Latent Structure of the Wisconsin Card Sorting Test in a Clinical Sample

Kevin W. Greve
University of New Orleans

Fred Ingram
The University of Texas Medical Branch at Galveston

Kevin J. Bianchini
University of New Orleans and East Jefferson General Hospital

The aims of this study were to: (a) examine the consistency of the published Wisconsin Card Sorting Test (WCST) factor structures; (b) determine the factor structure of the WCST in a large, heterogeneous sample; and (c) compare the WCST factor analytically with other neuropsychological procedures. Two WCST factors (concept formation/perseveration and Failure-to-Maintain-Set [FMS]) were consistently reported in the literature. Our analysis of data from 473 clinical cases replicated the two factors previously reported and revealed a third on which nonperseverative errors (NPE) was the sole salient variable. This pattern was maintained in three of four diagnostically distinct subgroups. These factors are potentially clinically meaningful, with each seeming to reflect one of three qualitatively different performance styles. In the construct validation factor analysis, WCST scores loaded independently of other neuropsychological variables, indicating that the WCST contributes uniquely to neuropsychological evaluation. Nevertheless, despite the rational interpretation of the factors, the cognitive processes underlying WCST performance remain poorly understood. Future directions for the application of these factor analytic findings are discussed. © 1998 National Academy of Neuropsychology. Published by Elsevier Science Ltd
Factor I (which accounts for 48–71% of the variance) has been interpreted variously as reflecting executive and memory function (either purely or interacting, depending on the sample) (Sullivan et al., 1993) or problem-solving ability (Greve, Williams, Haas, Littell, & Reinoso, 1996). The scores that load most highly on Factor I reflect two related aspects of executive function: (a) the ability to recognize the possible sorting concepts (percent conceptual level responses, categories completed, total correct responses); and (b) the inability to shift from an incorrect response set (perseverative errors, perseverative responses, total errors).

Factor II is comprised of scores that seem to measure the ability to: (a) quickly and efficiently test hypotheses and discover the correct dimension; and (b) maintain correct responding. Early studies suggested that scores related to Factor II reflect memory (Bowen, Kamienny, Burns, & Yahr, 1975), motivation (Fey, 1951), and attention (Fey, 1951). More recent studies have not provided much in the way of additional illumination. Following Fey (1951), Greve and colleagues (Greve et al., 1996, 1997) have further argued for the importance of attention in this factor on the basis of the effects of experimental manipulations of attentional processes. In contrast, others (Greve, Farrell, Besson, & Crouch, 1995; Paolo et al., 1995; Sullivan et al., 1993) have found these scores uncorrelated with clinical measures of either memory or attention; indeed, they were uncorrelated with any tasks these authors administered. Thus, the cognitive processes underlying this factor remain unclear.

The value of factor analytic study of the WCST derives from the information provided about the cognitive processes underlying performance on this task. However, the latent structure of a test as revealed by factor analysis is not necessarily equal to the cognitive processes required to perform the task. In other words, observation of a single factor does not imply the involvement of only one cognitive process or system. However, observing more than one factor most certainly implicates the involvement of more than one cognitive process. Examining the factor structure in a relatively homogenous sample of patients may give us insight into the nature of the spared and impaired cognitive abilities in such patients if the meaning (construct validity) of the observed factors is well established. This approach is of limited value, however, if we do not know whether the factors observed represent single cognitive processes or aggregates of several processes that are highly correlated in that particular sample. A better approach is to elucidate the factor structure of a test in the most heterogeneous samples possible, thereby increasing the likelihood that the individual factors represent relatively homogeneous cognitive processes (at least within the limits of the specific test or tests). The stability of the factors can then be examined in more homogenous samples.

If factor analysis of a particular set of variables produces a true picture of the cognitive processes underlying performance on those measures then one would expect some consistency in the variables comprising a given factor across studies and subgroups (Cattell & Baggageley, 1960). Traditionally, similarity of factor composition has been assessed rather informally via visual analysis. However, Cattell and Baggageley (1960) have argued that such an approach was inadequate. A number of statistical methods for measuring of the degree of similarity between two matched factors have been proposed (see Gorsuch, 1983, for formulae and discussion of several).

Reynolds and Harding (1983) compared six of these methods and found them to lead to similar conclusions. These authors recommend an approach in which the coefficient of congruence (c) is supplemented by the salient variable similarity index (s). “Only when both of these values are large, one may be relatively certain that the factor in question is indeed congruent across groups” (p. 728). The coefficient of congruence for an orthogonal factor is essentially the correlation coefficient for the pairs of factor scores (Gorsuch, 1983). The salient variable similarity index is a nonparametric measure of the proportion of variables common to two analyses that are salient (i.e., have a statistically significant loading) on a
given factor. Both measures are designed to compare loading patterns. Thus, the similarity between factors can be easily quantified and objectively evaluated. Given the apparent consistency of the number and composition of WCST factors reported in the literature, the question naturally arises as to whether that apparent consistency would be supported statistically.

Part 1 of the present study will use \( s \) and \( c \) to determine if the composition of the reported WCST factors is as consistent and reliable as it seems on informal analysis. Part 2 will further explore the WCST factor structure. Since the most information is to be found in heterogeneous populations, the factor structure will first be examined in a very large and diverse patient sample. Follow-up analyses will examine the latent WCST structure in several diagnostically based subgroups in order to assess the stability of the factor structure derived from the full, heterogeneous sample. Finally, in part 3, the WCST scores will be compared, factor analytically, to performance on several other neuropsychological procedures to help better define the WCST’s construct validity.

**PART 1: RELIABILITY OF THE WCST FACTOR STRUCTURE**

**Method**

The data for this analysis came from 10 factor analyses of the WCST reported in four papers (Goldman et al., 1996; Greve et al., 1997; Paolo et al., 1995; Sullivan et al., 1993) published in peer-reviewed journals. Each of these analyses (except Goldman et al.’s normal controls) resulted in at least two factors. Cattell’s salient variable similarity index (\( s \); Cattell & Baggaley, 1960; Tabachnick & Fidell, 1989) and the Coefficient of Congruence (\( c \); Gorsuch, 1983) were calculated independently for all pairs of samples for Factors I and II. The raw factor loading scores were entered into the formula presented in Gorsuch (1983) for the calculation of \( c \). For the calculation of \( s \) (please refer to Tabachnick & Fidell, 1989, for details of this analysis), the salient variables were identified according to Stevens (1992) converted to either 1 or \(-1\), depending on the sign of the loading. Nonsalient variables were converted to 0. Variables not common to both analyses in a pair were dropped from the calculation of \( s \).

**RESULTS**

Table 1 presents \( s \) and \( c \) for all pairs of comparisons on Factor I. Our analysis of the composition of Factor I demonstrated relatively high consistency between samples within those studies with multiple samples. The consistency of the Factor I pattern for all between-study comparisons, excluding those with a Goldman et al. (1996) sample, was also quite high (\( s > .91, c > .92 \)). The between-study comparisons involving the Goldman samples produced values of \( s \) and \( c \) of less than .50. Thus, as can be seen in Table 1, the factor patterns reported by Goldman et al. (1996) were inconsistent with those reported in the other three studies. This unreliable factor structure was likely due to the use of only five variables and the finding of only one factor for the normal controls. The variables that loaded consistently (i.e., were salient in all studies) on Factor I were: Percent Conceptual Level Responses, Categories Completed (not used in Goldman), Perseverative Errors, Perseverative Responses, and Total Errors. Total Correct was salient on Factor I in 50% of the analyses and so makes a significant but less reliable contribution.

The composition of Factor II was much more consistent across samples and studies (see Table 2). Only comparisons involving Paolo et al.’s (1995) Parkinson disease patients showed an unreliable pattern. This was because Nonperseverative Errors loaded saliently (but oppo-
TABLE 1
Salient Variable Similarity Index (s) and Coefficient of Congruence (c) Comparing the Factor I Pattern for All Pairs of Analyses

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mixed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sullivan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Elderly</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paolo</td>
<td>.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Parkinson</td>
<td>.91</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paolo</td>
<td>.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Normal</td>
<td>.29</td>
<td>.50</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.26</td>
<td>.34</td>
<td>.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Normal</td>
<td>.25</td>
<td>.44</td>
<td>.44</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.24</td>
<td>.36</td>
<td>.36</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Focal</td>
<td>.29</td>
<td>.50</td>
<td>.25</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.24</td>
<td>.13</td>
<td>.27</td>
<td>.84</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Diffuse</td>
<td>.29</td>
<td>.50</td>
<td>.25</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.20</td>
<td>.14</td>
<td>.25</td>
<td>.90</td>
<td>.86</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Combined</td>
<td>.91</td>
<td>1.00</td>
<td>.91</td>
<td>.33</td>
<td>.29</td>
<td>.33</td>
<td>.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greve</td>
<td>.92</td>
<td>.98</td>
<td>.97</td>
<td>.05</td>
<td>.04</td>
<td>.05</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Students</td>
<td>1.00</td>
<td>1.00</td>
<td>.91</td>
<td>.33</td>
<td>.29</td>
<td>.33</td>
<td>.33</td>
<td>.91</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Patients</td>
<td>1.00</td>
<td>1.00</td>
<td>.91</td>
<td>.33</td>
<td>.29</td>
<td>.33</td>
<td>.33</td>
<td>.91</td>
<td>1.00</td>
</tr>
<tr>
<td>Greve</td>
<td>.97</td>
<td>.98</td>
<td>.97</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
<td>.01</td>
<td>.98</td>
<td>.99</td>
<td></td>
</tr>
</tbody>
</table>

Note. There was only one factor found for the two Goldman Normal samples.

Failure-to-Maintain-Set was salient on Factor II in all studies. Total Correct also loaded on Factor II though, as on Factor I, somewhat less reliably. Similarly, Nonperseverative Errors loaded inconsistently on Factor II. The pattern of Factor III was not examined because a third factor was reported in only two studies and had no variables in common. In summary, visual analysis of the various reported factor structures of the WCST suggests that a consistent two factor pattern best accounts for performance on the WCST.

TABLE 2
Salient Variable Similarity Index (s) and Coefficient of Congruence (c) Comparing the Factor II Pattern for All Pairs of Analyses

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mixed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sullivan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Elderly</td>
<td>.67</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paolo</td>
<td>.50</td>
<td>.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Parkinson</td>
<td>.67</td>
<td>.67</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paolo</td>
<td>.89</td>
<td>.79</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Focal</td>
<td>.67</td>
<td>.67</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.85</td>
<td>.75</td>
<td>-.12</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Diffuse</td>
<td>.67</td>
<td>.67</td>
<td>.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman</td>
<td>.28</td>
<td>.90</td>
<td>.79</td>
<td>.92</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Combined</td>
<td>.67</td>
<td>1.00</td>
<td>1.00</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greve</td>
<td>.68</td>
<td>.97</td>
<td>.88</td>
<td>.96</td>
<td>.94</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Students</td>
<td>.67</td>
<td>1.00</td>
<td>.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greve</td>
<td>.51</td>
<td>.95</td>
<td>.85</td>
<td>.97</td>
<td>.95</td>
<td>.91</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. There was only one factor found for the two Goldman Normal samples.
This conclusion was supported statistically. Table 3 shows the composite factor structure for the WCST based on the present analysis of the published studies. The values in Table 3 will serve as the means of comparing the factor structure derived in the next section with the past literature.

**PART 2: WCST LATENT STRUCTURE**

**Method**

Data for this study were compiled retrospectively from the clinical files of 473 patients seen for neuropsychological evaluations at three sites (post acute brain trauma rehabilitation center; comprehensive hospital-based physical rehabilitation center; general neuropsychological outpatient practice) over the past 9 years. Subjects were included if they completed the WCST and Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981). See Table 4 for subject characteristics. These subjects also completed various other neuropsychological procedures though not all received exactly the same battery or completed all tasks; the specific procedures will be described in Part 3.

The WCST was administered and scored in accordance with the instructions in the revised manual (Heaton et al., 1993). Only scores described in the manual were included in the analysis. These were: Total Correct (TC); Perseverative Responses (PR); Perseverative Errors (PE); Nonperseverative Errors (NPE); Percent Conceptual Level Responses (%CLR); Categories Completed (CAT); and, Failure-to-Maintain-Set (FMS). Only raw scores were used. Several scores were excluded. Total Trials and Total Errors (TE) were not included because these scores are a linear combination of two or more included scores (TC, PE, NPE, and PE, NPE, respectively) and were thus redundant. Trials-to-complete-the-first-category and Learning-to-learn were not included because they required a degree of success many subjects did not achieve (the completion of one and three categories, respectively); those subjects would have been deleted automatically from the multivariate analyses, thus dramatically reducing and likely biasing the sample. Thus, all variables included in Table 2, except TE, were included in this analysis.
TABLE 4

Characteristics in the Full Sample and Construct Validation Subset

<table>
<thead>
<tr>
<th></th>
<th>Validation</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subset</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Percent female</td>
<td>28.3</td>
<td>30.2</td>
</tr>
<tr>
<td>Age (years)</td>
<td>12.6</td>
<td>19.2</td>
</tr>
<tr>
<td>Education (years)</td>
<td>85.7</td>
<td>12.2</td>
</tr>
<tr>
<td>WAIS-R VIQ</td>
<td>81.9</td>
<td>12.3</td>
</tr>
<tr>
<td>WAIS-R PIQ</td>
<td>83.2</td>
<td>12.8</td>
</tr>
<tr>
<td>WAIS-R FSIQ</td>
<td>64.7</td>
<td>17.3</td>
</tr>
<tr>
<td>Preservative responses</td>
<td>39.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Perseverative errors</td>
<td>32.8</td>
<td>22.1</td>
</tr>
<tr>
<td>Nonperseverative errors</td>
<td>21.0</td>
<td>22.1</td>
</tr>
<tr>
<td>Conceptual level responses</td>
<td>42.8</td>
<td>25.2</td>
</tr>
<tr>
<td>Categories completed</td>
<td>3.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Failure-to-Maintain-Set</td>
<td>1.1</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Diagnostic Composition of the Full Sample and Construct Validation Subset

<table>
<thead>
<tr>
<th></th>
<th>Validation Subset</th>
<th>Full Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Total sample</td>
<td>230</td>
<td>100.0</td>
</tr>
<tr>
<td>Brain trauma</td>
<td>81</td>
<td>35.2</td>
</tr>
<tr>
<td>Stroke</td>
<td>60</td>
<td>26.1</td>
</tr>
<tr>
<td>Dementia</td>
<td>14</td>
<td>6.1</td>
</tr>
<tr>
<td>Psychiatric</td>
<td>10</td>
<td>4.3</td>
</tr>
<tr>
<td>Mixed</td>
<td>45</td>
<td>19.6</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>2</td>
<td>.9</td>
</tr>
<tr>
<td>Seizure disorder</td>
<td>2</td>
<td>.9</td>
</tr>
<tr>
<td>Anoxia/hypoxia</td>
<td>5</td>
<td>2.1</td>
</tr>
<tr>
<td>Neoplasm</td>
<td>4</td>
<td>1.7</td>
</tr>
<tr>
<td>Infection</td>
<td>2</td>
<td>.9</td>
</tr>
<tr>
<td>Academic/developmental delay</td>
<td>2</td>
<td>.9</td>
</tr>
<tr>
<td>Other medical illness</td>
<td>4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Note. M = mean; SD = standard deviation; WAIS-R VIQ = Wechsler Adult Intelligence Scale-Revised Verbal Intelligence Quotient; WAIS-R PIQ = Wechsler Adult Intelligence Scale-Revised Performance Intelligence Quotient; WAIS-R FSIQ = Wechsler Adult Intelligence Scale-Revised Full Scale Intelligence Quotient; ns = nonsignificant.

RESULTS

The seven WCST scores (see Table 4 for descriptive statistics) were submitted to a principal components analysis with an orthogonal (varimax) rotation. A criterion of eigenvalue greater than 1.0 resulted in a three-factor solution that accounted for 94% of the variance (see Table 5 for the rotated factor structure). To examine the stability of the factor structure, the sample was randomly divided and the analysis repeated for each half. This split-half analysis resulted in factor structures that were identical to that of the entire sample (s = 1.00 and c > .97 for all factors in both samples). Thus the WCST factor structure observed in this sample was internally consistent.
To evaluate external consistency, each factor observed in this study was then compared to the corresponding factor derived from the literature (Table 3). The absence of a reliable third factor is addressed below. The variables loading on Factor I in this study were identical to those in the literature, but were opposite in sign. When the problem of sign was taken into account, $s$ was 1.00.

In the past, Factor II has been represented by FMS. In the present sample, FMS loaded on Factor III. Thus, the comparison of the composite FMS factor with our FMS factor produced an $s$ of .80. This study produced a separate factor on which NPE loaded significantly. An NPE factor has only been reported once (Sullivan et al., 1993); our NPE factor was moderately consistent with theirs ($s = .67, c = -.62$).

In summary, the factor pattern derived from the present data set is consistent with those reported in the literature. The salient variables on our first factor were identical to those in the literature. The difference in sign simply reflects the relatively greater influence of the variables, indicating positive performance (TC, %CLR, CAT) in our analysis because there were fewer variables indicating negative performance (PR, PE). In the previous analyses the negative indices had more influence because TE was included. To test this hypothesis, our data were reanalyzed with TE included. The resulting factor structure was identical but the signs of the variables on Factor I were now consistent with the literature. The FMS factor was also found in this analysis, though it loaded on a third factor rather than the traditional second. Thus, this study demonstrated the presence of the two factors found in all previous studies reporting more than one factor. The third factor reported (NPE) is consistent with the third factor reported by Sullivan et al. (1993).

To further study the stability of the WCST factor structure, separate factor analyses were computed for each of our four major diagnostic groups: brain trauma ($n = 161$), stroke ($n = 117$), dementia ($n = 35$), and psychiatric ($n = 33$). The three-factor solution accounted for 95.3, 95.0, 96.6, and 91.3% of the variance in each group, respectively. See Table 6 for the full factor structures. In three of the four subgroups the eigenvalue of the third factor was less than 1.0; the third factor was retained on the basis of the scree test criterion recommended by Stevens’ (1992) for samples under 200. In comparison to the full sample, Factor I continued to be stable and reliable except in the psychiatric sample. The apparent unreliability of the factor composition in the psychiatric group was due to the reversed sign on the salient variables, which lead to artificially low values of $s$; note that $c$ was $-.99$. The composi-
<table>
<thead>
<tr>
<th></th>
<th>Brain Trauma</th>
<th>Stroke</th>
<th>Dementia</th>
<th>Psychiatric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
<td>III</td>
<td>I</td>
</tr>
<tr>
<td>Percent variance</td>
<td>62.40</td>
<td>20.00</td>
<td>13.00</td>
<td>60.90</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>4.36</td>
<td>1.40</td>
<td>.91</td>
<td>4.26</td>
</tr>
<tr>
<td>$s^a$</td>
<td>.75</td>
<td>.86</td>
<td>.67</td>
<td>1.00</td>
</tr>
<tr>
<td>$c^a$</td>
<td>.49</td>
<td>.98</td>
<td>.98</td>
<td>.99</td>
</tr>
<tr>
<td>PR</td>
<td>-.98*</td>
<td>.00</td>
<td>-.16</td>
<td>-.89*</td>
</tr>
<tr>
<td>PE</td>
<td>-.98*</td>
<td>-.06</td>
<td>-.16</td>
<td>-.89*</td>
</tr>
<tr>
<td>%CLR</td>
<td>.84*</td>
<td>.49*</td>
<td>-.07</td>
<td>.93*</td>
</tr>
<tr>
<td>CAT</td>
<td>.83*</td>
<td>.49*</td>
<td>-.07</td>
<td>.91*</td>
</tr>
<tr>
<td>TC</td>
<td>.58*</td>
<td>.22</td>
<td>.69*</td>
<td>.83*</td>
</tr>
<tr>
<td>NPE</td>
<td>-.13</td>
<td>-.98*</td>
<td>.01</td>
<td>-.06</td>
</tr>
<tr>
<td>FMS</td>
<td>.01</td>
<td>-.09</td>
<td>.97*</td>
<td>.15</td>
</tr>
</tbody>
</table>

Note: CAT = categories completed; FMS = failure-to-maintain-set; NPE = nonperseverative errors; PE = perseverative errors; PR = perseverative responses; TC = total correct; %CLR = percent conceptual level responses.

*a*For subsample compared to the full sample.

*Statistically significant loading for this variable according to the Stevens (1992) criteria.
TABLE 7
Comparison of the Two Samples on the Neuropsychological Variables Included in the Construct Validation Factor Analysis

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th>Validation Subset</th>
<th>F(df*)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Verbal Comprehension Factor</td>
<td>82.5</td>
<td>14.8</td>
<td>89.2</td>
<td>13.7</td>
</tr>
<tr>
<td>Perceptual Organization Factor</td>
<td>73.2</td>
<td>16.6</td>
<td>79.9</td>
<td>15.9</td>
</tr>
<tr>
<td>Freedom from Distractibility</td>
<td>81.5</td>
<td>12.2</td>
<td>85.3</td>
<td>11.9</td>
</tr>
<tr>
<td>Logical Memory I (raw)</td>
<td>18.2</td>
<td>8.5</td>
<td>19.8</td>
<td>8.0</td>
</tr>
<tr>
<td>Logical Memory II (raw)</td>
<td>12.4</td>
<td>9.1</td>
<td>13.7</td>
<td>9.0</td>
</tr>
<tr>
<td>Visual Reproduction I (raw)</td>
<td>24.7</td>
<td>9.9</td>
<td>26.8</td>
<td>8.6</td>
</tr>
<tr>
<td>Visual Reproduction II (raw)</td>
<td>14.6</td>
<td>11.8</td>
<td>16.6</td>
<td>11.5</td>
</tr>
<tr>
<td>Trail-Making Test-A (seconds)</td>
<td>84.5</td>
<td>66.5</td>
<td>68.9</td>
<td>42.9</td>
</tr>
<tr>
<td>Trail-Making Test-B (seconds)</td>
<td>168.3</td>
<td>89.6</td>
<td>168.0</td>
<td>87.4</td>
</tr>
<tr>
<td>Digit-Span Backward (raw)</td>
<td>4.6</td>
<td>2.3</td>
<td>5.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>

Note. M = mean; SD = standard deviation; df = degrees of freedom; ns = nonsignificant.
'df varies because not all subjects in the full sample complete all tasks.

ation of Factor II (NPE) was variable in all subgroups. Nonetheless, NPE was consistently the most salient variable on this factor. Finally, Factor III, represented primarily by FMS, was consistent in all subgroups. In summary, the WCST factor structure reported in the literature was replicated in our full sample. This structure proved generally stable in three of four diagnostically based subgroups.

PART 3: CONSTRUCT VALIDATION

Method

The data used for the construct validation came from a subset of the patients described above. See Table 4 for descriptive statistics and comparison with the full sample for patient characteristics, diagnosis, and WCST scores. In addition to the seven WCST scores, the construct validation data set contained the following 10 variables representing a broad range of cognitive functions:

1. WAIS-R Verbal Comprehension factor
2. WAIS-R Perceptual Organization factor
3. WAIS-R Freedom from Distractibility factor (formulas for these factor scores are from Sattler, 1992)
4. Digit span backward raw score from the WAIS-R was used independently as a direct measure of working memory (see Reynolds, 1997, for a discussion of digits backward)
5. Raw scores (time in seconds) for Trails A and B (Reitan, 1958; War Department, Adjutant General’s Office, 1944)
6. Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987) raw scores for Logical Memory I and II (LM I and II) and Visual Reproduction I and II (VR I and II). Subjects with missing variables are automatically excluded from multivariate analyses, resulting in a validation sample of 230 subjects.

See Table 7 for descriptive statistics for these variables and comparisons with the full sample.
TABLE 8
Results of the Construct Validation Factor Analysis (Varimax Rotation) Using the Wisconsin Card Sorting Test Raw Scores

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent variance</td>
<td>42.1</td>
<td>14.4</td>
<td>8.8</td>
<td>7.2</td>
<td>6.9</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>7.15</td>
<td>2.45</td>
<td>1.49</td>
<td>1.23</td>
<td>1.17</td>
</tr>
<tr>
<td>Total correct</td>
<td>.89*</td>
<td>.06</td>
<td>.02</td>
<td>.12</td>
<td>.08</td>
</tr>
<tr>
<td>Perseverative responses</td>
<td>-.84*</td>
<td>-.36</td>
<td>-.20</td>
<td>.00</td>
<td>-.04</td>
</tr>
<tr>
<td>Perseverative errors</td>
<td>-.83*</td>
<td>-.38</td>
<td>-.19</td>
<td>.02</td>
<td>-.04</td>
</tr>
<tr>
<td>Percent Conceptual level responses</td>
<td>.74*</td>
<td>.34</td>
<td>.18</td>
<td>.09</td>
<td>.45</td>
</tr>
<tr>
<td>Categories completed</td>
<td>.61*</td>
<td>.44</td>
<td>.22</td>
<td>-.02</td>
<td>.53</td>
</tr>
<tr>
<td>Failure-to-Maintain-Set</td>
<td>.61*</td>
<td>-.22</td>
<td>-.21</td>
<td>.23</td>
<td>-.37</td>
</tr>
<tr>
<td>Trail-Making Test A</td>
<td>-.05</td>
<td>-.80*</td>
<td>-.15</td>
<td>-.03</td>
<td>-.06</td>
</tr>
<tr>
<td>Trail-Making Test B</td>
<td>-.08</td>
<td>-.78*</td>
<td>-.30</td>
<td>-.09</td>
<td>.02</td>
</tr>
<tr>
<td>Visual Reproduction I</td>
<td>.28</td>
<td>.75*</td>
<td>.00</td>
<td>.26</td>
<td>.14</td>
</tr>
<tr>
<td>Perceptual Organization</td>
<td>.27</td>
<td>.73*</td>
<td>.22</td>
<td>.20</td>
<td>.06</td>
</tr>
<tr>
<td>Visual Reproduction II</td>
<td>.29</td>
<td>.68*</td>
<td>-.02</td>
<td>.40</td>
<td>.10</td>
</tr>
<tr>
<td>Freedom from Distractibility</td>
<td>.13</td>
<td>.15</td>
<td>.87*</td>
<td>.22</td>
<td>.08</td>
</tr>
<tr>
<td>Digit-Span Backward</td>
<td>.11</td>
<td>.21</td>
<td>.83*</td>
<td>.11</td>
<td>.06</td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>.05</td>
<td>.14</td>
<td>.62*</td>
<td>.54</td>
<td>.03</td>
</tr>
<tr>
<td>Logical Memory I</td>
<td>.06</td>
<td>.26</td>
<td>.19</td>
<td>.87*</td>
<td>.11</td>
</tr>
<tr>
<td>Logical Memory II</td>
<td>.07</td>
<td>.19</td>
<td>.25</td>
<td>.85*</td>
<td>.11</td>
</tr>
<tr>
<td>Nonperseverative errors</td>
<td>-.09</td>
<td>-.02</td>
<td>-.04</td>
<td>-.21</td>
<td>-.89*</td>
</tr>
</tbody>
</table>

*Statistically significant loading for this variable according to the Stevens (1992) criteria.

RESULTS

This subsample did not differ from the full sample in terms of age or education. They were significantly less impaired than the full sample (see Tables 4 and 7) on WAIS-R variables (IQ’s and factor scores), WMS-R VR I, and Trails A. There were no group differences on any WCST variables. Prior to submitting the entire set of variables to factor analysis for construct validation, the factor structure of the WCST in the present subgroup was examined to insure that the original factor solution was preserved. This analysis produced a three-factor solution accounting for 95.1% of the variance. The loadings of the variables were almost identical to those in the original full sample analysis except that the FMS and NPE factors changed places, loading on Factor II and Factor III, respectively. The factor structure and composition observed in this subgroup were statistically highly consistent with the full sample.

The 17 variables used in the construct validation were then submitted to a principal components analysis with an orthogonal (varimax) rotation. A criterion of eigenvalue greater than 1 produced a five-factor solution accounting for 79.4% of the variance. Table 8 contains the loadings for the five-factor model. As can be seen, the WCST variables were represented on two factors (I and IV) with FMS loading rather weakly on Factor I. When a six-factor solution was forced (the scree plot indicated this was an acceptable solution), 85% of the variance was accounted for and FMS loaded on its own independent factor. None of the neuropsychological variables loaded significantly on any of the WCST factors. Thus, minimal information about the construct validity of the WCST was obtained from this analysis.
TABLE 9
Results of the Construct Validation Factor Analysis (Varimax Rotation) Using the Wisconsin Card Sorting Test (WCST) Factor Scores

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Variance</td>
<td>41.0</td>
<td>12.8</td>
<td>9.5</td>
<td>7.5</td>
<td>7.4</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>5.33</td>
<td>1.66</td>
<td>1.23</td>
<td>.98</td>
<td>.96</td>
</tr>
<tr>
<td>Visual Reproduction I</td>
<td>.82*</td>
<td>.29</td>
<td>.01</td>
<td>.06</td>
<td>.03</td>
</tr>
<tr>
<td>Perceptual Organization</td>
<td>.79*</td>
<td>.17</td>
<td>.25</td>
<td>.09</td>
<td>.04</td>
</tr>
<tr>
<td>Trail-Making Test A</td>
<td>−.76*</td>
<td>.03</td>
<td>−.19</td>
<td>−.24</td>
<td>.15</td>
</tr>
<tr>
<td>Visual Reproduction II</td>
<td>.76*</td>
<td>.42</td>
<td>−.03</td>
<td>−.02</td>
<td>.06</td>
</tr>
<tr>
<td>Trail-Making Test B</td>
<td>−.73*</td>
<td>−.07</td>
<td>−.31</td>
<td>−.08</td>
<td>.18</td>
</tr>
<tr>
<td>WCST Factor I</td>
<td>.68*</td>
<td>.10</td>
<td>.21</td>
<td>−.21</td>
<td>.11</td>
</tr>
<tr>
<td>Logical Memory I</td>
<td>.24</td>
<td>.89*</td>
<td>.16</td>
<td>.11</td>
<td>.00</td>
</tr>
<tr>
<td>Logical Memory II</td>
<td>.19</td>
<td>.89*</td>
<td>.21</td>
<td>.08</td>
<td>−.01</td>
</tr>
<tr>
<td>Freedom From Distractibility</td>
<td>.19</td>
<td>.23</td>
<td>.88*</td>
<td>.05</td>
<td>−.01</td>
</tr>
<tr>
<td>Digit-Span Backward</td>
<td>.25</td>
<td>.10</td>
<td>.85*</td>
<td>.05</td>
<td>−.02</td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>.13</td>
<td>.58</td>
<td>.59*</td>
<td>.01</td>
<td>−.07</td>
</tr>
<tr>
<td>WCST Factor II (NPE)</td>
<td>.09</td>
<td>.14</td>
<td>.07</td>
<td>.94*</td>
<td>.04</td>
</tr>
<tr>
<td>WCST Factor III (FMS)</td>
<td>−.02</td>
<td>−.02</td>
<td>−.04</td>
<td>.03</td>
<td>.97*</td>
</tr>
</tbody>
</table>

*Statistically significant loading for this variable according to the Stevens (1992) criteria.

Note. NPE = nonperseverative errors; FMS = Failure-to-Maintain Set.

In an attempt to further examine the construct validity of the factors themselves, the three WCST factor scores, instead of the individual WCST raw scores, were included in a second analysis. A five-factor solution was chosen based on an examination of the scree plot and the variance accounted for. As can be seen from Table 9, the overall structure is identical to that observed in the six-factor WCST raw score solution. The only exception is that WCST Factor I loads moderately on the overall Factor I. Examination of the variables suggests that this represents a general executive function factor. WCST Factors II and III continue to load independent of any other variables.

DISCUSSION

The present paper presents a factor analysis of the WCST in the largest clinical sample reported to date. Analysis of WCST variables alone yielded a three-factor solution. The first factor was labeled concept formation/perseveration while the remaining two factors were identified by the sole variable loading significantly on each (i.e., NPE and FMS, respectively). The composition of Factors I and III (FMS) is consistent with that described in earlier papers. The composition of Factor II (NPE) is less stable, but is consistent with Sullivan et al.’s (1993) third factor. The present structure was also internally consistent and was replicated in three of four diagnostic subgroups. In addition to being both internally and externally reliable, the observed factors also appear to have clinical relevance.

Specifically, each factor reflects one of three qualitatively different types of performance on the WCST. First, a high Factor I score might occur when the patient is unable to shift to a correct sorting principle and thus produces a high number of perseverative responses and completes few categories. Second, a high Factor 2 score may occur because the patient is
inefficient or unsuccessful in problem-solving, while clearly attempting to test different hypotheses (i.e., is not responding inflexibly). This patient is constantly shifting response set, an approach that produces few perseverations while generating large numbers of nonperseverative errors. Finally, a high Factor 3 score should be seen in the patient who discovers the correct sorting principle but has difficulty producing a completed run of 10 consecutive correct responses. It remains to be seen how these different performance styles relate to lesion locus and outcome.

The construct validation factor analysis produced the following factors: (a) general executive function (two parts: WCST concept formation/perseveration and spatial/organizational); (b) verbal new learning; (c) attention/concentration/verbal processing; (d) WCST FMS (WCST Factor III); and (e) WCST NPE (WCST Factor II). The present factor analytic findings are consistent with those of Paolo et al. (1995) and Greve et al. (1995) who found that the WCST scores tended to load independently of other neuropsychological measures. Similarly, Paolo et al. (1995) have noted that the WCST provides independent information about cognitive functioning. Nonetheless, it is somewhat disconcerting that these studies have found only minimal associations between other neuropsychological variables and the various WCST factors. Correlates of the FMS factor have been particularly elusive.

Surprisingly, while Greve et al. (1996) have argued that FMS represents attentional dysfunction, FMS did not to correlate with measures of either working or long-term memory or attention in this sample. This is consistent with the findings of Sullivan et al. (1993), Paolo et al. (1996), and Greve et al. (1995). Since FMS was affected by an experimental manipulation of attention (Greve et al., 1996), it may be that FMS is measuring a different form of attention than does digit span or WAIS-R arithmetic. The likely candidate is sustained attention. Intuitively, this makes sense as FMS is an indication of difficulty maintaining or sustaining responding to a given stimulus dimension for an extended period. Further, in Greve et al. (1996), subjects were assigned to groups based on performance on the Gordon Diagnostic System (GDS; Gordon, 1989), a measure of sustained attention. Since no measure of sustained attention was included in this analysis, an examination of its relationship to FMS may be worthy of future exploration.

In conclusion, this study has replicated the general factor structure of the WCST, which has been reported in several previous papers. Further, it has demonstrated that the WCST makes an independent contribution to neuropsychological evaluation. At the same time, given the limited association of the WCST with other neuropsychological variables both in the present study and in others reported in the literature, it is clear that the nature of the cognitive processes underlying performance on the WCST remains somewhat elusive. Further studies exploring the construct validity of the WCST would be valuable, particularly those that focus on the secondary factors (FMS, NPE). Finally, the relevance of factor analytic studies of the WCST goes well beyond simply attaining a better understanding of its psychometric properties. It is important to study these factors in different diagnostically based samples to clarify how the WCST can assist in the diagnosis of various forms of neuropathology and in the description of their cognitive effects. It is equally important to examine the relevance of these derived variables for rehabilitation and outcome prediction following brain injury.

REFERENCES
Latent Structure 609


