A Review of Neuropsychological Differences Between Paranoid and Nonparanoid Schizophrenia Patients

by Christine Zalewski, Margaret T. Johnson-Selfridge, Steven Ohriner, Karen Zarrella, and James C. Seltzer

Abstract

This review examines the literature on neuropsychological differences between paranoid and nonparanoid schizophrenia subjects. Thirty-two studies related to intellectual functioning, attention, memory, language, visual-spatial, and motor functions are discussed. Subjects with paranoid schizophrenia did not demonstrate higher intellectual functioning than those with nonparanoid schizophrenia, and both groups performed similarly on tests of verbal ability and visual-spatial functions. Several studies suggest that the paranoid subtype is associated with higher performance on tests of executive functions, attention, memory, and motor skills. However, the findings are inconsistent. Methodological issues in the literature are examined, including varying degrees of participants' chronicity and severity of illness among studies, criteria for diagnostic group membership, medication effects, reliability and validity of the neuropsychological measures, and statistical power.

Key words: Neuropsychology, paranoid, nonparanoid.


Studies investigating neuropsychological impairment associated with schizophrenia have consistently identified patterns of cerebral dysfunction associated with this diagnosis (Buchsbaum 1990; Carlsson and Carlsson 1990; Robbins 1990). Although the neuropsychological correlates of schizophrenia have been well researched (Levin et al. 1989; David and Cutting 1994), the considerable symptomatic heterogeneity within the disorder necessitates subtype comparison.

Theoretical and empirical evidence supports the notion that patients with a symptom configuration consisting of predominantly paranoid delusions, in the absence of other signs of thought disorder, represent a distinct taxonomic entity. The distinctiveness of a paranoid subtype can be traced back to the early work of Bleuler (1911/1950) and Kraepelin (1919/1971), through psychodynamic theorists, to more modern investigations (Magaro 1981). Tsuang and Winokur (1974) suggested that a smaller percentage of patients with paranoid subtype had psychomotor symptoms than hebephrenic patients. Paranoid patients were characterized by later onset of illness, less exclusiveness, less distractibility, fewer psychomotor symptoms, a higher incidence of marriage, more children, and less disruption of social and familial relationships. Others have differentiated between positive and negative syndromes, suggesting that deficit symptoms are associated with abnormal brain morphology (Crow 1980; Andreasen and Olsen 1982). Alternative subclassifications of schizophrenia (e.g., process vs. reactive, late vs. early onset, acute vs. chronic course of illness) may also intersect with the paranoid/nonparanoid distinction.

Discussions of cognitive differences, such as attention deficits, among subtypes have appeared in global reviews of neuropsychological correlates of schizophrenia (Levin et al. 1989), leading researchers to suggest that each subgroup of schizophrenia has a unique pattern of neuropsychological impairment. In fact, there is wide clinical acceptance that paranoid schizophrenia patients display less regression of mental faculties than their nonparanoid counterparts (Kaplan and Sadock 1994). Seidman’s (1983) comprehensive review of neurophysiological and neuropsychological findings, as well as Magaro’s (1981) review of information processing in schizophrenia, suggests meaningful distinctions in cognitive functioning among schizophrenia subtypes. However, no recent comprehensive review has specifically

Reprint requests should be sent to Dr. C. Zalewski, Pacific Graduate School of Psychology, 935 East Meadow Dr., Palo Alto, CA 94303.
addressed neuropsychological differences between the paranoid and nonparanoid subgroups.

Identification of subtype-specific cognitive profiles may help refine theories concerning multiple causal pathways in schizophrenia. For example, the vast heterogeneity in schizophrenia might reflect the difference between neurodevelopmental versus neurodegenerative disease processes or stem from a subtype-specific localization of cerebral dysfunction (Bilder et al. 1985). Delineation of the neuropsychological differences among schizophrenia subgroups could enable clinicians to more accurately identify classification boundaries and to develop treatment plans that capitalize on cognitive strengths.

This review examines the findings regarding neuropsychological differences between paranoid and nonparanoid schizophrenia samples from January 1975 to October 1995.

Literature Search and Study Selection

The empirical articles in this review were obtained from two English-language searches—PsycLIT (American Psychological Association 1991) and MedLine (National Library of Medicine 1992)—as well as reference lists from related articles.

A total of 32 studies examining cognitive differences between paranoid and nonparanoid schizophrenia subtypes on standard neuropsychological and intelligence measures were located. Studies of information processing, size estimation, subtype differences in cerebral organization, and interhemispheric transfer of information were excluded (e.g., McDowell et al. 1975; Asarnow and Mann 1978; Brennan and Hemsley 1984; Lubow et al. 1987; Overby et al. 1989; Ditchfield and Hemsley 1990; Carter et al. 1993). However, if a study included scores from a standard administration of a clinical neuropsychological measure, that portion was included in the review (Asarnow and Mann 1978; Pishkin and Lovullo 1986). When more than one study employed the same sample, only the first was reported (Rund 1983; Rund and Blakar 1986).

Findings

General Intellectual Functioning. Approximately one-third of the studies compared paranoid and nonparanoid schizophrenia subjects regarding level of overall cognitive functioning on varied intelligence measures, which ranged from full Wechsler Adult Intelligence Scale instruments (WAIS/WAIS–R; Wechsler 1955, 1981) (Kolb and Whishaw 1983) to vocabulary-based estimators of current intellectual functioning (Cox and Leventhal 1978; see table 1). Most studies failed to demonstrate IQ differences (Hirt et al. 1977; Neufeld 1977; Cox and Leventhal 1978; Kay 1979; Kolb and Whishaw 1983; Lyons and Fulkerson 1984; Wells and Leventhal 1984), although Langell et al. (1987) did report significant differences favoring the paranoid group on the Intellectual Processes Scale of the Luria-Nebraska Neuropsychological Battery (LNNB; Golden et al. 1980). One additional study entered the 11 WAIS subtests and 18 additional variables into a discriminant function analysis predicting subtype group membership (Goldstein and Halperin 1977). The authors reported a correct classification rate of 70 percent based on a discriminant function including the Arithmetic, Information, Comprehension, and Vocabulary subtests among the first 10 variables entered into the equation. However, the absolute differences between the two groups were small (0.2, 0.9, 0.4, and 0.7 scaled score points, respectively).

Four additional studies (Broga and Neufeld 1981; Highgate-Maynard and Neufeld 1986; Dobson and Neufeld 1987; George and Neufeld 1987) found no IQ differences between the groups based on the WAIS–Clarke Vocabulary Scale (Paitich and Crawford 1970), an unpublished instrument described in the manuscripts. However, all four studies excluded participants who obtained WAIS–Clarke scores below 80 (Broga and Neufeld 1981; Dobson and Neufeld 1987; George and Neufeld 1987) or 85 (Highgate-Maynard and Neufeld 1986) for methodological reasons related to their primary research questions. Similarly, Wells and Leventhal (1984) found no Shiley-Hartford Institute of Living Scale IQ (Shipley 1939) differences after excluding subjects with IQ scores below 80. Because of this confounding exclusionary criterion, these results may not reflect true group differences.

Executive Function/Problem-Solving. Only one of five studies examining executive functioning on the Wisconsin Card Sorting Test (WCST; Heaton and Pendleton 1981) found paranoid participants made significantly fewer perseverative errors than nonparanoid participants (Bornstein et al. 1990). However, Rosse et al.’s (1991) findings, which favored the paranoid group, did approach significance ($p = 0.054$). These authors also reported that the paranoid group sorted significantly more categories than the nonparanoid group. In addition, Kremen et al. (1994) cited data, which were not available for this review (Seidman and Kremen 1988), indicating that paranoid patients performed better than nonparanoid participants on the WCST. Goldstein and Halperin (1977) found that their paranoid sample made fewer errors on the Halstead-Reitan Category Test (Halstead 1947; Reitan 1966) than the nonparanoid sample. Unfortunately, the
Table 1. Neuropsychological measures and results

<table>
<thead>
<tr>
<th>Author</th>
<th>Groups (n)</th>
<th>Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asarnow and Mann (1978)</td>
<td>P/S (15) N/S (15)</td>
<td>WAIS-R Vocabulary</td>
<td>No difference on vocabulary test</td>
</tr>
<tr>
<td></td>
<td>Nonpsych (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bornstein et al. (1990)</td>
<td>P/S (28) N/S (27)</td>
<td>Halstead-Reitan Category Test, WAIS-R, WMS-R, Verbal Concept Formatio n Test, WCST</td>
<td>N/S were more consistently impaired than other schizophrenia subgroups; when effects of symptom severity, drug level, and education were statistically controlled, the magnitude and number of differences were substantially reduced</td>
</tr>
<tr>
<td></td>
<td>Schiz/aff (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broga and Neufeld (1981)</td>
<td>P/S (20) N/S (20)</td>
<td>WAIS-Clarke Vocabulary IQ, WAIS Digit Span, categorized list free recall</td>
<td>No IQ or digit span differences; no differences on low-organization word list recall; N/S better able to use organizational strategies in categorized word lists</td>
</tr>
<tr>
<td></td>
<td>Nonpsych (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psych (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dobson and Neufeld (1987)</td>
<td>Epi P/S (14) Rem P/S (14)</td>
<td>WAIS—Clarke Vocabulary IQ</td>
<td>No IQ differences</td>
</tr>
<tr>
<td></td>
<td>Epi N/S (14) Rem N/S (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finkelstein (1983)</td>
<td>P/S (12) N/S (12)</td>
<td>WAIS Vocabulary, auditory distraction test with digit lists</td>
<td>Groups matched by vocabulary scores; no significant differences on the auditory distraction task</td>
</tr>
<tr>
<td></td>
<td>Psych (12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>George and Neufeld (1987)</td>
<td>P/S (14) N/S (14)</td>
<td>WAIS-Clarke Vocabulary IQ</td>
<td>No IQ differences</td>
</tr>
<tr>
<td></td>
<td>Psych (14) Students (14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golden et al. (1980b)</td>
<td>P/S (24) Undiff/S (18)</td>
<td>Luria-Nebraska Neuropsychological Battery</td>
<td>No significant correlations between type of schizophrenia and Luria-Nebraska Neuropsychological Battery scores</td>
</tr>
<tr>
<td>Goldstein and Halperin (1977)</td>
<td>P/S &amp; N/S (140)</td>
<td>WAIS, Halstead-Reitan Category Tests, associated measures</td>
<td>Seventy percent correct classification rate based on discriminant function function</td>
</tr>
<tr>
<td>Gureje (1988)</td>
<td>P/S (36) N/S (34)</td>
<td>Assessment for hand, foot, and eye dominance</td>
<td>P/S showed significantly more mixed-handedness and less cross-dominance than N/S</td>
</tr>
</tbody>
</table>

Note.—See footnotes at end of table.

statistical significance of this measure was not reported, although it was not one of the first 10 test variables to enter into a stepwise discriminant function analysis. Taken together, the findings of executive function studies do not strongly support paranoid subtype superiority, but they do warrant further investigation.
Table 1. Neuropsychological measures and results (Continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Groups (n)</th>
<th>Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamlin and Folsom (1977)</td>
<td></td>
<td>Concrete and abstract vocabulary measures, single and multiple proverbs</td>
<td>N/S, equated for education, were inferior to P/S on total vocabulary and abstract vocabulary; no differences on proverb scores</td>
</tr>
<tr>
<td>Highgate-Maynard and Naufeld (1986)</td>
<td>P/S (20)</td>
<td>WAIS-Clarke Vocabulary IQ</td>
<td>No IQ differences</td>
</tr>
<tr>
<td>Hirt et al. (1977)</td>
<td>P/S (10)</td>
<td>Otis Test of Mental Abilities, eight timed card sorts of increasing complexity</td>
<td>No IQ differences; N/S had significantly longer sort times than P/S</td>
</tr>
<tr>
<td>Kay (1979)</td>
<td>P/S (12)</td>
<td>WAIS</td>
<td>No IQ differences</td>
</tr>
<tr>
<td>Kolb and Whishaw (1983)</td>
<td>P/S (9)</td>
<td>WAIS, WMS, Rey-Osterrieth Complex Figure Test (copy/delayed), Corsi Block Span Test, Mooney Closure Test, Chicago Word Fluency Test, Newcombe Word Fluency Test, Gotman Design Fluency Test, WCST, Left/Right Differentiation Test, Body Placing Test</td>
<td>No differences on any measure</td>
</tr>
<tr>
<td>Kremen et al. (1994)</td>
<td>P/S (11)</td>
<td>WAIS-R subtests, WRAT-R, Rey-Osterrieth Complex Figure Test, WMS-R, Benton Line Orientation Test, WCST, Finger Tapping Test, Nonverbal Cancellations, Auditory CPT</td>
<td>No differences on any measure among DSM-III-R subtypes; patients with a history of systematized delusions had better verbal skills than those without</td>
</tr>
<tr>
<td>Langell et al. (1987)</td>
<td>P/S (45)</td>
<td>Luria-Nebraska Neuropsychological Battery</td>
<td>P/S performed significantly better than N/S on Motor, Rhythm, Receptive Speech, Memory, and Intellectual Processes scales</td>
</tr>
<tr>
<td>Lyons and Fulkerson (1984)</td>
<td>P/S (20)</td>
<td>Ammons' Quick Test</td>
<td>No IQ differences</td>
</tr>
<tr>
<td>Magaro and Page (1983)</td>
<td>P/S (8)</td>
<td>Vocabulary test</td>
<td>P/S significantly better than N/S on vocabulary test</td>
</tr>
</tbody>
</table>

Note.—See footnotes at end of table.

Attention. The studies that assessed attentional skills are equivocal. Those that suggested higher functioning in the paranoid subtype, or a trend toward higher performance, used some form of distraction or interference task. Rund (1983) reported that the paranoid subtype demonstrated superior performance on an attentional digit span task. Relevant digits were presented to the left ear and recited by a female voice, while irrelevant digits were presented by a male voice to the right ear. Rund found that paranoid participants, in general, recalled more digits and were less vulnerable to distraction than their nonparanoid counterparts. Finkelstein (1983) asked participants to
Table 1. Neuropsychological measures and results (Continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Groups (n)</th>
<th>Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manschreck and</td>
<td>P/S (19)</td>
<td>Various measures of motor and sensory function</td>
<td>Pattern of neurological signs did not differ significantly among subtypes</td>
</tr>
<tr>
<td>Ames (1984)</td>
<td>Disorg/S (13)</td>
<td>and laterality</td>
<td>except for the frequency of stereognostic disturbances and left-sided</td>
</tr>
<tr>
<td></td>
<td>Undiff/S (21)</td>
<td></td>
<td>sensory errors, which occurred less frequently in P/S group</td>
</tr>
<tr>
<td></td>
<td>Affective (21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merrin (1984)</td>
<td>P/S (25)</td>
<td>Motor dominance questionnaire, hand grip strength; sight dominance</td>
<td>No differences in motor or sight dominance; right-handed P/S group had greater right-hand grip strength dominance than N/P group; no difference in absolute strength</td>
</tr>
<tr>
<td></td>
<td>N/S (27)</td>
<td>measure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Affective (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasrallah et al.</td>
<td>P/S (27)</td>
<td>Performance laterality scale for eye, hand, and foot dominance</td>
<td>No differences between P/S and N/S in eyedness or footedness; significantly more left-handedness was found in the P/S than in the N/S</td>
</tr>
<tr>
<td>(1982)</td>
<td>N/S (53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neufeld (1977)</td>
<td>P/S (11)</td>
<td>WAIS–Clarke Vocabulary IQ</td>
<td>No IQ differences</td>
</tr>
<tr>
<td></td>
<td>N/S (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paulman et al.</td>
<td>P/S (21)</td>
<td>Luria-Nebraska Neuropsychological Battery, WCST, Finger-Tapping Test</td>
<td>No differences on any measure</td>
</tr>
<tr>
<td>(1990)</td>
<td>N/S (19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pishkin and</td>
<td>P/S (10)</td>
<td>Vocabulary and abstraction scores from the Shipley Institute of Living Scale</td>
<td>No differences</td>
</tr>
<tr>
<td>Lovallo (1986)</td>
<td>N/S (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psych (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosse et al.</td>
<td>P/S (14)</td>
<td>WAIS–R Vocabulary, Peabody Picture Vocabulary Test, WCST</td>
<td>No vocabulary differences; P/S sorted significantly more categories on WCST</td>
</tr>
<tr>
<td>(1991)</td>
<td>N/S (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rund (1983)</td>
<td>P/S (9)</td>
<td>Digit span test with neutral and distractor portions</td>
<td>P/S recalled more digits than N/S; P/S less vulnerable to distraction than N/S</td>
</tr>
<tr>
<td></td>
<td>N/S (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psych (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russell and</td>
<td>P/S (16)</td>
<td>WAIS Vocabulary</td>
<td>No vocabulary differences</td>
</tr>
<tr>
<td>Page (1976)</td>
<td>N/S (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sengel and</td>
<td>P/S (12)</td>
<td>Shipley-Hartford Institute of Living Scale—Vocabulary; recall performance of two word lists</td>
<td>No vocabulary differences; N/P benefited more from cuing than P/S</td>
</tr>
<tr>
<td>Lovallo (1983)</td>
<td>N/S (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depressed (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sengel et al.</td>
<td>P/S (12)</td>
<td>Shipley-Hartford Institute of Living Scale—Vocabulary and abstraction scores; verbal recall under low-, moderate-, and high-interference conditions</td>
<td>Significance of vocabulary and abstraction differences not reported; no difference at low or moderate interference; N/S performed significantly worse than P/S on recall tasks at high interference level</td>
</tr>
<tr>
<td>(1985)</td>
<td>N/S (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Depressed (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manic (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—See footnotes at end of table.

recall only the digits recited by a female voice. The differences were not significant, although a trend was noted for the nonparanoid group to be more distractable than the paranoid group (p < 0.10). Goldstein and Halperin (1977) reported WAIS Digit Span scores for paranoid and nonparanoid groups. Although the significance of this mea-
Approximately half of the studies examining memory found paranoid participants to perform significantly better than nonparanoid participants. Bornstein et al. (1990) found nonparanoid participants to be more impaired than paranoid and schizoaffective participants on percent retention of both verbal and nonverbal components of the Wechsler Memory Scale-Revised (WMS–R; Wechsler 1987). Kremen et al. (1994) reported higher verbal memory skills (WMS–R), but not higher visual memory skills (WMS–R; Rey-Osterrieth Complex Figure Test [Osterrieth 1944]), in patients with a history of systematized delusions, as compared with patients without systematized delusions. However, these differences were not observed when comparing the paranoid and nonparanoid subtypes diagnosed using DSM–III–R criteria (American Psychiatric Association 1987). Langell et al. (1987) found that paranoid participants performed significantly better than nonparanoid participants on the Memory Scale of the LNNB, but Golden et al. (1980b) and Paulman et al. (1990) found no such differences on the LNNB.

Yesavage et al. (1983) tested paranoid and chronic schizophrenia patients on free recall and recognition in both auditory and visual modes. The paranoid and chronic patients were comparable on visual memory task, with both groups performing more poorly than the normal controls. When the words were presented auditorily, only the chronic group performed more poorly than the normal control group. The paranoid and chronic groups also performed comparably with regard to recall, but the paranoid group had significantly higher scores on recognition.

Contrary to these results, Kolb and Whishaw (1983) found that paranoid participants did not differ from nonparanoid participants on visual or auditory memory when assessed using the WMS–R and the Rey-Osterrieth Complex Figure Test. The comparison between subtypes was not the primary objective in this study, however, and the paranoid group consisted of only nine participants.

### Table 1. Neuropsychological measures and results (Continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Groups (n)</th>
<th>Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N/S (10)</td>
<td>Living Scale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psych (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yesavage et al. (1983)</td>
<td>P/S (15)</td>
<td>Recall and recognition of two word lists presented in auditory and visual modes</td>
<td>No differences on visual or recall tasks; P/S performed better on auditory and recognition tasks</td>
</tr>
<tr>
<td></td>
<td>Undiff/S (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonpsych (15)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes.*—P/S = paranoid schizophrenia; N/S = nonparanoid schizophrenia; N/S+ = nonparanoid schizophrenia with a history of systematized delusions; N/S− = nonparanoid schizophrenia without a history of systematized delusions; Psych = psychiatric control group; Nonpsych = nonpsychiatric control group; Schiz/aff = schizoaffective group; Epi = episodic; Rem = remitted; Disorg/S = disorganized schizophrenia; Undiff/S = undifferentiated schizophrenia. Scales: Ammons Quick Test (Ammons and Ammons 1962); Benton Line Orientation Test (Benton et al. 1978); Body Placing Test (Semmes et al. 1963); Chicago Word Fluency Test (Milner 1984); Corsi Block Span Test (Milner 1971); CPT = Continuous Performance Task (Rosvold et al. 1956); Digit Span Test (Oltmanns and Neale 1975); DSM–III–R = Diagnostic and Statistical Manual of Mental Disorders—Revised (American Psychiatric Association 1987); Finger Tapping Test (Reitan and Wolfson 1985); Gotman Diagnostic Memory Test (Jones-Gotman and Milner 1977); Halstead-Reitan Category Test (Halstead 1947, Reitan 1966); Left/Right Differentiation Test (Semmes et al. 1963); Luria-Nebraska Neuropsychological Battery (Golden et al. 1980a); Mooney Closure Test (Mooney 1956); Newcombe Word Fluency Test (Newcombe 1969); Nonverbal Cancellations (Lezak 1983); Otis Test of Mental Abilities (Otis 1962); Peabody Picture Vocabulary Test (Dunn 1965); Rey-Osterrieth Complex Figure Test (Osterrieth 1944); Shipley-Hartford Institute of Living Scale (Shipley 1939); Verbal Concept Formation Test (Bornstein and Leason 1985); WAIS-Clarke Vocabulary IQ (Paitich and Crawford 1970); WCST = Wisconsin Card Sorting Task (Heaton and Pendleton 1981); Wechsler Adult Intelligence Scale (WAIS/WAIS–R; Wechsler 1955, 1981); WMS–R = Wechsler Memory Scale–Revised (Wechsler 1987); WRAT–R = Wide-Range Achievement Test–Revised (Jastak and Wilkinson 1984).
Several other studies (Broga and Neufeld 1981; Sengel and Lovallo 1983; Sengel et al. 1985) showed some paranoid versus nonparanoid distinctions, but the results are difficult to compare across investigations because nonstandard assessment tasks were used. Broga and Neufeld (1981) asked participants to recall words from visually presented low-organization word lists and categorized word lists. Recall did not differ between paranoid and nonparanoid participants on the low-organization word list, but nonparanoid participants were better able to use organizational strategies in the categorized word lists (i.e., under some conditions the nonparanoid patients performed better than the paranoid subtype). Sengel and Lovallo (1983) reported that nonparanoid participants benefited more from cuing than paranoid participants as measured by the number of words recalled. Sengel et al. (1985) found that paranoid and nonparanoid groups did not differ in verbal recall at low and moderate interference levels, but that nonparanoid participants performed significantly worse than paranoid participants at the high interference level.

In summary, there is a trend in the literature suggesting that, under certain conditions, the paranoid subtype may demonstrate better memory functioning than the nonparanoid subtype. However, the complex interactions among assessment variables, including presentation (e.g., visual vs. auditory), stimuli (e.g., verbal vs. nonverbal), and interference (high vs. low), preclude any simple conclusions.

**Verbal Abilities.** Most of the studies measured some form of verbal ability, most often vocabulary, which was tested with several standardized instruments: the Vocabulary subtest from various versions of the WAIS, the Shipley-Hartford Institute of Living Scale Vocabulary subtest (Shipley 1939), and the Mill Hill Vocabulary Test (Raven 1982). Of the investigations specifically examining vocabulary test scores (Russell and Page 1976; Hamlin and Folsom 1977; Neufeld 1977; Asarnow and Mann 1978; Cox and Leventhal 1978; Kay 1979; Magaro and Page 1983; Sengel and Lovallo 1983; Sengel et al. 1985; Fishkin and Lovallo 1986; Rosse et al. 1991), only two (Hamlin and Folsom 1977; Magaro and Page 1983) reported significant differences, with the paranoid subtype having better vocabulary performance in each case. As mentioned earlier, Goldstein and Halperin (1977) reported mean subgroup differences of under one scaled score point on the WAIS Vocabulary subtest, and five others found no differences after using vocabulary-based IQ exclusionary criteria (Broga and Neufeld 1981; Wells and Leventhal 1984; Highgate-Maynard and Neufeld 1986; Dobson and Neufeld 1987; George and Neufeld 1987). Finkelstein (1983) was excluded from consideration because the schizophrenia groups were matched to controls according to their vocabulary score.

Although most studies focused on vocabulary, several other verbal skills were examined as well. Kolb and Whishaw (1983) found no verbal fluency differences between the subgroups. Bornstein et al. (1990), on the other hand, reported significant verbal fluency group differences, with the schizoaffective and nonparanoid groups performing most poorly, followed by the paranoid and control groups, respectively. These differences did not remain after covariation for medication level, total symptoms, and education, however. Similarly, these authors reported no mean verbal IQ differences after covariation, although the nonparanoid subgroup obtained the lowest score prior to covariation. Of the studies using the LNNB Expressive Speech, Receptive Speech, Writing, and Reading subscales, none found significant differences between the subtypes (Golden et al. 1980b; Paulman et al. 1990) with the exception of Langell et al. (1987), who found that the paranoid group performed better on Receptive Speech and reported a nonsignificant trend toward better performance on Expressive Speech and Writing. Finally, Kremen et al. (1994) found that patients with systematized delusions demonstrated better verbal abilities (a composite score from WAIS–R Vocabulary and Similarities subtests and WRAT–R Spelling) than participants without systematized delusions. However, there were no significant differences when the sample was defined by DSM–III–R subtype criteria.

Generally, the studies under review do not suggest that paranoid and nonparanoid schizophrenia subjects perform differently on standardized tests of verbal skills. This finding is particularly striking given that aberrant verbal behavior (e.g., marked loosening of associations) is a common criterion for assignment to the nonparanoid group. The tests commonly used to assess verbal skills do not appear to be sensitive to the disturbances in thought that typically manifest in patients' speech.

**Visual-Spatial Skills.** Little evidence was presented to suggest that subgroups differ on general measures of visual-spatial ability, and inconsistent findings were reported for specific visual-spatial tasks. Bornstein et al. (1990) found that nonparanoid participants had the lowest mean WAIS–R Performance IQ score, followed by paranoid, schizoaffective, and control participants; however, this difference was not significant when total symptoms, medication level, and education were controlled. Kay (1979) and Kolb and Whishaw (1983) did not find significant differences between paranoid and nonparanoid groups on WAIS Performance IQ. The three studies using
the LNNB Visual subscale (Golden et al. 1980b; Langell et al. 1987; Paulman et al. 1990) reported no subtype differences, although Langell et al. (1987) did note a trend favoring the paranoid group.

Some subtype differences were found on other perceptual measures. Hirt et al. (1977) found that nonparanoid participants required longer overall times than paranoid participants on increasingly complex sorting tasks. As task complexity increased, the nonparanoid participants were first to experience deterioration in performance. Kolb and Whishaw (1983) found no differences on the copy component of the Rey-Osterrieth Complex Figure Test, the Mooney Closure Test (Mooney 1956), or the Gotman Design Fluency Test (Jones-Gotman and Milner 1977).

Motor. Studies comparing the motor functioning of paranoid and nonparanoid schizophrenia subjects yielded inconsistent results in most domains (Goldstein and Halperin 1977; Golden et al. 1980b; Nasrallah et al. 1982; Manschreck and Ames 1984; Merrin 1984; Langell et al. 1987; Gureje 1988; Bornstein et al. 1990; Paulman et al. 1990; Kremen et al. 1994). Bornstein et al. (1990) found the nonparanoid subtype performed worse than all other groups on the Grooved Pegboard (Matthews and Klove 1964) and Finger Agnosia tasks (Boll 1974). Goldstein and Halperin (1977) reported greater dominant-hand grip strength in the paranoid group than in the nonparanoid group, although the significance of this variable was not reported. Also in this study, a variable representing the length of time needed to write the word “television” with the nondominant hand was the first to enter into a stepwise discriminant function analysis that correctly predicted 70 percent of paranoid and nonparanoid participants. Merrin (1984) found that paranoid and nonparanoid inpatients did not differ in absolute grip strength, but that right-handed paranoid patients exhibited greater right-hand grip strength dominance than nonparanoid patients.

Neither Paulman et al. (1990) nor Kremen et al. (1994) found subtype differences on the Finger-Tapping Test. Similarly, Manschreck and Ames (1984) found no differences among subtypes on neurological measures of motor functioning, although paranoid subjects had lower frequencies of certain sensory deficits. Langell et al. (1987) found that the paranoid group performed better on the LNNB Motor Scale than the nonparanoid group but Golden et al. (1980b) and Paulman et al. (1990) found no subtype differences on this scale.

Studies of motor dominance likewise produced variable results. Gureje (1988) noted more mixed-handedness in paranoid participants as compared with nonparanoid participants. Nasrallah et al. (1982) reported significantly more left-handedness in paranoid participants. Merrin (1984), however, did not find any group differences in handedness among psychiatric inpatient subtypes.

Computerized Taxonomic Procedures. In addition to the studies listed in table 1, several researchers have examined the neuropsychological profiles of schizophrenia subtypes without specifically comparing paranoid and nonparanoid subgroups. These investigations described how multivariate relationships among symptoms and neuropsychological measures can be used to generate subtype distinctions in contrast to the more traditional, and possibly artificial, diagnostic subtypes (Bilder et al. 1985; Liddle 1987; Gruzelier et al. 1988; Liddle et al. 1989). The findings from these studies suggest that three or more subgroups may exist: (1) a globally impaired group characterized by avolation, social withdrawal, and affective flattening; (2) a group with less global dysfunction and a general disorganization of behavior and thought (rather than a general decrease in behavior), particularly alogia, bizarre behavior, attentional impairments, and positive formal thought disorder; and (3) a neuropsychologically intact group characterized by systematized paranoid delusions. While these results cannot be compared directly with the neuropsychological profiles of paranoid and nonparanoid schizophrenia, the trend suggests that the paranoid group has less cognitive impairment.

Methodological Issues

Although paranoid schizophrenia subjects appear to demonstrate more intact functioning and a distinct cognitive style in some samples, the lack of consistent results is troubling. Characteristics other than subgroup membership may contribute to these inconsistencies. Factors likely to affect the findings of differential neuropsychological performance have been organized into nine primary categories: (1) diagnostic criteria; (2) severity of illness; (3) chronicity of illness; (4) state versus trait effects; (5) medication effects; (6) sampling bias; (7) criterion measures; (8) sample size, effect size, and power; and (9) the matching fallacy.

Diagnostic Criteria. Both the criteria used to establish the presence of schizophrenia and the manner in which the schizophrenia subtypes are determined may affect the neuropsychological profiles of the subgroups. Given the changes in the standard classification systems over time, the comparability of earlier studies to more recent efforts is limited.

The diagnostic systems used for the assignment of participants to the paranoid and nonparanoid subgroups
vary widely across the literature. Before the 1980s, studies frequently failed to report specific diagnostic criteria. More recent investigations carefully integrate standardized procedures associated with specific classification systems and report interrater reliability coefficients (Kremen et al. 1994). The procedures used in most studies range from objective diagnostic interviewing or clinical judgment based on standard classification systems (Kolb and Whishaw 1983; Bornstein et al. 1990) to clinical judgment in the absence of any reported classification system (Cox and Leventhal 1978). Studies that objectively measured subgroup symptomatology most often used the Maine Scale for Paranoid and Nonparanoid Schizophrenia (Magaro et al. 1981) and the Positive and Negative Syndrome Scale for Schizophrenia (Kay et al. 1987).

A related issue involves the placement of all nonparanoid subtypes into a single group. Most studies classified participants as either paranoid or nonparanoid, with far fewer examining specific subtypes within the nonparanoid group (e.g., hebephrenic, catatonic, and undifferentiated), despite cautions against grouping all nonparanoid subtypes together (Berkowitz 1981).

The type of diagnostic system used may also influence the severity of symptoms present in a particular group. Sørensen et al. (1988) suggest that applying certain criteria may artificially restrict the amount of nonparanoid behavior present in participants designated as “paranoid,” which could result in an overrepresentation of severely impaired participants in the nonparanoid group. Subjects with paranoid schizophrenia are often identified as having paranoid delusions without thought disorder or other behavioral disorganization. Thus, it may be the absence of these latter symptoms, rather than the presence of paranoia, that is related to cognitive functioning and cerebral pathology.

Conversely, Kremen et al. (1994) suggest that the presence or absence of systematized delusions, rather than the traditional DSM–III–R subtype criteria, should be used to divide the subgroups. They divided a sample of patients based on the presence or absence of systematized delusions (i.e., some patients in the DSM–III–R nonparanoid group were combined with the paranoid group due to a history of systematized delusions, even though they also demonstrated other disorganized or undifferentiated subtype symptoms). The schizophrenia patients with a history of systematized delusions demonstrated superior verbal abilities and a trend toward better verbal memory compared with those without a history of systematized delusions. However, no neuropsychological differences emerged between subtypes when the DSM–III–R criteria were applied. Given this promising line of research, comparisons of illness-related variables among groups that are distinguished by cognitive and symptom profiles should be pursued (Bilder et al. 1985; Liddle et al. 1989).

Finally, attention to concomitant diagnosis is remarkably lacking in this literature. Although several studies excluded subjects who met criteria for additional psychiatric diagnoses, such as affective or schizoaffective disorders (e.g., Merrin 1984), the majority reported no information related to potentially confounding comorbidity. The occurrence of schizomimetic symptoms in psychotic depression, personality disorders, bipolar illness, substance abuse, and other nonpsychiatric central nervous system disorders heightens the need for clear diagnostic procedures.

**Severity of Illness.** Many studies described the treatment setting from which the sample was recruited as a gross measure of severity of illness. Most studies used inpatient samples, followed by mixed or separate inpatient/outpatient samples (Golden et al. 1980b; Rund 1983; Manschreck and Ames 1984; Dobson and Neufeld 1987) and primarily outpatient samples (Wells and Leventhal 1984; Bornstein et al. 1990). Several studies gave few or no details regarding setting (e.g., Rosse et al. 1991).

As discussed earlier, cognitive differences between subtypes may reflect lower illness severity in the paranoid group as opposed to fundamental correlates of cerebral pathology. In the study by Bornstein et al. (1990), covariation of illness severity, as well as medication levels and education, attenuated subtype differences. These results suggest that neuropsychological findings may reflect the severity of psychosis rather than differential functioning of psychiatric subtypes. However, controlling for severity of symptoms may obscure meaningful subtype distinctions. In addition, Buchanan et al. (1994) found that changes in symptom rating scales and changes in neuropsychological performance were not significantly correlated.

Although further investigation of the relationship between symptom severity and cognitive test performance is needed, the measurement of severity of illness itself may be problematic. The validity of scales using linear or additive strategies (e.g., Brief Psychiatric Rating Scale [BPRS; Overall and Gorham 1962]) is questionable. The presence of extreme forms of any one symptom may indicate high severity not reflected in scale elevation due to the absence of other symptoms. For example, a BPRS total score of 12 (e.g., extremely severe suspiciousness = 6, extremely severe unusual thought content = 6) may reflect a more debilitating illness than the higher total score of 20 (e.g., mild somatic concern = 2, moderate anxiety = 3, moderate conceptual disorganization = 3, mild guilt = 2, moderate tension = 3, moderate
depression = 3, mild blunted affect = 2, and mild uncooperative ness = 2). Scales measuring more specific syndromes or adaptive functioning may provide an alternative. Examples of these scales include the BPRS Syndrome Scales (Overall and Gorham 1962), the Positive and Negative Syndrome Scales (Kay et al. 1987), and several "quality of life" and "current adjustment" scales reviewed by Corrigan (1989).

Because setting-related patient demographics have altered considerably in recent years due to changes in mental health care provision, generalization of findings across the literature may be limited. For instance, partial hospitalization programs now provide treatment for many patients who previously would have remained in inpatient units. Thus, comparisons across studies may be complicated by differences in severity of illness, particularly if samples were drawn from different treatment settings or from similar settings at different times.

Chronicity of Illness. Chronicity of illness may also influence results related to neuropsychological testing. Patients with a long history of illness, especially institutionalized individuals, may demonstrate fatigue, lethargy, passivity, and other acquired behavioral traits associated with a persistent disorder. Goldstein and Halperin (1977) assessed cognitive abilities in a sample that was grouped and analyzed according to three criteria: paranoid versus nonparanoid, neurologically normal versus neurologically abnormal, and long-term versus short-term institutionalization. They suggested that length of institutionalization may be more closely associated with cognitive abilities than either subtype or positive neurological evaluations. Yesavage et al.'s (1983) findings of impaired recall in chronic undifferentiated schizophrenia subjects, as compared with paranoid schizophrenia subjects of unspecified chronicity, may also be due to chronicity effects rather than a subtype-specific impairment.

In the present review, various measures of chronicity were used, such as duration of illness, length of hospitalization, and number of hospitalizations. Several studies excluded participants based on years of total hospitalization (e.g., Magaro and Page 1983; George and Neufeld 1987). Some studies described their samples as acute, chronic, or mixed, and a minority provided means and standard deviations for duration of illness, length of hospitalization, and/or number of hospitalizations.

Many of these factors are difficult to compare across studies due to recent changes in clinical practice and insurance coverage for psychiatric illness. For example, a sample with higher numbers of hospitalizations may reflect the trend toward shorter lengths of stay and increased readmission rates rather than a more chronic or treatment-resistant group. Thus, time since onset of illness may be more useful to assess chronicity than duration of treatment or number of hospitalizations.

State Versus Trait Effects. As Levin et al. (1989) have noted, the issue of whether deficits found through testing are state- or trait-related is important. This concern applies not only to the need to control for the effects of group differences with regard to clinical state, but also to the stability of subtype classifications. Schizophrenia symptoms change over time, potentially altering an individual's subtype diagnosis (i.e., a paranoid schizophrenia patient might be placed in the nonparanoid group as the illness progresses). Unfortunately, a minority of the studies employing inpatient samples specified the time of testing relative to the current hospitalization.

In addition, the severity of schizophrenia symptomatology is only one state variable that can affect neuropsychological functioning. Others include levels of anxiety, depression, fatigue, and motivation, none of which are routinely assessed in the literature. Ideally, the effects of concomitant state variables should be carefully considered when evaluating subtype differences.

Medication Effects. Because observed cognitive functioning may reflect a state-related medication effect (Sweeney et al. 1991), knowledge of medication levels across the subgroups is important. In a recent review, Cassens et al. (1990) found that chronic neuroleptic administration may improve performance on tasks requiring sustained attention and visuomotor problem-solving abilities. Acute administration of neuroleptic medications did, however, impair performance on some tasks requiring sustained vigilance, attention, and motor abilities. It is unclear whether improvements resulted directly from drug effects or were secondary to improved clinical status. The anticholinergic effects of neuroleptic drugs and anticholinergic medications may also impair neuropsychological function, particularly memory (Cassens et al. 1990; Sweeney et al. 1991).

Most recent investigations have documented neuroleptic medication dosage, and approximately half of all studies either assessed differences in group mean dosage or correlated dosage with performance on neuropsychological measures. These studies have typically used prescribed dosage in chlorpromazine (CPZ) equivalents as a measure of medication amount. However, the CPZ equivalents may not be useful for some medications (e.g., risperidone and haloperidol decanoate). Most of the studies that specifically examined the effects of medication on dependent variables found no significant medication effects (e.g., Fishkin and Lovallo 1986).

Spohn and Strauss (1989) have suggested that indices such as dose per day or dose per body weight per day are
poor indicators of physiological drug levels. An analysis of neuroleptic levels in serum may be more useful; however, blood levels may not reflect the potency of the medication. Also, patients selected from day treatment or out-patient settings may not adhere to their prescriptions (i.e., prescribed dosage cannot be used as a measure for actual dosage). In addition, the effects of nonstandard neuroleptics (e.g., clozapine, risperidone) on cognitive performance need to be clarified. Studies done before the use of these atypical neuroleptics may not be comparable to current research.

Sampling Bias. The potential for sampling bias is difficult to determine in this literature because so few studies described procedures for subject selection. Results from studies using consecutively admitted patients (e.g., Nasrallah et al. 1982; Kolb and Whishaw 1983; Kremen et al. 1994) could differ dramatically from those studies using widely solicited volunteers, clinical referrals, chart-based retrospective analyses, or some combination of these (e.g., Golden et al. 1980b). For example, participants assessed for clinical reasons may demonstrate more severe cognitive deficits than those randomly selected for research (i.e., neuropsychological referrals are often based on professional concern regarding cognitive functioning). Further, patients who refuse to complete the testing may be more paranoid or socially withdrawn than those who consent. Other bias may occur if samples have skewed distributions regarding gender, genetic risk (e.g., family history of schizophrenia), premorbid functioning, or other related variables. Comparing the selected participants to the excluded participants on a range of demographic variables (e.g., age at onset of illness, schizophrenia subtype, severity of illness) may help delineate potential sampling bias.

The variation in application of exclusionary criteria also increases the potential for sampling bias. More than two-thirds of the studies used specific inclusion/exclusion criteria for subject participation to control for potential confounding influences on the neuropsychological measures. These criteria excluded patients based on the presence of conflicting psychiatric conditions, age, presence of central nervous system disorders or injuries, presence of substance abuse, IQ, duration of treatment, chronicity of illness, or history of electroconvulsive therapy. However, the criteria were not uniformly employed across studies, and the actual number of participants excluded from each study was rarely reported.

Approximately half of the studies excluded participants for organic dysfunction, including brain damage, neurological disease, epilepsy, electroencephalogram abnormalities, and lobotomy. However, a striking minority used all of these variables as exclusionary criteria (e.g., Broga and Neufeld 1981; Kremen et al. 1994). The procedures for determining the presence of organic dysfunction were rarely documented, and reliance on medical records or structured interviews may not provide reliable assessments of the presence or history of these conditions.

Neuropsychological profiles among schizophrenia subtypes may also be altered by substance abuse or dependence, yet a minority of studies included participants on the basis of drug and alcohol abuse. In more recent investigations, patients with comorbid substance abuse were typically excluded, although specific criteria related to substance usage varied across studies. The exclusion of all patients with a history of substance abuse may be problematic in that it would limit the generalizability of findings and reduce the number of patients available for research. Clearly, however, the interactions among substance use, schizophrenia subtype, and cognitive functioning warrant empirical investigation.

Additional exclusionary criteria included age and age at onset of illness. For example, several studies (Highgate-Maynard and Neufeld 1986; George and Neufeld 1987) excluded participants under 18 or over 60 years of age. Although exclusion of patients older than 60 years may reduce the effects of aging on cognitive performance, it may also affect selection with regard to other factors, such as severity.

Age at onset of illness was documented by approximately one-third of the studies reviewed. However, reliable determination of age at onset is difficult in retrospective studies. In addition, age at onset of the full syndrome and age at onset of prodromal symptoms were rarely distinguished. Differences in the age at onset between subgroups or across studies may be critical in understanding the neurodevelopment of schizophrenia. Further, the relationships among age of onset and potentially confounding variables such as education warrant careful attention, particularly as they relate to those instruments measuring general fund of knowledge (e.g., vocabulary).

Several investigations excluded participants with low IQ scores (e.g., Broga and Neufeld 1981) and those with mental retardation (e.g., Finkelstein 1983). Exclusion of participants on the basis of low IQ may enhance the interpretability of neuropsychological test results by omitting the confounding presence of premorbid cognitive deficits. However, this practice may also artificially and differentially reduce the number of severely cognitively impaired patients within subgroups. More obviously, the absence of subgroup IQ differences in those studies excluding low IQ participants for methodological reasons should not be overinterpreted as evidence of subtype comparability (e.g., Dobson and Neufeld 1987; George and Neufeld 1987).
Criterion Measures. The tests used as criterion measures vary considerably in their level of standardization and in the specific processes they are intended to address (see table 1). In addition, the assignment of individual tests to specific areas of cognitive function is often problematic. For instance, common digit span tests could be considered measures of attention or of immediate recall. Examination of subtype differences on specific cognitive tasks may be overly simplistic because many functional areas are interrelated (e.g., verbal IQ and auditory attention are both related to verbal list learning).

The reliability and validity of the measures used by the studies in this review varied considerably. Some measures, such as the WAIS-R, WAIS, LNNB, and the Halstead-Reitan Category Test, have been well standardized and clinically utilized. The reliability and validity of other measures, such as certain card-sorting tasks (Hirt et al. 1977), was less clear. In addition, some tests may have adequate reliability when administered to nonpsychiatric populations but will demonstrate poor test-retest reliability when given to schizophrenia patients. The use of such measures makes comparisons among studies, interpretation of results, and replication of findings more difficult.

Levin et al. (1989) also noted that the neuropsychological performance of schizophrenia subjects is often compared with the performance of adult neurological patients in an attempt to localize areas of brain dysfunction. The dysfunction in schizophrenia, however, may be developmental and not comparable to acquired brain dysfunctions. Furthermore, studies using only single measures are difficult to interpret without knowledge of other areas of functioning. More consistent use of standardized batteries of tests with normative data may greatly enhance this area of research.

Sample Size, Effect Size, and Power. Traditionally, behavioral science researchers have ignored the importance of statistical power analysis (Cohen 1992), and the current literature, unfortunately, does not deviate from this standard. As a simplified example, consider a three-group analysis of variance (ANOVA) that tests for differences on a given neuropsychological measure. In this analysis, 52 participants per group are needed to achieve power of 0.80, assuming a medium effect size and alpha = 0.05 (Cohen 1992). If medium effect size differences actually existed in the population, the chances of obtaining significant results would be about 8 in 10. However, Cohen (1988) concludes that medium effect sizes are uncommon in the behavioral sciences, and power estimates based on them may be optimistic. In an ANOVA using 20 participants per group, the power (assuming medium effect size and alpha = 0.05) would be 0.17. Even if true population differences at the medium-effect-size level existed, there would be only one chance in six of rejecting the null hypothesis (i.e., no group differences). The power estimate drops considerably to 0.09 when hypothesizing a small effect size; that is, less than 1 chance in 10. The issue is not irrelevant to the current literature; most of the studies reviewed used groups of 20 or fewer participants, and many employed groups of 15 or fewer. Concluding that no group differences exist based on analyses conducted with insufficient power is both premature and misleading.

A related issue in this literature is the high probability of committing a Type I error, or falsely rejecting the null hypothesis. This concern is particularly salient in the many studies that employ numerous univariate analyses. Unfortunately, applying a correction for high familywise error, such as a Bonferroni adjustment, is directly related to decreased power, which many studies cannot afford. Similarly, multivariate methods also reduce power, although they may offer a more sophisticated approach to analyzing the data. Overall, the high prevalence of small sample sizes, coupled with the widespread practice of employing numerous univariate analyses, increases the probability of both Type I and Type II error.

In addition to Type I and Type II error problems, variation in sample sizes and statistical techniques (e.g., multiple univariate vs. multivariate, correction vs. no correction for familywise error) restricts cross-literature comparisons. For instance, Langell et al. (1987) reported 5 of 14 LNNB measures to differ between the paranoid and nonparanoid groups after applying a Bonferroni adjustment that reduced alpha from 0.05 to 0.0036. Had the authors neglected to adjust for familywise error and retained a 0.05 alpha level, 9 of the 14 measures would have “differed significantly.” Thus, simple calculations of the percentage of studies with significant findings may be misleading.

Application of meta-analytic techniques involving the comparison of effect sizes across studies can help to summarize the literature more accurately. Unfortunately, the high variability of methods and measurements relative to the small number of studies reporting usable data within each cognitive domain precluded a meta-analytic approach to this review. However, two of the domains contained sufficient numbers of calculable effect sizes to allow an informed examination of the issue. To maximize the number of studies for which an exact effect size could be calculated, and to improve cross-literature comparisons of effect size, r was calculated instead of d by the procedures outlined in Rosenthal (1991).

Within the domain of general intellectual functioning, 13 r values were calculable, ranging from -0.28 to +0.34
Neuropsychological Differences

... (with positive values reflecting greater impairment in the nonparanoid group). Eight of the values were positive, and nine were within the range of -0.15 to +0.15. Within the domain of verbal abilities (which partly overlaps with the domain of general intellectual functioning due to the use of some verbal measures as general IQ estimates), 17 r’s were calculable, ranging from -0.28 to +0.33. Twelve of these values were positive, and more than half were within the range of -0.15 to +0.15. When considering all of the domains together, only a small minority of studies reported differences greater than r = 0.30 (a medium effect size), generally considered to be the smallest absolute value to have clinical meaning with regard to psychological assessment (e.g., Greene 1987).

Matching Fallacy. Meehl (1970) delineated several inherent problems associated with attempts to balance groups with respect to various nuisance variables, collectively referenced in the literature as the “matching fallacy.” In schizophrenia subtype research, the relevance of the matching fallacy can be summarized as follows: When researchers attempt to either match groups or statistically control for nuisance variables that are not extraneous to the subjects’ clinical history, they may unwittingly mismatch the groups on other important factors or sample from subpopulations that differ in some systematic way from the true population. In the schizophrenia literature, the matching fallacy as it relates to education has received notable attention; educational differences have been consistently demonstrated to be a likely consequence of the disease process (Resnick 1992; Kremen et al. 1995). Even variables such as age, which cannot be logically considered a consequence of the disease process, should be managed thoughtfully. For instance, if one subtype is characterized by a later onset of illness (e.g., paranoid), matching on age will result in a systematic mismatch on duration of illness (i.e., the groups will have the same mean age, but the nonparanoid group will have a longer mean duration of illness).

Integration. Synthesis of results across studies requires simultaneous consideration of methodological variation in statistical procedures and sample size and treatment of such factors as diagnostic criteria, clinical state, setting, age, gender, medication, premorbid functioning, duration of illness, and age at onset. Unfortunately, the number of methodological differences and their interactions surpass the number of studies in most domains, substantially limiting the adequacy of any comprehensive synthesis.

Consideration of several of the most obvious divisions according to methodology failed to offer further useful information. For example, within each domain, those studies employing DSM-III-R did not yield substantially different results, as a group, than those using other criteria. Similarly, examination of effect sizes (presumably negating differential results due to variation in sample size and power) did not alter the general conclusions of this review. Finally, examination of those studies that reported positive results within each domain did not appear to differ from the remaining studies in any meaningful way with regard to sample demographics, clinical setting, medication, or chronicity.

Summary

The notion of higher IQ in paranoid schizophrenia subjects as compared with nonparanoid schizophrenia subjects was commonly examined in this literature; however, little empirical support was found for this hypothesis. Similarly, no consistent support was found for subtype differences in verbal abilities or visual-spatial skills. Limited support for neuropsychological differences favoring the paranoid group was found in executive functioning/problem-solving, attention, memory, and motor skills.

The present review offers only modest support for the notion of cognitive differences associated with paranoid and nonparanoid schizophrenia. Given the inconsistencies across studies, statements concerning subtype differences in localization of cerebral dysfunction or subgroup-specific etiological pathways based on neuropsychological findings are premature at this time. Whereas the paranoid-nonparanoid distinction has been supported in other realms, such as course of illness, premorbid functioning level, and familial tendency (McGlashan and Fenton 1991), evidence from the neuropsychological realm is not consistent.

Central to a discussion of these findings is the methodological variability evident in the literature. The wide range of diagnostic criteria employed and the varying degrees of reliability and validity among the neuropsychological measures limit both interpretation and comparisons across studies. Subject variables, such as severity of illness, clinical state, medication effects, and sampling bias, also complicate reconciliation of conflicting results. The suboptimal statistical power evident in many of the studies further limits clear interpretation of the findings.

Based on the present review, several recommendations for future research are suggested. First, sufficient samples sizes are vital to ensure adequate statistical power. Clearly, the literature does not support the expectation of large effect sizes between paranoid and nonparanoid schizophrenia groups on standard neuropsychologi-
neuropsychological test performance across samples might be more easily interpreted. Other guidelines might include (1) use of structured clinical interviews using rigorous criteria can facilitate comparisons across studies, such as core procedures or measures to be used when assessing neuropsychological functioning in schizophrenia. For instance, if all studies uniformly report the mean time since onset of the full schizophrenia syndrome, then the effects of length of illness on neuropsychological test performance across samples might be more easily interpreted. Other guidelines might include (1) use of structured clinical inter-

views and validated diagnostic systems for selecting participants; (2) assessment of the severity of illness and adaptive functioning; (3) application of specific operational definitions for exclusion criteria (e.g., age, substance abuse, organicity); (4) consistent reporting of CPZ and anticholinergic dose equivalents prescribed to the participants; (5) estimates of general intellectual functioning (e.g., WAIS–R Full-Scale IQ or Shipley-Hartford Institute of Living Scale); (6) when possible, indicators of premorbid functioning; and (7) matching subjects on variables that are extraneous to their clinical history (e.g., parental socioeconomic status).

Continued research aimed at specifying neuropsychological profiles among schizophrenia subtypes may provide useful information in identifying pathological cerebral processes, but the paranoid versus nonparanoid dichotomy may be too simplistic to yield meaningful results. Future research might define subtypes by simultaneously integrating a larger number of clinical variables (e.g., course of illness, positive/negative symptoms, comorbidity), physiological findings, genetic information (familial history of schizophrenia), treatment response (e.g., medications and rehabilitation), and neuropsychological measures.

Although advances in neuroimaging provide more direct measures of cerebral pathology, the indirect inferences of brain dysfunction made from neuropsychological assessments may nevertheless be useful in identifying etiological subtypes. For example, complex or subtle cerebral dysfunctions may be more readily identified based on behavioral manifestations evident in cognitive test performance than on neuroimaging output.

In addition, tests should be selected based on ecological validity, or measures of the patients’ adaptive functioning should be included to provide more useful information for treatment.

Future studies might also examine a wider spectrum of disorders. Instead of limiting the population studied to schizophrenia and/or schizoaffective diagnoses, the population might be defined as any patient with a history of psychosis (including psychotic affective disorders). Because many neuropsychological and physiological abnormalities found in schizophrenia may not be specific to the disorder, identification of neuropsychological correlates of paranoia, per se, may provide additional information.

Finally, researchers must acknowledge the substantial limitations in methodology inherent in work with the schizophrenia population. For example, it is unreasonable to expect that age at onset of illness or history of specific symptoms can be identified retrospectively. Similarly, structured clinical interviews using rigorous criteria can
promote highly reliable judgments but do little to ensure the validity of subtype diagnosis. Given that many of the patients have substantial mistrust, poor insight, and a variety of deficits in cognitive processes, it may be unreasonable to expect them to provide researchers with accurate data for symptom ratings. The inconsistencies in findings across studies may reflect these limitations. The longitudinal high-risk studies represent one solution to the difficulties of characterizing the clinical course of illness.

There are two possible solutions to problems of validly rating the presence and quality of delusional thought content, other positive symptoms, and affective states over the course of illness. First, expand the populations studied to include a broader spectrum of related psychopathology, thereby eliminating many of the diagnostic limitations associated with temporal measurement of the fine distinctions between schizophrenia, schizoaffective disorder, and affective psychosis. Second, more closely examine correlations between neuropsychological performance and the presence of different observable symptoms (e.g., loose association, flat affect, alogia, avolition) as opposed to symptoms that must be inferred from patients’ verbal descriptions of ideational content (e.g., nature of delusions or extent of systematization).

In conclusion, replicating patterns of functional deficits in schizophrenia subgroups may lead to a better understanding of underlying psychophysiological processes. However, to date, the literature does not offer strong support for the notion of paranoid versus nonparanoid neuropsychological differences as measured by the most common clinically employed measures. Although the literature suggests a limited trend for the paranoid subtype to exhibit higher functioning in some cognitive domains, definitive conclusions regarding the interaction between schizophrenia subtype and neuropsychological functioning are premature.

References


Carter, C.; Robertson, L.; Nordahl, T.; O’Shora-Celaya, L.; and Chaderjian, M. Abnormal processing of irrelevant


Kolb, B., and Whishaw, I.Q. Performance of schizophrenic patients on tests sensitive to left or right frontal,


Osterrieth, P.A. Le test de copie dune figure complexe. *Archives de Psychologie*, 32:672–674, 1944.


### Acknowledgments

This project was supported by a grant from the Chapman Research Fund to Dr. Christine Zalewski and a grant from the Institute of Living, Hartford, CT, to Dr. James C. Seltzer.

### The Authors

Christine Zalewski, Ph.D., is Associate Professor, Pacific Graduate School of Psychology; Margaret T. Johnson-Selfridge, M.A., is a Doctoral Student, Pacific Graduate School of Psychology, Palo Alto, CA; Steven Ohriner, M.A., is in Private Practice, Redwood City, CA; Karen Zarrella, M.A., is in Private Practice, Winter Park, FL; and James C. Seltzer, Ph.D., is Program Director, Schizophrenia Rehabilitation Center, Institute of Living, Hartford Hospital’s Mental Health Network, Hartford, CT.