Stability and Course of Neuropsychological Deficits in Schizophrenia

Robert K. Heaton, PhD; Julie Akiko Gladsjo, PhD; Barton W. Palmer, PhD; Julia Kuck, PhD; Thomas D. Marcotte, PhD; Dilip V. Jeste, MD

Background: Neuropsychological deficits in schizophrenia appear to predate clinical symptoms of the disease and become more pronounced at illness onset, but controversy exists about whether and when further neuropsychological progression may occur.

Objective: To identify and characterize any subset of patients who evidenced progressive neuropsychological impairment, we compared the longitudinal stability of neuropsychological functioning in schizophrenic outpatients and normal comparison subjects.

Methods: One hundred forty-two schizophrenic outpatients and 206 normal comparison subjects were given annually scheduled comprehensive neuropsychological evaluations during an average of 3 years (range, 6 months to 10 years). Clinically and demographically defined subgroups were compared, and test-retest norms were used to identify individual patients who showed unusual worsening over time.

Results: The schizophrenic group was neuropsychologically more impaired than the normal comparison subjects but showed comparable test-retest reliability and comparable neuropsychological stability over both short (mean, 1.6 years) and long (mean, 5 years) follow-up periods. No significant differences in neuropsychological change were found between schizophrenic subgroups defined by current age, age at onset of illness, baseline level of neuropsychological impairment, improvement or worsening of clinical symptoms, and occurrence of incident tardive dyskinesia. Norms for change also failed to show neuropsychological progression in individuals with schizophrenia.

Conclusions: Neuropsychological impairment in ambulatory persons with schizophrenia appears to remain stable, regardless of baseline characteristics and changes in clinical state. Our results may not be generalizable to the minority of institutionalized poor-outcome patients.

Arch Gen Psychiatry. 2001;58:24-32

There is a high prevalence of neuropsychological impairment in persons with schizophrenia, ranging from mild deficits to frank dementia. Consistent with a neurodevelopmental view, some such deficits appear to predate clinical symptoms and exacerbate with typical illness onset during late adolescence or early adulthood.

There remains considerable controversy about whether there is further progression of neuropsychological deficits after the onset of the illness. With a few notable exceptions, cross-sectional studies generally have not found evidence of increased neuropsychological deficits in association with duration of illness or (relative to age-matched controls) in older than in younger patients with schizophrenia. Definitive answers to questions regarding possible progression of neuropsychological deficits in schizophrenia must come from longitudinal studies. The available studies, however, provide conflicting results and have a variety of methodologic limitations (eg, small or nonrepresentative samples, no controls, brief follow-ups, and/or limited neuropsychological testing).

Both neurodevelopmental and neurodegenerative views of neuropsychological deficits in schizophrenia seemingly remain viable. While progression of deficit after illness onset clearly is not universal, or even typical of the disorder, a subset of persons with schizophrenia may evidence cognitive deterioration over time. The size and nature of that subgroup are unclear, but the existing studies suggest that potentially relevant issues include subject age, duration of illness, level of initial neuropsychological impairment, improvement or deterioration.
SUBJECTS AND METHODS

SUBJECTS

Subjects included 142 outpatients with schizophrenia and 206 NCs, who had completed at least 2 comprehensive neuropsychological evaluations. Each patient and NC was a participant in 1 of 3 university-based clinical research centers, and most had contributed baseline neuropsychological data to previously published reports.37-40 Diagnostic procedures for subjects in both groups included the Structured Clinical Interview for the DSM-III-R or DSM-IV.41,42 Subjects were also screened by a physician using a physical examination and a structured medical history questionnaire; those with current or past medical conditions likely to affect central nervous system functioning (such as significant head injuries, seizure disorder, or acute medical conditions), as well as those meeting DSM-III-R53 or DSM-IV54 criteria for current alcohol or substance abuse or dependence, were excluded. All subjects provided written informed consent before participation in the research.

NEUropsychological EXaminations

The patients and NCs were assessed with an annually scheduled comprehensive neuropsychological test battery (in actuality, the mean test-retest interval between the first 2 of these evaluations was 16.6 months [SD, 8.6 months; range, 6-81 months]). Eighty-nine schizophrenic patients and 119 NCs completed at least 1 additional neuropsychological follow-up evaluation. The total number of evaluations ranged from 2 to 10 (mean, 3.38; SD, 1.57; median, 3.00). The interval between the first and last neuropsychological evaluations ranged from 6 to 125 months and was not significantly different for the schizophrenic patients vs NCs (mean and SD, 37.0 and 22.3 months vs 37.1 and 25.7 months, respectively; t_{\text{df}}=0.05, P = .96).

Except where otherwise indicated, each of the neuropsychological measures was part of the expanded Halstead-Reitan battery, and details of administration and scoring were as described by Reitan and Wolfson45 and Heaton and colleagues.46-47 The individual measures were grouped into 7 ability areas based on the neuropsychological constructs that they are putatively designed to measure. The verbal ability area included the Aphasia Screening Test, Boston Naming Test, and the Controlled Oral Word Association Test.48-49 Scores on the latter test were unavailable for a few subjects, so the Thurstone (written) Word Fluency task was substituted for 2 subjects at the baseline evaluation and 5 subjects at the first follow-up. The psychomotor ability area included the block design, object assembly, and digit symbol subtests from the Wechsler Adult Intelligence Scale–Revised (WAIS-R)50 and the Trail Making Test Part A (total time), Tactual Performance Test (total time), and Digit Vigilance Test (time). The abstraction and cognitive flexibility ability area included the Category Test (errors), Trail Making Test Part B (total time), and Wisconsin Card Sorting Test51 (perseverative responses). The attention ability area included the digit span and arithmetic subtests from the WAIS-R,50 the Rhythm Test, Speech Sounds Perception Test, and Digit Vigilance (errors). The learning and delayed recall ability areas included those scores from the Figure and Story Memory Tests, and the California Verbal Learning Test.52 The motor skills ability area included the right- and left-hand scores on Finger Tapping, Grooved Pegboard, and Hand Dynamometer.

We also had data from baseline and at least 1 retest evaluation with an even more comprehensive neuropsychological evaluation on a subsample of patients and NCs, permitting calculation of the WAIS-R Verbal, Performance, and Full-Scale IQs55 for 81 patients and 86 NCs, and the Halstead-Reitan battery Average Impairment Rating56 for 74 patients and 82 NCs.

Neuropsychological test raw scores were converted to standardized scaled scores (mean and SD, 10 and 3, respectively, in normal subjects), sex-corrected T-scores (mean and SD, 50 and 10, respectively, in normal subjects), by means of previously published normative data.46-48,51,54 We then calculated the mean scaled scores and T-scores within each ability area, and a composite global scaled score and T-score.

PSYCHIATRIC AND MOTOR SYMPTOM RATINGS

Longitudinal ratings of psychiatric symptoms were also obtained for a subset (n = 116) of the schizophrenic patients, as well as a smaller subset of NCs (n = 86 to 88). Psychiatric symptom rating scales included the Brief Psychiatric Rating Scale57 and the Scales for the Assessment of Positive and Negative Symptoms (SAPS and SANS, respectively).58 Tardive dyskinesia was assessed with the Abnormal Involuntary Movement Scale.57

STATISTICAL METHODS AND ANALYSES

The test-retest reliability of neuropsychological measures and symptom rating scales was assessed in terms of

Continued on next page

RESULTS

BASELINE CHARACTERISTICS

As shown in Table 1, there were small but statistically significant group differences in age, education, and test-
scores were different from those without only in age
3 research centers from which the present sample was
differences in neuropsychological protocols among the
different demographic recruiting emphases and slight
Impairment Rating scores, which primarily reflected
with and without WAIS-R IQ scores and Average
sex or ethnicity. As expected, the patients had substan-
ters were compared by means of Fisher
r-to-z transformations for intraclass correlation coeffi-
patients and NCs were compared by means of Fisher
were used for these analyses, since the lat-
neuropsychological T-scores, rather than T-scores, were used for these analyses, since the lat-
groups included 22 elderly patients (age at study entry,
65 years; mean, 69.2 years; SD, 3.3 years) vs 120
younger patients (age at study entry, <65 years; mean,
37 years; SD, 13.7 years); 24 patients with late-onset
patients and NCs, and then again in terms of changes from
patients in terms of sex and initial level of neuropsycho-
ate least 3 points. Separate analyses were conducted for
the SAPS- and SANS-defined change groups.
We also examined significant decreases in global neu-
psychological T-scores of individual subjects, by means
of the reliable change index method with adjustments for
practice effect.59,60 This approach involves constructing
prediction intervals around each subject’s expected
follow-up score. The predicted follow-up score is the sub-
ject’s baseline score, adjusted for practice effects among
cognitively stable individuals (as determined from the
mean change observed among the NCs). The boundary
values of the prediction interval around the predicted
follow-up score are determined by the normal variability
of baseline to follow-up changes (SE of the difference)
determined from the NC group. Specifically, 90% predic-
tion intervals were built around these predicted follow-up
scores by multiplying the SD of the difference among NCs
by 1.64. Subjects whose observed follow-up scores were
below the lower limits of the 90% prediction interval (ie,
the bottom 5% of normal controls) were considered to
have shown significant declines in neuropsychological
functioning. These procedures were conducted twice: once
to evaluate changes from baseline to first retest among all
patients and NCs, and then again in terms of changes from
the first to last neuropsychological assessment among sub-
jects with long follow-ups.
Two-tailed tests were used for all analyses. To help
avoid type I errors associated with multiple comparisons,
significance was defined as P<.01.

McGraw and Wong’s58 intraclass correlation coefficients
for degree of consistency between measurements at the
baseline and 1-year follow-up evaluations. The magni-
tudes of the intraclass correlation coefficients for the neu-
ropsychological and symptom rating scale scores of
patients and NCs were compared by means of Fisher

To examine the effects of length of follow-up on neu-
ropsychological stability, patients and NCs were divided
into diagnostic—by—length of follow-up groups, where
“short” follow-up was defined as less than 36 months and
“long” follow-up was defined as 36 months or longer. This
yielded 121 NCs and 75 schizophrenic patients with short
follow-up (mean, 19.5 months; SD, 6.8 months) and 85 NCs
and 67 schizophrenic patients with long follow-up (mean,
59.7 months; SD, 19.6 months).

In addition to the primary subject groupings
below the lower limits of the 90% prediction interval (ie,
the bottom 5% of normal controls) were considered to
have shown significant declines in neuropsychological
functioning. These procedures were conducted twice: once
to evaluate changes from baseline to first retest among all
patients and NCs, and then again in terms of changes from
the first to last neuropsychological assessment among sub-
jects with long follow-ups.

Two-tailed tests were used for all analyses. To help
avoid type I errors associated with multiple comparisons,
significance was defined as P<.01.

McGraw and Wong’s58 intraclass correlation coefficients
for degree of consistency between measurements at the
baseline and 1-year follow-up evaluations. The magni-
tudes of the intraclass correlation coefficients for the neu-
ropsychological and symptom rating scale scores of
patients and NCs were compared by means of Fisher

To examine the effects of length of follow-up on neu-
ropsychological stability, patients and NCs were divided
into diagnostic—by—length of follow-up groups, where
“short” follow-up was defined as less than 36 months and
“long” follow-up was defined as 36 months or longer. This
yielded 121 NCs and 75 schizophrenic patients with short
follow-up (mean, 19.5 months; SD, 6.8 months) and 85 NCs
and 67 schizophrenic patients with long follow-up (mean,
59.7 months; SD, 19.6 months).

In addition to the primary subject groupings
below the lower limits of the 90% prediction interval (ie,
the bottom 5% of normal controls) were considered to
have shown significant declines in neuropsychological
functioning. These procedures were conducted twice: once
to evaluate changes from baseline to first retest among all
patients and NCs, and then again in terms of changes from
the first to last neuropsychological assessment among sub-
jects with long follow-ups.

Two-tailed tests were used for all analyses. To help
avoid type I errors associated with multiple comparisons,
significance was defined as P<.01.

retest interval. There were no significant differences in
sex or ethnicity. As expected, the patients had substi-
tially greater global neuropsychological impairment and
more severe clinical symptoms than the NCs.

There were some differences between subjects
with and without WAIS-R IQ scores and Average
Impairment Rating scores, which primarily reflected
different demographic recruiting emphases and slight
differences in neuropsychological protocols among the
3 research centers from which the present sample was
drawn. Schizophrenic patients who had WAIS-R IQ
scores were different from those without only in age
(mean, 41.4 years [SD, 16.6 years] vs 55.9 years [SD,
9.5 years], respectively; t131 = 6.55; P<.001). Relative
to NCs without WAIS-R IQ scores, those with WAIS-R
IQ scores were younger (mean age, 45.4 years [SD,
16.8 years] vs 56.2 years [SD, 21.1 years], respectively;
t201 = 4.09; P<.001); completed slightly more educa-
tion (mean education, 14.8 years [SD, 2.4 years] vs
13.7 years [SD, 2.9 years], respectively; t200 = 3.12;
P<.002); had higher baseline neuropsychological
performance (mean global neuropsychological
T-score, 50.8 [SD, 4.0] vs 48.6 [SD, 4.9], respectively;
t204 = 3.46, P=.001); and were more likely to be male

(Reprinted) Arch Gen Psychiatry/Vol 58, Jan 2001
26

©2001 American Medical Association. All rights reserved.
Among the NCs, were not significantly different from those observed.

Table 1. Baseline Characteristics of Normal Comparison Subjects and Patients With Schizophrenia*

<table>
<thead>
<tr>
<th>Normal Comparison Subjects (n = 206)</th>
<th>Patients With Schizophrenia (n = 142)</th>
<th>t or χ²</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>51.7 (20.1)</td>
<td>47.6 (15.7)</td>
<td>2.12</td>
<td>340.63</td>
</tr>
<tr>
<td>Education, y</td>
<td>14.2 (2.8)</td>
<td>12.8 (2.3)</td>
<td>4.85</td>
<td>332.70</td>
</tr>
<tr>
<td>Sex, % M</td>
<td>64.1</td>
<td>69.7</td>
<td>1.20</td>
<td>1</td>
</tr>
<tr>
<td>Ethnicity, % white</td>
<td>81.6</td>
<td>79.6</td>
<td>0.21</td>
<td>1</td>
</tr>
<tr>
<td>Retest interval, mo</td>
<td>15.6 (5.3)</td>
<td>18.1 (11.7)</td>
<td>2.35</td>
<td>181.10</td>
</tr>
<tr>
<td>Total No. of neuropsychological evaluations</td>
<td>3.4 (1.7)</td>
<td>3.3 (1.4)</td>
<td>1.04</td>
<td>336.09</td>
</tr>
<tr>
<td>Global neuropsychological T-score</td>
<td>49.5 (4.7)</td>
<td>41.9 (6.5)</td>
<td>12.00</td>
<td>239.57</td>
</tr>
<tr>
<td>BPRS</td>
<td>22.2 (3.7)</td>
<td>33.8 (10.2)</td>
<td>12.13</td>
<td>180.75</td>
</tr>
<tr>
<td>SAPS</td>
<td>1.0 (1.4)</td>
<td>5.8 (3.8)</td>
<td>13.45</td>
<td>181.56</td>
</tr>
<tr>
<td>SANS</td>
<td>1.3 (1.7)</td>
<td>8.6 (5.1)</td>
<td>15.37</td>
<td>171.78</td>
</tr>
<tr>
<td>Age at onset of schizophrenia, y</td>
<td>NA</td>
<td>28.2 (14.0)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Duration of illness</td>
<td>NA</td>
<td>18.8 (14.3)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Paranoid subtype, %</td>
<td>NA</td>
<td>45.8</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Median CPZE, mg</td>
<td>NA</td>
<td>313 (n = 101)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Median BZE, mg</td>
<td>NA</td>
<td>4 (n = 49)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

*Values represent means (with SD) or proportions for all variables other than medication dosages; these are presented as medians for the subgroups taking the respective types of medication. Significance of differences was assessed with independent t tests for variables involving means and with Pearson χ² for those involving proportions. The mean daily dose of antipsychotic medication was 993.1 (SD, 1742.3; range, 29.0-12 250.0) mg CPZE. The mean daily anticholinergic dose was 9.4 (SD, 28.1; range, 0.5-200.0) mg BZE. Twenty-five patients were taking an atypical antipsychotic medication at some point in their participation in this longitudinal study (11 risperidone, 9 clozapine, 2 olanzapine, and 3 ramoxipride). Repeated-measures analysis of variance showed no significant differences in the daily chlorpromazine equivalent (CPZE) and the daily benztropine equivalent (BZE).

To further examine the change in cognitive functioning over time, we compared longitudinal global neuropsychological performance changes in subgroups of schizophrenic patients defined on the bases of demographic and clinical variables (as listed and defined above). These ANOVAs disclosed no significant group effects. As shown in Table 4, with 1 exception there also were no significant group × time interactions, indicating that the characteristics defining the various groups were not related.

©2001 American Medical Association. All rights reserved.
to changes in neuropsychological performance over time. The significant interaction reflected the fact that the schizophrenic subgroup with short duration of illness (mean, 2 years at baseline) had a slightly larger neuropsychological improvement than the subgroup with long duration (mean, 24 years at baseline). The smaller “improvement” (practice effect) shown by the long-duration group was the same as that of the NCs (both means, 1.3 T-score points).

Using the SAPS to categorize clinical change status resulted in 11 patients being classified as worse, 27 as better, and 78 as stable. With the use of repeated-measures ANOVAs, no significant group × time interactions were found for the neuropsychological global score or for scores on any of the 7 ability areas (Table 4). Categorization of schizophrenic patients on the basis of SANS scores yielded 19 subjects identified as significantly worse, 80 subjects as stable, and 17 subjects as better. The results of repeated-measures ANOVAs for the neuropsychological global score were similar to those with the SAPS change groups, ie, no significant group × time interactions.

We also compared patients with initial high vs low global neuropsychological performance (global neuropsychological T-score $\geq 45$ vs $\leq 39$, respectively), and those with vs without tardive dyskinesia as defined by the Schooler and Kane criteria. Using the SAPS to categorize clinical change status resulted in 11 patients being classified as worse, 27 as better, and 78 as stable. With the use of repeated-measures ANOVAs, no significant group × time interactions were found for the neuropsychological global score or for scores on any of the 7 ability areas (Table 4). Categorization of schizophrenic patients on the basis of SANS scores yielded 19 subjects identified as significantly worse, 80 subjects as stable, and 17 subjects as better. The results of repeated-measures ANOVAs for the neuropsychological global score were similar to those with the SAPS change groups, ie, no significant group × time interactions.

Table 2. Test-Retest Reliability of Neuropsychological and Psychiatric Symptom Rating Scale Scores for Normal Comparison Subjects vs Patients With Schizophrenia*

<table>
<thead>
<tr>
<th></th>
<th>Normal Comparison Subjects, ICC (C,1)</th>
<th>Patients With Schizophrenia, ICC (C,1)</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global neuropsychological scaled score</td>
<td>0.94 (n = 206)</td>
<td>0.93 (n = 141)</td>
<td>0.52</td>
<td>.61</td>
</tr>
<tr>
<td>Average impairment rating raw score</td>
<td>0.88 (n = 82)</td>
<td>0.86 (n = 74)</td>
<td>0.42</td>
<td>.67</td>
</tr>
<tr>
<td>WAIS-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>0.86 (n = 86)</td>
<td>0.91 (n = 81)</td>
<td>1.31</td>
<td>.19</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>0.72 (n = 86)</td>
<td>0.80 (n = 81)</td>
<td>2.97</td>
<td>.003</td>
</tr>
<tr>
<td>Full-Scale IQ</td>
<td>0.85 (n = 86)</td>
<td>0.92 (n = 81)</td>
<td>2.19</td>
<td>.03</td>
</tr>
<tr>
<td>Specific neuropsychological ability area scaled scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>0.82 (n = 183)</td>
<td>0.86 (n = 139)</td>
<td>1.33</td>
<td>.18</td>
</tr>
<tr>
<td>Psychomotor</td>
<td>0.92 (n = 167)</td>
<td>0.88 (n = 139)</td>
<td>2.14</td>
<td>.03</td>
</tr>
<tr>
<td>Abstraction/cognitive-flexibility</td>
<td>0.86 (n = 201)</td>
<td>0.81 (n = 135)</td>
<td>1.22</td>
<td>.22</td>
</tr>
<tr>
<td>Attention</td>
<td>0.80 (n = 184)</td>
<td>0.81 (n = 133)</td>
<td>0.41</td>
<td>.68</td>
</tr>
<tr>
<td>Learning</td>
<td>0.83 (n = 204)</td>
<td>0.81 (n = 140)</td>
<td>0.64</td>
<td>.52</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>0.63 (n = 206)</td>
<td>0.72 (n = 140)</td>
<td>1.45</td>
<td>.15</td>
</tr>
<tr>
<td>Motor</td>
<td>0.80 (n = 199)</td>
<td>0.85 (n = 153)</td>
<td>1.57</td>
<td>.12</td>
</tr>
<tr>
<td>Psychiatric symptom scales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPRS</td>
<td>0.48 (n = 88)</td>
<td>0.45 (n = 116)</td>
<td>0.27</td>
<td>.79</td>
</tr>
<tr>
<td>SAPS</td>
<td>0.42 (n = 86)</td>
<td>0.57 (n = 116)</td>
<td>1.37</td>
<td>.17</td>
</tr>
<tr>
<td>SANS</td>
<td>0.52 (n = 87)</td>
<td>0.56 (n = 116)</td>
<td>0.41</td>
<td>.68</td>
</tr>
</tbody>
</table>

* Test-retest reliability values are intraclass correlation coefficients for degree of consistency between measurements (ICC [C,1]). The z and P values reflect the comparison of ICC magnitude among normal comparison subjects vs patients with schizophrenia. WAIS-R indicates Wechsler Adult Intelligence Scale–Revised; BPRS, Brief Psychiatric Rating Scale; and SAPS and SANS, Scales for the Assessment of Positive and Negative Symptoms, respectively.

Table 3. Change in Neuropsychological T-Scores (First Retest Minus Baseline T-Score) for Normal Comparison Subjects vs Patients With Schizophrenia*

<table>
<thead>
<tr>
<th></th>
<th>Normal Comparison Subjects</th>
<th>Patients With Schizophrenia</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary score changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global neuropsychological T-score</td>
<td>1.3 (2.4) (n = 206)</td>
<td>1.6 (3.2) (n = 142)</td>
<td>1.30</td>
<td>251</td>
<td>.20</td>
</tr>
<tr>
<td>Average impairment rating T-score</td>
<td>1.6 (9.2) (n = 82)</td>
<td>3.9 (7.9) (n = 74)</td>
<td>0.94</td>
<td>154</td>
<td>.35</td>
</tr>
<tr>
<td>WAIS-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal IQ T-score</td>
<td>0.9 (4.8) (n = 86)</td>
<td>1.3 (5.6) (n = 81)</td>
<td>0.44</td>
<td>165</td>
<td>.66</td>
</tr>
<tr>
<td>Performance IQ T-score</td>
<td>3.3 (6.9) (n = 86)</td>
<td>1.9 (5.7) (n = 81)</td>
<td>1.46</td>
<td>165</td>
<td>.15</td>
</tr>
<tr>
<td>Full-Scale IQ T-score</td>
<td>2.2 (5.1) (n = 86)</td>
<td>1.6 (4.8) (n = 81)</td>
<td>0.73</td>
<td>165</td>
<td>.47</td>
</tr>
<tr>
<td>Neuropsychological ability area T-score changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>1.5 (5.6) (n = 183)</td>
<td>0.9 (5.4) (n = 140)</td>
<td>0.99</td>
<td>321</td>
<td>.32</td>
</tr>
<tr>
<td>Psychomotor</td>
<td>1.2 (4.3) (n = 167)</td>
<td>1.6 (4.7) (n = 140)</td>
<td>0.87</td>
<td>305</td>
<td>.38</td>
</tr>
<tr>
<td>Abstraction/cognitive-flexibility</td>
<td>3.0 (7.3) (n = 201)</td>
<td>4.5 (7.3) (n = 136)</td>
<td>1.87</td>
<td>335</td>
<td>.06</td>
</tr>
<tr>
<td>Attention</td>
<td>0.7 (4.6) (n = 184)</td>
<td>1.1 (5.0) (n = 134)</td>
<td>0.59</td>
<td>316</td>
<td>.49</td>
</tr>
<tr>
<td>Learning</td>
<td>1.4 (5.9) (n = 205)</td>
<td>2.8 (6.9) (n = 141)</td>
<td>1.94</td>
<td>269</td>
<td>.03</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>1.5 (6.7) (n = 206)</td>
<td>1.7 (7.4) (n = 140)</td>
<td>0.23</td>
<td>344</td>
<td>.82</td>
</tr>
<tr>
<td>Motor</td>
<td>1.0 (5.0) (n = 199)</td>
<td>0.6 (5.7) (n = 135)</td>
<td>0.78</td>
<td>332</td>
<td>.44</td>
</tr>
</tbody>
</table>

* Values reflect the mean (and SD) of the difference scores between each subject’s score at the first follow-up assessment, and that at baseline. Therefore, positive values reflect improved test performance. WAIS-R indicates Wechsler Adult Intelligence Scale–Revised.
Again, the groups did not differ significantly with respect to the change in the global neuropsychological score or the 7 neuropsychological ability domain scores (Table 4).

**SUBJECTS WITH UNUSUAL DECREASES IN GLOBAL NEUROPSYCHOLOGICAL T-SCORES**

The modified Reliable Change Index method, described above, was used to identify and compare the percentages of individual NCs and schizophrenic patients who evidenced unusual neuropsychological worsening from baseline to first retest. With the use of the 90% prediction interval, 10 NCs were so identified (about 5% of the total sample); the latter proportion did not differ significantly from the 5.6% of patients with schizophrenia who evidenced unusual worsening ($\chi^2_{1, N=348} = 0.10; P = .75$).

The above procedures were also used to evaluate baseline to final follow-up scores of the 85 individual NCs and 67 patients with schizophrenia with follow-up intervals of 36 months or longer (mean, 59.7 months; SD, 68 months). Again, the proportion of patients below the 90% prediction interval was not significantly different from that of the NCs (7.5% vs 4.7%, respectively; $\chi^2_{1, N=152} = 0.51; P = .48$).

**COMMENT**

To our knowledge, this is the first longitudinal study to compare comprehensive neuropsychological test results of large samples of NCs and patients with schizophrenia during a multiyear follow-up period. After comparing the two groups with respect to test-retest reliabilities, we attempted to determine whether the schizophrenic patients, or even a subset of that group, evidenced progressive neuropsychological decline.

Consistent with previous reports, presence of psychosis did not appear to affect the reliability of neuropsychological test performance. Across the entire neuropsychological battery, reliability estimates were quite high, and were at least as high for the schizophrenic group as for the NCs. This was true even though clinical symptoms were relatively variable over time (Table 2), and even though the test-retest intervals were longer than those in most previous reliability studies with NCs. These stability results support the view that neuropsychological deficits in schizophrenia are stable traitlike dimensions of the disorder, rather than reflecting state-related features.

Acceptable test-retest reliability does not rule out the possibility that neuropsychological deficits may be progressive in at least some patients with schizophrenia. Our results, however, provided no evidence of a deteriorating neuropsychological course in the total schizophrenic group, or in subgroups defined on the basis of age, sex, early vs late or recent vs remote onset of illness, baseline level of neuropsychological impairment, or length of follow-up. In each analysis, the schizophrenic groups showed slight improvements that were comparable with those evidenced by the NCs, and these likely represented practice effects. Even rather extreme clinical changes did not appear to influence the subjects’ neuropsychological performance.

Longitudinal research in institutional settings has shown evidence of neuropsychological decline in some low-functioning, chronically hospitalized patients with schizophrenia. Even in those studies, however, such worsening was observed only in small subsets of the groups being followed up. In the absence of longitudinal data from a neurologically stable comparison group, it is unclear whether these observed neuropsychological changes represent true neuropsychological decline, as in a neurodegenerative disorder, or whether they represent the tail of the distribution of test-retest fluctuations in neurologically stable (albeit very low-functioning) persons. It might be argued that the types of change observed in these patients (eg, change in the rated level of dementia) are too gross to be considered “normal fluctuation” and are pathognomonic of a progressive disorder. During approximately a 4-year follow-up interval, however, Ivnik and colleagues found remarkable test-retest differences in neuropsychological performances of some elderly subjects who did not have any neuromedical disorder likely to affect cognition, eg, on neuropsychological factor scores with IQ-type scaling, test-retest differences of 20 or more points were not unusual. Therefore, the outer range of possible fluctuation in performance of neurologically stable persons cannot be assumed and should be established by means of an appropriate comparison group.

The mean global neuropsychological score for our patient group was 1.62 SDs below that of NCs (Table 1), consistent with the range of effect sizes in a recent metaanalytic review of studies comparing controls and schizophrenic patients. Although severely demented patients were not represented in our sample, our “low-functioning” subgroup ($n=45$; Table 4) was very significantly impaired (ie, more than 3 SDs below NCs on the global neuropsychological T-score).

Limitations of this study include absence of chronically institutionalized patients, and therefore its results may not be generalizable to that minority of patients (but...
most contemporary schizophrenic patients are not institutionalized.\textsuperscript{68-70} Also, there was limited representation of elderly subjects, and subjects who evidenced substantial clinical change or incident tardive dyskinesias during follow-up (Table 4). Although relatively few of our patients were tested initially very early in the course of their illness, 2 recent studies of first-break schizophrenia showed no significant cognitive decline during the first several years of illness.\textsuperscript{35,36} Moreover, despite the limited power associated with some of the subgroup analyses in the current study, the data summarized in Table 4 show that these various subject characteristics were associated with clinically trivial effect sizes for changes in cognitive performance.

Considered together, the results of available longitudinal studies strongly suggest that the large majority of people with schizophrenia do not experience progressive neuropsychological decline after the initial onset of their illness. We found no evidence of such decline, even in our ambulatory patients with long follow-up, who were observed for an average of 5 years. It remains possible that a subset of patients with very poor outcome (not represented in our study) do experience progressive neuropsychological decline. Future research with this population should attempt to document such decline by ruling out nonsignificant fluctuations in performance as evidence of progressive impairment. This might be done by using neuropsychological norms for change developed with a similarly impaired but neurologically stable comparison group, and determining whether any unusual neuropsychological worsening in the schizophrenic patients remains stable or progresses further during a subsequent follow-up period. If future research should conclusively establish progressive neuropsychological decline in a subset of schizophrenic patients, questions arise as to why only a minority are so affected. Is this another example of the heterogeneous manifestations of a common disease,\textsuperscript{71-74} or are other factors involved, such as treatment history\textsuperscript{75} or comorbid neuromedical conditions (eg, age-related neurodegenerative changes, possibly in combination with schizophrenia-related low “cognitive reserve”)?\textsuperscript{76} There is a need for larger, collaborative studies using the same methods with patients in different treatment settings.

Accepted for publication July 20, 2000.

This study was supported in part by grants MH43695, MH45131, MH49671, MH19934, MH01452, and MH45294 from the National Institute of Mental Health, Bethesda, Md; State of California (Sacramento) Department of Mental Health grant DMH 89-7000; and the Department of Veterans Affairs, Washington, DC.

Preliminary versions of this work were presented in part at the Mt Sinai Conference on the Role of Cognitive Dysfun-

We thank Lou Ann McAdams, PhD, for her comments on an earlier draft of this article.

Reprints: Robert K. Heaton, PhD, Department of Psychiatry, University of California, San Diego, 140 Arbor Dr, San Diego, CA 92103 (e-mail: rheaton@ucsd.edu).

REFERENCES

30. Spree O, Strauss E. A Compendium of Neuropsychological Tests: Administra-


Andreasen NC, Olsen S. Negative vs positive schizophrenia: definition and validation. Arch Gen Psychiatry. 1982;39:789-794.


