INTRODUCTION

Neuropsychological impairment in alcoholics is found across different cultures. Studies in the US (Parsons, 1977; Tarter, 1980; Beatty et al., 1996; Nixon and Bowlby, 1987), Canada (Wilkinson, 1987), England (Ackr, 1986), Australia (Cala et al., 1978), Sweden (Bergman, 1985), and Spain (Nicolas et al., 1993) have shown deficits in cognitive flexibility, problem solving, verbal and non-verbal abstraction, visuo-motor co-ordination, learning, conditioning, and memory. General intellectual functions seem to be impaired less frequently (Tarter, 1980; Parsons, 1987). Whereas many of the above-mentioned studies used similar test batteries derived from the Halstead–Reitan Battery (Halstead, 1947) and the Wechsler Memory Scale (WMS; Wechsler, 1945), we combined the Trail Making Test B (TMT B), subtests of the WMS and the Benton Visual Retention Test (BVRT; Benton, 1972) with standardized tests commonly used in German neuropsychological research, e.g. the Leistungs Prüf System (LPS; Sturm and Willmes, 1983) as a measure of primary functions of intelligence. In this way, the first objective of our study was to investigate cognitive deficits in a sample of German alcoholics using international as well as domestic neuropsychological assessment methods. The correlations between cognitive deficits and the duration of alcohol dependency or abstinence prior to admission were examined as well.

The second goal of our study was to determine the effect of controlled abstinence on the recovery of psychological functions. Some studies on the reversibility of cognitive deficits in abstinent alcoholics have used independent group designs [reviewed by Tarter and Edwards (1985) and Goldman (1995)]. These cross-sectional studies usually compare groups of patients with different lengths of abstinence, e.g. recently detoxified alcoholics with age-matched subjects recruited from Alcoholics Anonymous samples. They do not require retesting of the subjects. Long-term abstinent alcoholics were shown to have normal memory (Reed et al., 1992) and normal TMT performance, compared with controls (Templer, 1975). Nevertheless, the patients did not always attain the level
of healthy controls (Parsons and Leber, 1981; Brandt et al., 1983; Muuronen et al., 1989).

Longitudinal studies, most of them uncontrolled, have shown significant improvements over time in a variety of variables, such as perceptual-motor speed and visual search time (Schafer et al., 1991), visual retention, verbal memory, and motor functions (Grünberger, 1989), verbal thinking, numerical thinking, and attention (Steck et al., 1982), and visual perception of mechanical causality (Poinmann et al., 1990). Controlled test–retest studies of the same cohort of patients are difficult to conduct and more costly than cross-sectional studies (Fabian and Parsons, 1983; Yohman et al., 1985). However, it was decided to retest patients and matched controls in order to assess learning effects and to pursue the extent and rate of cognitive deficits and their recovery during abstinence.

METHODS

Setting and subjects

The study was conducted at the in-patient treatment unit for alcohol-dependent patients at the University of Tübingen, Germany. All patients took part in a 6-week residential programme. They were admitted and discharged at the same time (closed group approach). Treatment elements comprised group therapy, role playing, and relapse prevention; individual counselling was done with patients and relatives (Mann and Batra, 1993). Specific training for cognitive dysfunction was not included.

Forty-nine alcohol dependent male patients and 49 male controls, matched for age, education, and marital status, were recruited. All patients met the criteria for alcohol dependency as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) of the American Psychiatric Association (1987), and the Munich Alcoholism Test (MALT; Feuerlein et al., 1980). Exclusion criteria included: dependence on psychotropic substances other than alcohol; history of signs of Wernicke–Korsakoff’s syndrome; evidence of head or brain injury (with amnesia of more than 12 h), and a history of signs of epilepsy. Patients were assessed by a standardized clinical interview (Mann, 1992), whose reliability and validity have been established (Mann et al., 1995). They received a standardized psychiatric and neurological examination, and had to fulfil at least one criterion of social adjustment, e.g., possessing regular employment and/or living with a partner (Mann et al., 1998).

Male controls were recruited by local newspaper advertisements. They were paid for participation and had to be abstainers or to have consumed no more than 20 g alcohol/day. Of 63 respondents meeting inclusion criteria, four were not willing to participate in the study, four could not be matched, and six showed elevated levels of γ-glutamyl transpeptidase (GGT) or mean corpuscular volume (MCV).

The age of subjects ranged from 24 to 60 years with a mean (± SD) of 41.7 ± 9.0 and 41.8 ± 8.9 for patients and controls respectively. In 29 matched pairs, age in years was identical for each patient and control. Seventeen pairs differed by 1 year and three pairs by 2 years. Years of education were equivalent \( t = 0.46, 96 \text{ df}, P < 0.64 \). Thirty-eight patients and 38 controls had no more than 9 years of school education, i.e. completion of extended elementary school. Six patients and four controls had completed non-classical secondary school (usually 10 years), and five patients and seven controls had completed classical secondary school (13 years). The hierarchical organization of German schooling makes educational level a highly reliable indicator of (pre-morbid) formal intelligence. Thirty-three patients and 33 controls were either married or lived in a stable relationship.

Patients had been alcohol-dependent for an average of 11.4 ± 5.5 years with durations of between 2 and 26 years. The onset of alcohol dependency was defined by the first signs of physical dependency (Schuckit et al., 1993). The average daily alcohol consumption was 214 ± 87 g (about 18 standard drinks; 1 standard drink = 12 g of pure ethanol), ranging from 100 to 450 g per day. Seven patients showed withdrawal symptoms which had to be treated with oral chlormethiazole for 3–6 days. They were tested after their withdrawal symptoms had disappeared (according to a set of specified withdrawal symptoms, such as elevated blood pressure, tachycardia, tremor, sweating, etc.), and after medication had been discontinued. The time of abstinence prior to admission for treatment was an average of 17.8 ± 29.4 days. Seven patients had not been abstinent when admitted, 19 patients had been abstinent for 1 to 4 days, and 23 patients had been abstinent for 6 or more days. On admission, GGT
values averaged 105.7 IU with an SD 128.5 IU (range 12 to 569 IU). Upon discharge, the average GGT was 34.8 IU. Seventeen patients (34.7%) showed symptoms of peripheral neuropathy and 19 patients (38.8%) had signs of cerebellar ataxia, as assessed by a standardized neurological examination (Mann, 1992). Seventeen patients (35%) had a positive family history of alcoholism (at least one first-degree relative with a history of alcohol dependency), 37 patients (76%) had lost their driver’s licence because of drunk driving, and 20% of the patients were unemployed, as compared with 12% of the controls.

**Neuropsychological testing**

All patients were tested at the beginning (T1) and at the end (T2) of their 6-week in-patient treatment. Abstinence had to be established at least 5 days before testing. Intervals between T1 and T2 were the same for both samples.

A multiple choice vocabulary test, the Mehrfachwahl Wortschatz Test (MWT-B; Lehrl, 1977), was administered only at baseline. This is designed to measure verbal intelligence and is thought to be an indicator of pre-morbid intelligence which is only affected by severe cerebral diseases (Lehrl et al., 1991). All the other neuropsychological tests (see Table 1) were performed at T1 and T2. They were selected to cover a broad range of cognitive functions. Accordingly, the battery was related to different aspects of intelligence, to attention and concentration, as well as to visuo-perceptual co-ordination and psychomotor speed. Our battery was comprised of the TMT B (Halstead, 1947; Reitan and Davison, 1974), the BVRT (Benton, 1972), an adapted version of the logical memory subtest of the WMS (Wechsler, 1945), and a number of subtests from the LPS (Horn, 1962). The LPS is a standardized test battery commonly used in German neuropsychological research to measure general intelligence. The LPS has been based on, but is not identical to, Thurstone’s primary mental abilities (Thurstone, 1938; Thurstone and Thurstone, 1941). Split-half reliabilities for

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trail Making Test B</td>
<td>Perceptual-motor speed, cognitive flexibility (part of the Halstead–Reitan Battery, Halstead, 1947; Reitan and Davison, 1974)</td>
</tr>
<tr>
<td>Logical Memory Subtest</td>
<td>Logical verbal memory (adopted from subtest 4 of the Wechsler Memory Scale, Wechsler, 1945)</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>Verbal ‘short-term’ memory</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>Verbal ‘long-term’ memory</td>
</tr>
<tr>
<td>Benton Visual Retention Test</td>
<td>Visual recognition of patterns, visual memory (Benton, 1972)</td>
</tr>
<tr>
<td>Leistungs Prüf System</td>
<td>Primary functions of intelligence (Horn, 1962, adapted by Sturm and Willmes, 1983)</td>
</tr>
<tr>
<td>Subtest LPS-U1 + U2</td>
<td>Verbal knowledge</td>
</tr>
<tr>
<td>Subtest LPS-U4</td>
<td>Reasoning (non-verbal)</td>
</tr>
<tr>
<td>Subtest LPS-U5</td>
<td>Word fluency (verbal)</td>
</tr>
<tr>
<td>Subtest LPS-U9</td>
<td>Spatial imagination (non-verbal)</td>
</tr>
<tr>
<td>Subtest LPS-U10</td>
<td>Flexibility of closure (non-verbal)</td>
</tr>
<tr>
<td>Subtest LPS-U12</td>
<td>Verbal closure</td>
</tr>
<tr>
<td>Revisionstest</td>
<td>Sustained attention and concentration (Marschner, 1980)</td>
</tr>
<tr>
<td>Auditory Verbal Learning Test (AVLT)</td>
<td>Immediate verbal memory span (recall of 15 words, adapted from the AVLT by Rey, 1964)</td>
</tr>
<tr>
<td>Mehrfachwahl Wortschatz Test</td>
<td>Verbal IQ (pre-morbid intelligence, Lehrl, 1977)</td>
</tr>
</tbody>
</table>
subtests range from 0.90 to 0.99. Our selection of LPS subtests was based on the work of Sturm and Willmes (1983) who conducted a thorough psychometric evaluation of this instrument. Furthermore, our test battery included an adaptation of the Auditory Verbal Learning Test (AVLT) originally introduced by Rey (Rey, 1964; Spreen and Strauss, 1991), and finally the Revisionstest (REVT) (Marschner, 1980), a well-known standardized test for concentration and attention consisting of simple arithmetical operations.

The neuropsychological test battery was administered to all patients. There was only one case of partially missing data, because of an organizational regulation that testing should not interfere with the existing treatment regimen. Written informed consent was obtained from all subjects after the procedure had been fully explained.

Statistical procedure
Apart from simple descriptive statistics and Pearson correlations, we used multiple univariate repeated-measures analyses (ANOVAs) to assess the main effects of groups and of time and the statistical interaction of group membership and time. Significant effects of group × time and consonant changes in means will be interpreted as indicators of patients’ cognitive recovery beyond mere retest effects and learning. Multiple ANOVAs seemed to be appropriate and sufficient, because comparisons aim at each variable separately (cf. Huberty and Morris, 1989). Since a Bonferroni–Holm correction (Holm, 1979) seemed to be too conservative, we decided on single hypotheses as the conceptual units for error rate (Kirk, 1968). All data analyses were performed using the Statistical Analysis System (SAS). The GLM procedure was used for repeated-measures analyses of variance, including Tukey’s test for the comparison of means.

RESULTS

Group differences in neuropsychological variables
On the MWT-B, the multiple choice vocabulary test, which was administered only at T1, patients and controls had means (±SD) in the medium range (107.2 ± 12.8 and 116.2 ± 15.4, respectively). Contrary to our expectations, there was a significant difference between both samples (t = 3.17, 96 d.f., P < 0.01).

The ANOVA results showed significant group differences for five out of 12 parameters from the main battery (see Table 2). Except for the BVRT and delayed logical memory test, where patients and controls achieved almost equivalent results, patients consistently scored lower than controls. Significant impairment was confirmed for a number of LPS subtests, including verbal knowledge, abstract reasoning, and spatial imagination. Significant differences also resulted for perceptual motor speed (TMT-B) and auditory verbal learning.

At T1, all five significant group differences (patients vs controls) were confirmed by Tukey’s test for the comparison of means.

At T2, only the differences in immediate recall in auditory verbal learning remained significant (Tukey’s test). Although previously significant differences disappeared at T2, patients still had lower scores than controls on most of the tasks (see Table 2 for details).

ANOVA interaction effects (group × time) were significant for three out of 12 variables. These were the TMT-B (F = 4.43, 1/95 d.f., P < 0.04), the subtests verbal knowledge (F = 5.70, 1/94 d.f., P < 0.02), and spatial imagination (F = 4.71, 1/94 d.f., P < 0.04) of the LPS. Moreover, for the abstract reasoning subtest of the LPS, the interaction effect approached the 5% significance level (F = 3.48, 1/94 d.f., P < 0.07). All the interaction effects had the same pattern: patients showed distinct deficits at T1 compared with the controls, but improved considerably over time. Thus, in these domains, the performance of the alcoholics tended to approach that of the controls. Respective pre-post-test differences ranged from 0.5 to about 2.0 SD.

In addition, ANOVA main effects of time, which are thought to represent non-specific improvement (learning effects) for alcoholics and for controls alike, could be confirmed on the 5% level for all variables, except for auditory verbal learning and the number of correct responses in the BVRT.

Drinking and neuropsychological variables at T1 and T2
The number of days of abstinence prior to admission (17.8 ± 29.4 days) was not related to neuropsychological performance, including the MWT-B. Correlations between duration of abstinence and all neuropsychological tests were below 0.25 and thus failed to reach conventional significance levels.
Years of alcohol-dependence correlated significantly with poor performance in spatial imagination at T2. In accordance with our expectations, patients who had longer periods of dependence showed poorer results in this LPS subtest \( r = -0.32, P < 0.03 \). However, at T1, no significant relationship between these two variables could be established \( r = -0.05, P < 0.75 \). The remaining correlations, including the MWT-B as a measure for pre-morbid intelligence, were well below 0.20 and therefore of no statistical significance.

**DISCUSSION**

The absence of any significant correlation between duration of abstinence prior to admission and all neuropsychological tests suggests that minor symptoms which may persist after acute withdrawal had little, if any, impact on neuropsychological performance.

At the beginning of the rehabilitation treatment, alcoholics scored significantly lower than matched controls on five out of 12 neuropsychological parameters. Impairment was found for verbal as well as for non-verbal performance. In agreement with current research, these results provide further support for the ‘mild generalized dysfunction hypothesis’ (Tivis et al., 1995).

Contrary to our expectations, we found no differences between the two groups on the BVRT and logical memory test. On both of these tasks, which are strongly related to non-verbal and verbal memory functions respectively, patients and controls performed equally well.

Whereas logical memory and visual retention did not differ between the two samples, patients demonstrated an evident and persistent handicap on a test involving memory for verbal material, a procedure which is similar to Rey’s (1964) AVLT. This supports research that suggests that the AVLT may be more sensitive for detecting memory deficits in clinical populations and less amenable

---

**DISCUSSION**

The absence of any significant correlation between duration of abstinence prior to admission and all neuropsychological tests suggests that minor symptoms which may persist after acute withdrawal had little, if any, impact on neuropsychological performance.

At the beginning of the rehabilitation treatment, alcoholics scored significantly lower than matched controls on five out of 12 neuropsychological parameters. Impairment was found for verbal as well as for non-verbal performance. In agreement with current research, these results provide further support for the ‘mild generalized dysfunction hypothesis’ (Tivis et al., 1995).

Contrary to our expectations, we found no differences between the two groups on the BVRT and logical memory test. On both of these tasks, which are strongly related to non-verbal and verbal memory functions respectively, patients and controls performed equally well.

Whereas logical memory and visual retention did not differ between the two samples, patients demonstrated an evident and persistent handicap on a test involving memory for verbal material, a procedure which is similar to Rey’s (1964) AVLT. This supports research that suggests that the AVLT may be more sensitive for detecting memory deficits in clinical populations and less amenable
to compensatory strategies than prose learning (Peaker and Stewart, 1989).

With respect to initial testing, another point is worth making. Based on the hierarchical organization of the German educational system, with its homogeneity within the different levels of education, patients and controls were carefully matched, therefore a major difference in pre-morbid intelligence was not to be expected. However, the significant difference in the MWT-B casts some doubt on the validity of this test as a precise measure of pre-morbid intelligence in alcoholics. Although there are seemingly no other direct comparisons between alcoholics and matched controls for the MWT-B, corresponding evidence has been provided by Carroll et al. (1992). They found that the National Adult Reading Test (NART), which like the MWT-B, is considered to be a good indicator of pre-morbid intelligence, was invalid in a sample of alcoholic patients suffering from Korsakoff’s syndrome. The issue of pre-morbid intelligence continues to be a matter for debate. Whereas Parsons (1987) and Acker et al. (1984) suggest that such results reflect true differences in pre-morbid intelligence, Ar buckle et al. (1994) consider them to be alcohol-related effects. More conclusive evidence should be obtainable from prospective studies.

After several weeks of abstinence, differences between control subjects and patients were reduced to non-significant levels for all but one parameter (verbal short-term memory, AVLT). Significant interaction effects also confirmed that patients’ neuropsychological performance increased significantly more than that of controls for verbal and non-verbal tests, as well as for complex cognitive tasks such as the TMT. However, differences remained stable for immediate recall in auditory verbal learning.

In addition to significant interaction effects and group differences, non-specific learning or retest effects were observed for patients and controls alike. This underlines the importance of using control groups in assessing neuropsychological recovery in a repeated measurements study. In addition to general considerations, e.g. effect sizes and statistical power, it is evident that the confirmation of group differences and over-proportional improvement depends to a great extent on temporal conditions, especially on the gradient of recovery for the respective functions under investigation. The temporal aspect creates certain difficulties for the comparison of different studies. For example, Yohman et al. (1985), in their study of 37 male alcoholics, found clear time effects but no interaction effects for alcoholics and controls. Similar results in a study with 25 female alcoholics and matched controls were reported by Fabian and Parsons (1983).

Recovery is not only time-dependent, but also experience-dependent. This has been shown by repeated-measurement studies (Forsberg and Goldman, 1985; Goldman, 1987, 1990). Aside from abstinence, non-specific treatment effects might also have contributed to the recovery process, although no explicit cognitive training was part of the treatment programme. The recovery process might also be more pronounced for time measures, as compared to accuracy measures, and patients who have resumed drinking might score even better in some measures on retesting than abstainers (Glenn et al., 1994), although they have been observed to perform significantly poorer than abstainers and normal controls on an overall neuropsychological performance index (Parsons and Nixon, 1996).

No substantial correlations between duration and dependence and neuropsychological performance could be established. Therefore, linear relationships between duration of dependence and neuropsychological performance may be too simple a way to describe the dose–response relationships between alcohol and neuropsychological performance. In fact, most studies have found that an alcoholic’s drinking history does not relate significantly to the speed or extent of recovery (Goldman, 1995). In the future, specific drinking patterns, including repeated withdrawal, physical impairment induced by alcohol, and perhaps personality characteristics should be considered as moderators when assessing dose–response patterns.

In summary, our study, using a powerful controlled test–retest design including normal controls, provided further evidence that chronic alcoholism (intermediate stage) has distinct detrimental effects on a variety of cognitive functions. It also shows that recovery from cognitive dysfunction on a variety of tasks is relatively rapid when patients remain abstinent.

Acknowledgements — This work was supported by a grant from ‘Deutsche Forschungsgemeinschaft — DFG’ (SFB 307,B3). The authors thank C. Wegner, B. Overberg, S. Taubert, A. Schupmann, H. W. Schied, J. Jojart, and R. U’ren for their assistance with this study.
REFERENCES


