validity of factor-derived g**

Bivariate correlations revealed moderate education correlations with g in both cases \[r_{(three-factor \ g)} = 0.528; \]
\[r_{(six-factor \ g)} = 0.413; \]
\[r_{(composite \ g)} = 0.489\] and controls \[r_{(three-factor \ g)} = 0.328; \]
\[r_{(six-factor \ g)} = 0.202; \]
\[r_{(composite \ g)} = 0.284.\]

**Discussion**

The current study comprises one of the largest single-site schizophrenia samples cognitively profiled. The primary objectives were to elucidate the latent cognitive structure that underlies neuropsychological performance in schizophrenia, establish if cognitive structures were similar between schizophrenia and healthy samples, and to examine the discriminant properties of various latent cognitive factors. Secondary objectives were to examine properties of factor analytic-derived g and conventionally derived g composite scores from a neuropsychological battery—to establish a viable cognitive phenotype that is adequately robust for subsequent studies (Dickinson et al. 2013; Donohoe et al. 2013).
Sample description

Case–control differences across cognitive tests up to two standard deviations below the mean performances of healthy controls were similar to those in previously published work (Saykin et al. 1991; Harvey & Keefe, 1997; Heinrichs & Zakzanis, 1998; Aleman et al. 1999; Bokat & Goldberg, 2003; Harvey et al. 2003, 2004; Harvey & Keefe, 2009; Keshavan et al. 2008; Szöke et al. 2008; Mesholam-Gately et al. 2009).

Examining covariance across test performance

Six factors were obtained in preliminary EFA in cases and controls, respectively, which appear to correspond to test-specific factors: F1 (CPT-IP hits); F2 (executive function/spatial reasoning); F3 (WCST-64); F4 (semantic fluency); F5 (CPT-IP commission errors); and F6 (CPT-IP false alarms). Initial factors appeared to be broadly consistent with previously reported cognitive factors in the literature (online Supplementary Table S1). However, methodological variance may distort the initial cognitive architecture. Data-driven EFA methods extracted broadly similar cognitive domains as in previous reports (Genderson et al. 2007; Wang et al. 2010). However, a model driven by the method-specific variances might not be useful in detecting neurobiological abnormalities or intervention effects.

Addressing methodological variance

Further refinement of neuropsychological test measures resulted in derivation of four- and three-factor models in controls and schizophrenia subjects, respectively. The three-factor model corresponded to S1 (executive function), S2 (fluency/memory) and S4 (speed/vigilance). While speed and vigilance did not appear separable in schizophrenia subjects, they did in healthy individuals. Factors appeared to represent commonly accepted domains of cognition and associated test modalities (see online Supplementary Table S1; Fioravanti et al. 2005; Dickinson & Gold, 2008).

CFA model fitting for derived factors

Nine models were tested. Fit indices across models were reasonable. Models that were reduced and refined were comparable with larger, more complex models that included method variances. Results may suggest several cognitive substrata of neuropsychological architecture, such that the initial extracted architecture could represent how test batteries are organized and administered, or method variance. Further refinement of the factor structure via CFA uncovered a simpler but more parsimonious architecture that represents underlying cognitive processes responsible for test performance.

While fit indices were generally acceptable for the nine models that were tested, there appears to be a trend that first-order correlated factors fit better than hierarchical models estimating latent g. An interpretation of this phenomenon may be related to first-order correlated models being more flexible in accounting for variances in the data, while hierarchical models tend to constrain the estimation of a second-order latent factor g. However, further exploration of the data is necessary to investigate if factor structures are similar after adjusting for g, and if g plays a direct role in test performance that is not otherwise mediated by latent cognitive domains (Gignac, 2008).

Cognitive factor analytic studies seek to identify separable cognitive domains subserved by specific neural processes. Factors derived from traditional neuropsychological tests are posited to be sufficiently independent to permit assay of discrete neural systems.

<table>
<thead>
<tr>
<th>Model</th>
<th>Type</th>
<th>Sample</th>
<th>NFI</th>
<th>RFI</th>
<th>IFI</th>
<th>TLI</th>
<th>CFI</th>
<th>RMSEA (90% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Six-factor correlated</td>
<td>HC</td>
<td>0.960</td>
<td>0.932</td>
<td>0.977</td>
<td>0.960</td>
<td>0.976</td>
<td>0.036 (0.028–0.044)</td>
</tr>
<tr>
<td>Model 2</td>
<td>Six-factor correlated</td>
<td>SCZ</td>
<td>0.964</td>
<td>0.940</td>
<td>0.981</td>
<td>0.968</td>
<td>0.981</td>
<td>0.039 (0.030–0.049)</td>
</tr>
<tr>
<td>Model 3</td>
<td>Six-factor hierarchical</td>
<td>HC</td>
<td>0.934</td>
<td>0.903</td>
<td>0.953</td>
<td>0.930</td>
<td>0.953</td>
<td>0.048 (0.041–0.054)</td>
</tr>
<tr>
<td>Model 4</td>
<td>Six-factor hierarchical</td>
<td>SCZ</td>
<td>0.923</td>
<td>0.886</td>
<td>0.941</td>
<td>0.912</td>
<td>0.941</td>
<td>0.065 (0.057–0.073)</td>
</tr>
<tr>
<td>Model 5</td>
<td>Four-factor correlated</td>
<td>HC</td>
<td>0.961</td>
<td>0.900</td>
<td>0.978</td>
<td>0.942</td>
<td>0.978</td>
<td>0.035 (0.018–0.051)</td>
</tr>
<tr>
<td>Model 6</td>
<td>Three-factor correlated</td>
<td>SCZ</td>
<td>0.985</td>
<td>0.962</td>
<td>0.994</td>
<td>0.984</td>
<td>0.994</td>
<td>0.032 (0.000–0.055)</td>
</tr>
<tr>
<td>Model 7</td>
<td>Four-factor hierarchical</td>
<td>HC</td>
<td>0.915</td>
<td>0.808</td>
<td>0.933</td>
<td>0.846</td>
<td>0.932</td>
<td>0.057 (0.043–0.071)</td>
</tr>
<tr>
<td>Model 8</td>
<td>Three-factor hierarchical</td>
<td>SCZ</td>
<td>0.980</td>
<td>0.929</td>
<td>0.989</td>
<td>0.961</td>
<td>0.989</td>
<td>0.034 (0.006–0.059)</td>
</tr>
</tbody>
</table>

CFA, Confirmatory factor analysis; NFI, normed fit index; RFI, relative fit index; IFI, incremental fit index; TLI, Tucker–Lewis index; CFI, comparative fit index; RMSEA, root mean square of approximation; CI, confidence interval; HC, healthy controls; SCZ, schizophrenia cases.
Evaluation of model discrimination of case-control status

The reduced and refined three-factor model best discriminated cases and controls. Results support the notion that the factor structure more proximal to cognitive domains is probably more sensitive in separating cases and controls.

Follow-up logistic regression indicated that executive function and speed/vigilance demonstrated superior discriminant properties in our sample. Executive function and speed/vigilance demonstrated superior domains is probably more sensitive in separating cases and controls. Results support the notion that the factor structure more proximal to cognitive domains is high and may reduce orthogonality, the resultant cognitive architecture may represent the subtle facets of cognition that may nevertheless be valuable to future biological research in schizophrenia.

An evaluation of g estimation

Calculating g is a valuable approach due to the phenomenon of ‘positive manifold’ (Carroll, 1993; Jensen, 1998), its widespread practicality, associations with biological variables (Jensen, 1992, 1998, 2002) and its genetic contributions (e.g. Davies et al. 2011). The theory of indifference suggests, given the administration of a sufficiently large and diverse selection of neuropsychological instruments, that the composite cognitive scores of any given battery are likely to be similar to that of any other battery (Jensen, 1998; Johnson et al. 2004, 2008; Hunt, 2011; Mackintosh, 2011; Deary, 2012). This encourages the use of g as a candidate phenotype in inter-center replication studies (Loo et al. 2012).

Here, we found differences between conventional composite score-type calculation of g versus factor analytic-derived g. The factor analytic estimation may not have been optimal in this study because of the inherent limitations of using a second-order factor where estimation becomes largely dependent on the first-order loadings (for an in-depth review, see Gignac, 2008). It is also possible that measures originally validated in Caucasian samples (Keefe et al. 2004) may not capture latent factors specific of an Asian sample, hence there may be subtle differences in the estimation of true factor scores. The use of a second-order hierarchical factor could therefore have amplified these differences. Perhaps a much larger and diverse battery.
would have better stabilized the hierarchical model in our sample. For this reason, either a conservative approach of calculating composite scores or alternative factor analytic approaches for \(g\) estimation in Asian samples may be required (Gignac, 2006, 2008).

Two aspects of \(g\) estimation should be considered in subsequent studies. First, although discriminant properties of both methods of calculating general cognition appear comparable, differences at the level of score distribution are probably related to score calculation methods. Factor analytic \(g\) appeared slightly more sensitive to associations with education. However, further evaluation of sensitivity of either index is necessary (e.g. in genetic association studies where covariances with DNA polymorphism of either measure can be thoroughly evaluated). Second, in our review, the number of impaired cognitive domains ranged from six to 22 factors, which may be limited by heterogeneous sample characteristics and methodologies (Wilk et al. 2004; Foriavantti et al. 2005), statistical inadequacies in addressing latent substructure (Genderson et al. 2007) and methodological variances in test administration (Nuechterlein et al. 2004; Kraus & Keefe, 2007). The potential presence of several cognitive architectures within a battery of tests may suggest possibilities where the diverse factor structures previously reported in the literature can be reconciled in subsequent studies. In this context, the calculation of composite \(g\) may still be required in cross-center collaborations.

**Limitations**

The current study benefits from its large sample sizes of controls and patients. Our study is not based on an exhaustive battery of tests covering the entirety of cognitive aspects classically assessed in patients, as decisions for test inclusion had to balance comprehensiveness of the assessment with its practicality and tolerability. Also, although medication type was recorded, its effects were not explicitly tested as part of the factor model, as the complex task of reviewing lifetime case records is disproportionate with respect to the scope of the study.

**Conclusions**

As the field moves towards a more dimensional approach of understanding complex psychiatric illness, measurement of cognition and refinement of measures will continue to be important in clinical practice and research (Collinson et al. 2010). Separable cognitive factors were identified in the current study that may be valuable in capturing subtle aspects of cognitive processes. Cognitive and neuropsychological researchers familiar with the inherent strengths and weakness of neuropsychological testing are in the position to further develop innovative strategies, refine neuropsychological procedures, and maximize what tests can reveal about the complex nature of cognitive deficits and their underlying neural substrates in schizophrenia.

**Supplementary material**

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291714001020

**Acknowledgements**

The Singapore Translational and Clinical Research in Psychosis is supported by the National Research Foundation Singapore under the National Medical Research Council Translational and Clinical Research Flagship Programme (grant no. NMRC/TCR/003/2008). M.L. is supported by the National Medical Research Council Training Fellowship (grant no. MH095: 003/008-1014) during the preparation of the current paper.

S.-A.C. has full access to all data generated by the consortium and is responsible for the integrity and accuracy of data reported. The authors also acknowledge Dr Mythily Subramaniam and Dwight Dickinson for their advice and inputs to the manuscript and analysis of the current study.

**Declaration of Interest**

R.S.E.K. currently or in the past 24 months has received investigator-initiated research funding support from the Department of Veterans Affairs, Feinstein Institute for Medical Research, GlaxoSmithKline, National Institute of Mental Health, Novartis, Psychogenics, Research Foundation for Mental Hygiene, Inc. and the Singapore National Medical Research Council. R.S.E.K. currently or during the past 24 months has received honoraria, or served as a consultant or advisory board member for Abbvie, Akebia, Amgen, Astellas, Asubio, BiolineRx, Biomarin, Boehringer-Ingelheim, Bristol-Myers Squibb, Eli Lilly, EnVivo, Helicon, Lundbeck, Merck, Mitsubishi, Otsuka, Pfizer, Roche, Shire, Sunovion, Takeda and Targacept. R.S.E.K. receives royalties from the BACS testing battery and the MATRICS Battery (BACS Symbol Coding). He is also a shareholder in NeuroCog Trials, Inc., Durham NC. The other authors have no conflicts of interest to declare.

**References**


Gignac GE (2006). Evaluating subtest ‘g’ saturation levels via the single trait-correlated uniqueness (STCU) SEM approach: evidence in favor of crystallized subtests as the best indicators of ‘g’. Intelligence 34, 29–46.


Harvey PD, Keefe RSE (1997). Cognitive impairment in schizophrenia and implications of atypical neuroleptic treatment. CNS Spectrums 2, 201–222.


Reilly JL, Harris MSH, Khine TT, Keshavan MS, Sweeney JA (2008). Reduced attentional engagement contributes to deficits in prefrontal inhibitory control in schizophrenia. Biological Psychiatry 63, 776–783.


