Abstract

Approaches to cognitive remediation have differed across studies. Most of the larger studies have concentrated on group treatments designed without the benefit of recent laboratory-based studies. The current study describes a randomized trial of an intensive cognitive remediation program involving individual daily sessions of 1 hour for up to 3 months. It targets executive functioning deficits (cognitive flexibility, working memory, and planning) that are known to be problematic in people with schizophrenia. Procedural learning, as well as the principles of errorless learning, targeted reinforcement, and massed practice, was the basis of the intervention. The program was compared with an alternative therapy (intensive occupational therapy) to control for some of the effects of therapeutic contact. Some improvements in cognition followed both therapies. A differential effect in favor of cognitive remediation therapy was found for tests in the cognitive flexibility and the memory subgroups. There was a trend for those receiving atypical antipsychotic medication to benefit more from cognitive remediation for tests of cognitive flexibility. Although there were no consistent changes in symptoms or social functioning between groups, if improvement in cognitive flexibility tasks reached a threshold then there is some evidence that social functioning improved, even over the short duration of the trial. In addition, cognitive remediation differentially improved self-esteem. This study supports the view that cognitive remediation can reduce cognitive deficits and that this reduction may affect social outcome, at least in the short term.

Key words: Cognition, remediation.


People with a diagnosis of schizophrenia complain of cognitive deficits. They describe problems with concentrating on simple tasks, both within and between episodes of illness, that may be so severe as to affect every aspect of life. Despite our patients' concern, most therapeutic attention has been placed on the reduction and recurrence of positive and negative symptoms, which until recently have been treated with medication only (Garety et al. 1994; Kuipers et al. 1997).

A wealth of evidence now points to the influence of these deficits on current and future social functioning (Wykes and Dunn 1992; Buchanan et al. 1994) and on treatment outcome (Mueser et al. 1991). They also have a role in vulnerability-stress models in the prediction of both the onset and recurrence of the disorder (Nuechterlein et al. 1994). Recently it has been suggested that these deficits are more directly related to functioning outcomes than symptoms are, which accounts for the paucity of predictive relationships in previous studies linking symptoms and future outcome (Green 1998). Reducing these deficits would provide obvious benefits to both the patients and the psychiatric services. Not only would quality of life improve, there would also be less dependence on psychiatric care and fewer hospital admissions. The overall social and financial costs of the disorder would therefore be decreased. Only during the past few years has cognitive remediation been a therapeutic focus. Given the articles set out in this issue of Schizophrenia Bulletin, this tardiness may be unforgivable.

Choosing the right target for cognitive remediation has proven problematic. We have lacked a clear theoretical framework to direct our efforts to particular deficits (Green 1993), and a general pessimism about the likelihood of changing such deficits has pervaded any attempts. On purely pragmatic grounds, choosing those cognitive disabilities that contribute to functional outcomes would at least affect independent living, quality of life, and cost of care. However, longitudinal studies have generally
tested few cognitive variables and replications of positive findings are few (Green 1996). Our approach is to combine these predictive data with information on the pattern of deficits found in people with schizophrenia.

The most commonly reported deficits are reductions in the executive functions, often in the context of average IQ (e.g. Levin et al. 1989; Crawford et al. 1993; Evans et al. 1997). **Executive functioning** is a term suggested by Alan Baddeley (1986) and Shallice (1988) to describe the way information is controlled and processed. These processes are essential in a number of different situations, for example, in planning and decision making, in error correction, and for novel responses. Patients with schizophrenia have shown reductions in performance on neuropsychological tests reputed to tap these abilities (*working memory*, Goldman-Rakic 1986; *cognitive flexibility*, Goldberg et al. 1990; *planning*, Morris et al. 1995). In a recent small study involving a group of remitted patients living in the community, 94 percent were found to have a deficit in one of the three components of executive functioning (Morice and Delahunty 1996). Because these deficits are so prevalent, and also because they have been linked to functioning outcomes (Jaeger and Douglas 1992; Green 1996), they are the focus of this clinical intervention.

In designing the remediation of these executive functions, we noted two sorts of studies of cognitive remediation. The first group changes performance on individual tasks to tease out specific training elements that are essential to improvement. Although the results are variable and may be related to chronicity of illness, severity of deficits, and sophistication of the training methods adopted, some optimistic data were produced. Errorless learning, immediate feedback, and reinforcement improved performance on a variety of tasks. For example, the Wisconsin Card Sorting Test (Summerfelt et al. 1991; Benedict et al. 1994; Stratta et al. 1994; Vollema et al. 1995; Bellack et al. 1996) and a recent study of massed practice (Wexler et al. 1997) showed improvements on both perceptual and fine motor tasks. But there has been little evidence that improvements were generalized to other neuropsychological tasks, let alone to general problem solving.

In contrast, the key idea for the second group of studies is that changes in cognitive performance will lead directly to improvements in social functioning. Although some of the same training elements are involved, the emphasis is on the practice and rehearsal of specific cognitive abilities. One of the first of these more comprehensive programs is called Integrated Psychological Therapy (IPT; Brenner et al. 1994) and involves a number of different subprograms, only one of which focuses on cognitive abilities. All are presented in group format. The data from controlled trials of this program are variable. Most show improvements in cognitive functioning, but there is little support for subsequent effects on social skills (Brenner et al. 1994). Some evidence does suggest a subsequent effect on readmissions. However, it is not possible to conclude that any improvements are specific to the cognitive subprogram rather than to the remaining psychosocial subprograms. Spaulding and colleagues (1998) in Nebraska have addressed this problem, and their results suggest a specific effect of the cognitive subprogram on improvements in social skills; improvements in cognitive functioning seem to result from nonspecific effects of the therapy. Further results are also presented in this issue of the Bulletin.

Another dimension to cognitive remediation is the adoption of a group or individual approach. The pattern of disabilities in schizophrenia, although showing some general trends (Shallice et al. 1991), can also show some individual variation, particularly in executive functions (Evans et al. 1997). The variable response of patients to the IPT group treatment has sometimes been attributed to these individual differences, in response to different stages of the subprogram. It may be difficult to provide an optimal learning environment when patients differ in their abilities and when each may have different practice requirements. For these reasons, and because of the difficulty of incorporating the proven training elements into group treatments, our research team adopted an individual approach. This approach facilitated the use of errorless learning, immediate feedback, and nondidactic training, which have been successful in laboratory studies.

The program adopted by this research group was developed in Australia by Ann Delahunty and Rod Morice (1993). It has a strong theoretical base in the deficit profile of executive processing in schizophrenia. The program consists of three modules: cognitive flexibility, working memory, and planning. Each module involves a series of tasks, rated from extremely easy to easy, effectively providing an errorless learning environment. In the cognitive flexibility module, patients are given practice in engagement, disengagement, and re-engagement activities for a particular cognitive set or between two sets. For instance, they are given a page with a set of numbers on it and are asked to cross off the odd or even numbers. This activity requires them to maintain a set but also to shift set when requested. The working memory module requires the person to maintain two sets of information simultaneously and to carry out transformations on a held information set. For example, the subject recalls the numbers of tokens in series of lines but also transforms this information by recalling the lines in different orders. This task requires encoding as well as transformations of that
encoding. The emphasis is on categorizing and chunking information as well as on some self-instructional training in the use of mnemonic strategies. In the planning module, the participant has to plan a sequence of moves to achieve a goal. The emphasis is to organize information and to create and use subgoals. Preliminary evidence of efficacy has been encouraging. Delahunty et al. (1993) found improvement on the Wisconsin Card Sorting Test and Delahunty and Morice (personal communication) showed significant differential gains on tests for cognitive flexibility and memory following training. The latter study did not investigate social functioning, so generalization could not be measured. Rossell and David (1997) have suggested that “modifications which promote planning, pacing, and error detection” aid performance on the Wisconsin Card Sorting Test. These interventions are exactly the types promoted in our training program, although we hope to be able to change performance on more than one type of task with little overlap between the training program and the outcome neuropsychological tests.

All neurocognitive training programs are highly intensive and require much contact with a therapist, so effectiveness of therapy is confounded with therapist involvement. This trial was designed to test the gains in neuropsychological test performance and social functioning specific to neurocognitive remediation rather than therapist contact. The gains were therefore compared with those produced by an alternative therapy with good face validity. The control therapy chosen is intensive occupational therapy, which is similar to the occupational therapy activities offered in most psychiatric settings but more structured. Because of the heterogeneity of the deficits among patients with schizophrenia, participants were chosen for entry only if they showed deficits on both social functioning and cognitive tests. The durability of training 6 months after treatment and changes in regional brain activity associated with any cognitive changes (using functional magnetic resonance imaging) have also been investigated. These latter two effects will be described in future reports.

Design

Patients who agreed to take part in the trial were randomly allocated to one of two treatment groups (neurocognitive remediation or intensive occupational therapy). The therapies were provided in addition to and independently of day program activities in which the participants were enrolled. Patients were assessed before and after treatment and 6 months after completing therapy on key outcome variables chosen a priori from a variety of social, symptomatic, and cognitive tests.

All ratings of symptoms were carried out by a single rater who was blind to group assignment. Although the cognitive test data were not collected by blind assessors, many of the scales were computer driven and scored. Key workers who provided the social functioning data were not blind to group assignment, but anecdotal evidence suggests that each worker believed both treatment groups were therapeutically active.

Subjects

All participants were recruited from community psychiatric clinics that serve a large geographical area in South London and fulfilled the following criteria:

1. Satisfied DSM-IV (American Psychiatric Association 1994) criteria for schizophrenia (from chart reviews by T.W.)
2. In touch with services for at least 2 years
3. Showed evidence of cognitive difficulties (scored below the 16th percentile on the Wisconsin Card Sorting Test) and social functioning problems (scored on at least one problem on the Social Behavior Scale)
4. Presented no evidence of organic brain disease and no primary diagnosis of substance abuse
5. Had no plans to change medication during the treatment phase

Recruitment was systematic; each community team was asked to identify all likely participants who might fulfill the criteria for entry and these potential recruits were then screened on the above criteria. Our data suggest that this sample was similar to an epidemiologically defined sample drawn from the same population, although the current sample scored higher on both symptoms and social functioning problems, which might be expected given the criteria for entry (comparison data drawn from Wykes et al. 1998). The majority of eligible participants were treated as outpatients, but a few were in residential rehabilitation programs in the psychiatric hospital.

Basic Demographic Information

Basic information such as age, sex, and marital status as well as information on current and past functioning, such as peak job performance, were collected from case notes and the key workers. Clinical information on service use and medication (in chlorpromazine equivalents) was also collected.
Cognitive Measures

The cognitive measures are categorized as cognitive flexibility, planning, or working memory. This broad categorization was informed by the authors' description of the tasks and a systematic review of the literature. Each task may contain elements of other cognitive processes. In addition, the cognitive flexibility tests were identified in a cognitive analysis by Eslinger and Grattan (1993). General ability as measured by the vocabulary subtest from the Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler 1981) and a premorbid measure of IQ using the National Adult Reading Test (NART; Nelson and Willison 1991) were also collected.

Cognitive Flexibility

Hayling Sentence Completion Task (Burgess and Shallice 1996). In the first section of the task, participants complete a sentence with a word so that the sentence makes sense (compatible condition). In the second part, the participant provides a word so that the resulting sentence does not make sense (incompatible condition). The number of errors produced in the incompatible condition was the key outcome variable (scoring manual provided by Burgess and Shallice 1996).

Trails (Reitan 1958). This is a computerized version of the letters and numbers task (trails A and B) in the Halstead-Reitan battery. The key variable was the time taken to carry out a letters and numbers task minus the time required to carry out a numbers task.

Response inhibition (Wykes et al. 1990). This task measures the ability of the participant to inhibit a previously learned response and to learn a novel one. In the initial trials, the participant has to press the button adjacent to the light, which is an easily learned and “natural” response (compatible condition). In the later trials, the participant must press a different button that is not adjacent to the light (incompatible condition). The required flexibility is to inhibit the previously learned response and generate the new one. Previous research has shown that this task is an indicator of future functioning and may also be a vulnerability indicator for schizophrenia (Wykes et al. 1990; Wykes 1994). The key variable was the reaction time in the incompatible condition, in which there are four lights and four buttons, less the movement time, that is, the simple reaction time.

Controlled Oral Word Fluency Test (Spreen and Benton 1977). The FAS version of this task was used. Subjects had to provide as many different words as possible beginning with a particular letter within 1 minute. The key measure was the age-scaled score.

Stroop Neuropsychological Screening Test (Trenerry et al. 1989). In this paper-and-pencil version of the ubiquitous Stroop test, the participant has to name the ink color of printed words that describe a different color. The key outcome variable was the number of correct ink color names within 2 minutes in the color word task.

Wisconsin Card Sorting Test (WCST; Milner 1963). The standard form of the test, which has been investigated in many studies of schizophrenia, was used. The key variable was the age- and education-scaled percentile of categories achieved, scored as a categorical variable where 1 = 1st percentile and 5 = 16th percentile or greater, following the manual (Heaton et al. 1993). This variable was chosen because it is the final outcome of a number of different patterns of error performance, such as high levels of perseverative errors and/or nonperseverative errors, and failure to maintain set.

Planning

Tower of London (Morris et al. 1995). This computerized version of the Tower of London task was adapted from Shallice (1982). Subjects move discs on a touch-sensitive screen from a starting position to match a pattern shown on the top of the screen. The key measure was the number of moves required to complete the task minus the minimum number of moves required for this level. Results from previous studies have shown that people with schizophrenia perform poorly on this task, even after their overall slowness to respond has been controlled for (Morris et al. 1995).

Six elements. In this modification of the Shallice and Burgess (1991) test, participants perform three tasks, each of which has two sections (A and B). They are given 10 minutes in which to complete as many tasks as possible, but they are not allowed to do section B immediately after section A of the same task. The key indicator of improvement was the number of tasks completed less the number of rule breaks.

Memory and Working Memory

Visual span. Participants were required to reproduce from memory increasingly complex figures presented on a grid. The key measure was the highest level at which two out of four figures were correctly recalled.

Sentence span. This test was based on the Daneman and Carpenter (1980) sentence span task. Groups of sentences are read to the participant, who must recall the last word in each sentence after hearing the whole group. The number of sentences within a group increases throughout the test, and the span was the number of sentences in the group where the last word was recalled correctly.

Digit span. This was a subtest from the WAIS–R. The key variable was the age-scaled score.
Dual span (Della Sala et al. 1995). This task measures working memory as defined by Baddeley (1986), whose concept of working memory is the processing capacity required to carry out two tasks at the same time. This dual span task involves a visual task (tracking) and a verbal task (remembering numbers). The participant carries out each task separately so capacity on each task can be assessed, then carries out both tasks together at the level just below his or her capacity level on each task individually. The measure of working memory was defined by a formula produced by Baddeley and Della Sala (1996, p. 1399), which combines performance on the memory and tracking tasks. Not all participants completed the final version of this task (n = 25).

General Functioning Measures

To control for the effects of general functioning and premorbid ability, both the National Adult Reading Test (NART), which provides a measure of premorbid IQ, and the vocabulary subtest of the WAIS–R, which provides a measure of current functioning, were also collected.

Symptoms and Social Functioning

Social Behaviour Schedule (SBS; Wykes and Sturt 1986). This schedule was rated from information given by a key informant who had known the participant for at least 1 month. Detailed and anchored ratings of 21 items of behavior are scored on a scale of 0 to 4. The key outcome variable was the total score; high scores indicate problematic behavior.

Present State Examination (PSE; Wing et al. 1974). This scale was rated from an interview with the patient. The key outcome variable was the total PSE score.

Brief Psychiatric Rating Scale (BPRS; Ventura et al. 1993). Each symptom on the 24-item scale was rated on a scale of 1 to 7. The key score was the total for the complete scale.

Rosenberg Self Esteem Schedule (Rosenberg 1965). This scale measures both self-esteem and self-deprecation. The key measure was the total score, which reflects overall esteem.

Therapies

The therapies were matched for therapist contact and length of treatment. The regimen was 1-hour daily sessions over 40 days. Our aim was to have patients attend at least 3 days per week and preferably 5 days per week. The pace was set by the client but remained within the confines of the program structure.

Intensive Occupational Therapy. The intensive occupational therapy program consisted of activities often offered by occupational therapy services within psychiatric settings, although we excluded those that require extensive preparation, such as painting. Our elements included relaxation, assertiveness training, life diary, comprehension of social information, and role playing. We tried to develop a program that had little overlap in form with the experimental therapy, although in practice this was extremely difficult. For example, the life diary requires the participant to remember past information and order it, which often requires overt cueing by the therapist. The program was presented in a manual, and participants were expected to try each of the activities in a set timetable.

Neurocognitive Remediation. The procedure for cognitive remediation was set out in the manual (Delahunty and Morice 1993). In each session, a variety of tasks were presented to practice each of the component processes in complex planning or problem solving. For instance, each session of the cognitive shift module consisted of oculo-motor, perceptual, fine motor, and conceptual tasks. The tasks were graded in difficulty, with easier ones being presented early in the program. Tasks involving fine motor material (e.g., finger tapping and finger sequencing) were always presented last because their ease ensured that the session always ended on a positive note. At first, information-processing strategies or means to organize behavior were incorporated into the task, because the participants were unable to undertake these independently (e.g., the therapist may regulate performance speed by pointing to each item at a fixed pace). The therapist also discussed information-processing strategies and the regulation, organization, and monitoring of behavior for each task in order to produce few errors. If necessary, the therapist demonstrated these methods of dealing with information. The three steps of this process were: (1) the therapist demonstrated the information processing overtly, (2) the patient used such methods overtly, and (3) the patient used the methods covertly. The amount of covert use of the processing strategies increased as the participant moved through the program.

Statistical Analysis

The analyses were carried out on an intention-to-treat basis (i.e., all possible patients who began therapy, irre-
spective of completion of therapy, are included in each statistical test). The statistical analyses were divided into two sets. In the first set, analyses of covariance were utilized to assess differential treatment on each individual outcome measure. These analyses were carried out using BMDP-V5, which employs a multivariate regression/likelihood analysis approach that is valid when data are missing at random (versus missing "completely at random" as required by analysis of variance and multivariate analysis of variance approaches; see Everitt 1998 for details).

The second, more exploratory set of analyses investigated whether a priori levels of improvement in the cognitive measures were related to other changes, particularly social functioning. First, an index was calculated for each test, which was the change in test score divided by the standard error of the whole sample at baseline. This variable was similar to a measure of the effect size and allowed improvements across different tests for each individual to be equated within a category. These proxy effect size measures were further transformed to define a threshold of improvement for each of the three main categories of neuropsychological tests (cognitive flexibility, memory, and planning) for each individual. This threshold was calibrated as "improvement" if 50 percent or more of the tests in that category showed performance increases (defined as an increase of at least one standard error of the whole sample's baseline scores for that test). If less than 50 percent of the tests improved, the index was calibrated as "little or no improvement."

Results

Descriptive Information. A total of 33 people were recruited to the trial, 16 in the control therapy intensive occupational therapy group (IOT) and 17 in the neurocognitive remediation group (NCR). Four people did not complete the whole course of treatment (three in the NCR group and one in the control therapy). Some data are missing from each group at each time point as patients refused to take part in a particular test. All possible data were entered into the analyses.

The sociodemographic information and clinical history variables of the two treatment groups are shown in table 1. The groups were representative of the types of patients with chronic schizophrenia who attend the inner-city services. The length of contact was defined as a minimum of 2 years, but the groups had been in contact for considerably longer. Most had been able to live away from their parents and psychiatric services at some time.

The symptom and social functioning levels at entry to the trial are shown in table 2.

The groups were clearly well matched. All patients achieved a clinical diagnosis of schizophrenia, and many showed continuing high levels of positive symptoms as measured by the PSE, which allowed a current symptomatic diagnosis of schizophrenia for 82 percent of the patients in the NCR group and 88 percent for the IOT group. In addition to their functioning and cognitive difficulties, many also continued to experience hallucinations.
Table 2. Social and symptomatic variables

<table>
<thead>
<tr>
<th>Factor</th>
<th>Neurocognitive remediation group (n = 17)</th>
<th>Intensive occupational therapy group (n = 16)</th>
<th>Statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social behavior problems</td>
<td>12.9 (range = 1–24)</td>
<td>11.1 (range = 1–31)</td>
<td>NS</td>
</tr>
<tr>
<td>PSE total</td>
<td>23.6 (4–74)</td>
<td>26.8 (3–52)</td>
<td>NS</td>
</tr>
<tr>
<td>BPRS total</td>
<td>43 (25–66)</td>
<td>45.6 (26–60)</td>
<td>NS</td>
</tr>
<tr>
<td>Rosenberg Self Esteem</td>
<td>32.4 (18–50)</td>
<td>29.3 (12–47)</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Medication**

- **Neuroleptic** (typicals in chlorpromazine daily equivalents):
  - Typical antipsychotics (n = 8) 807 mg (SD = 760)
  - Clozapine (n = 6) 426 mg (range 300–700)
  - Risperidone (n = 2) 6–8 mg
  - Olanzapine (n = 1) 10 mg

- **Anticholinergic** (procyclidine): n = 7
  - Average dose 12 mg (SD = 4.8)

- **Typical antipsychotics** (n = 9) 867 mg (SD = 751)
- Clozapine (n = 5) 450 mg (range 300–700)
- Risperidone (n = 2) 6–12 mg

**Statistical significance**: NS

**Note.** — NS = not significant; PSE = Present State Examination; BPRS = Brief Psychiatric Rating Scale; SD = standard deviation.


and to report passivity phenomena and delusions. There were no differences in medication dosage, range of antipsychotic medications, or levels of anticholinergic medication prescribed for the two treatment groups.

There were no differences between treatment groups on any of the key measures on the cognitive tests at entry to the trial. The groups were selected because they had cognitive deficit and performance on the majority of tests fell well below normal levels, yet a range of scores was still discernible on each task. Premorbid IQ as estimated by the NART was 104 (range = 86–123) and the current level of functioning on the vocabulary subscale of the WAIS–R was 8.4 (range = 3–16), indicating a reduction from premorbid levels. Neither of these two variables differed between the groups.

**Dropouts During Treatment.** Only four people dropped out of treatment. They had more symptoms but fewer social functioning problems than those who completed the trial and generally performed worse on all tests but significantly so only on verbal fluency (mean score: dropouts = 17, completers = 33; t = 4.9, df = 7.4, p < 0.001). The analysis technique, which uses all data combined with the relatively small numbers of dropouts, suggests that the reported results should not be affected greatly by this problem (Everitt 1998).

**Medication.** Data for medication are given in table 2; there were no differences between the groups in the type or dosage of the medication prescribed. Patients who were taking clozapine had been on the drug for at least 6 months (usually a year) prior to the trial. Over the course of the trial, only two patients, one from each group, changed type of medication (one from typical to clozapine and one from risperidone to clozapine). In addition, 12 patients changed the dosage of their medication (NCR: three decreases, four increases in dose; IOT: three decreases, two increases in dose). There were no significant differences between therapy groups in the amount of change in drug dose within categories of drugs (i.e., typical: t = 1.27, p = 0.22; atypical: t = 0.87, p = 0.4). Nor was there evidence of significant changes in anticholinergic medication. Medication changes did not reflect changes in symptom scores, which were equivalent between drug groups (PSE: t = 0.48, p = 0.6; BPRS: t = 0.20, p = 0.8). There were no significant correlations between changes in test performance and medication dose changes or between dose changes and social functioning, self-esteem, or symptoms.

**Improvements With Both Therapies.** These analyses investigate whether there are nonspecific aspects of therapy that can improve cognitive and social functioning over time. Many studies have repeatedly shown that performance on executive functioning tasks is poor in patients with schizophrenia and remains stable over time in the absence of any intervention (e.g., Shallice et al. 1991). However, some evidence of improvement in test performance over time was found in both therapy groups as assessed by paired t tests (for means and confidence intervals, see table 3). The improvements were most noticeable in the cognitive flexibility group of tests:
Table 3. Mean changes over time in cognitive and functioning variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline</th>
<th>Posttreatment</th>
<th>Confidence interval of the mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>33.2 (12.4)</td>
<td>37.5 (12.9)</td>
<td>1.07, 7.55†</td>
</tr>
<tr>
<td>Hayling</td>
<td>11.1 (6.65)</td>
<td>5.87 (5.51)</td>
<td>2.2, 8.3†</td>
</tr>
<tr>
<td>Trails</td>
<td>29.0 (21.2)</td>
<td>20.5 (14.1)</td>
<td>2.6, 14.5†</td>
</tr>
<tr>
<td>WCST</td>
<td>2.67 (1.32)</td>
<td>2.53 (1.41)</td>
<td>−0.32, 0.59</td>
</tr>
<tr>
<td>Response inhibition</td>
<td>599 (588)</td>
<td>464 (303)</td>
<td>−69, 339</td>
</tr>
<tr>
<td>Stroop</td>
<td>72.5 (23.0)</td>
<td>81.5 (20.5)</td>
<td>3.3, 14.6†</td>
</tr>
<tr>
<td><strong>Planning tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower of London</td>
<td>7.7 (5.4)</td>
<td>7.4 (5.42)</td>
<td>−1.9, 2.5</td>
</tr>
<tr>
<td>Modified Six Elements</td>
<td>3.37 (1.4)</td>
<td>4.3 (2.3)</td>
<td>0.03, 1.84†</td>
</tr>
<tr>
<td><strong>Memory tests</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit span</td>
<td>8.17 (2.9)</td>
<td>8.7 (3.1)</td>
<td>−0.2, 1.27</td>
</tr>
<tr>
<td>Sentence span</td>
<td>1.96 (0.98)</td>
<td>2.26 (1.05)</td>
<td>−0.09, 0.7</td>
</tr>
<tr>
<td>Visual span</td>
<td>5.13 (1.73)</td>
<td>5.96 (1.73)</td>
<td>0.27, 1.4†</td>
</tr>
<tr>
<td>Dual span</td>
<td>86.5 (13.9)</td>
<td>87.26 (17.2)</td>
<td>−8.4, 9.8</td>
</tr>
<tr>
<td><strong>Functioning</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-esteem</td>
<td>31.1 (7.7)</td>
<td>32.6 (7.8)</td>
<td>−0.65, 3.47</td>
</tr>
<tr>
<td>Social behavior</td>
<td>12.7 (7.8)</td>
<td>12.5 (8.6)</td>
<td>−3.35, 3.7</td>
</tr>
<tr>
<td>PSE total</td>
<td>23.7 (14.4)</td>
<td>18.7 (13.6)</td>
<td>2.1, 7.95†</td>
</tr>
<tr>
<td>BPRS total</td>
<td>43.8 (10.2)</td>
<td>41.1 (11.2)</td>
<td>−0.59, 6.06</td>
</tr>
</tbody>
</table>

Note.—SD = standard deviation; WCST = Wisconsin Card Sorting Test; PSE = Present State Examination; BPRS = Brief Psychiatric Rating Scale.

†Statistically significant effect.

Trails, fluency, Stroop, and Hayling Sentence Completion Task. For memory tests, the only improvement was in the visual span task; for the planning set of tests, there was an improvement only in the Six Elements task.

For the symptom and functioning measures, the only general improvement was in the symptom score on the PSE, although the BPRS scores showed a trend for improvement. The improvement in symptoms was not accounted for by changes in any particular factor scores; rather the pattern of change was different for each individual.

Changes in test performance could be due to changes in symptoms rather than the interventions. However, there were few significant correlations between changes in test performance and changes in symptoms. The improvement on the PSE was positively correlated with WCST improvement but was negatively related to improvements in the Six Elements tests (PSE: \( r = 0.34, p = 0.035 \); WCST: \( r = 0.38, p = 0.023 \)). The BPRS was related only to changes in fluency (\( r = 0.36, p = 0.028 \)). Exploratory stepwise multiple regression analyses of the symptom scores indicated no significant predictors of outcome levels other than initial scores.

Specific Effects of Therapy on Outcome Measures. These measures are divided into cognitive functioning, symptoms, and social functioning changes. Analysis of covariance of the posttreatment scores, using the corresponding baseline measure as a covariate, was used to calculate confidence intervals for the treatment difference for each variable. This type of analysis is more powerful than alternatives (e.g., analyzing change scores; see Senn 1997). Further analyses of these individual outcome variables, including the medication factor (typical vs. atypical neuroleptics), found no drug differences and no drug-by-treatment interaction for any individual outcome variable.

Cognitive performance. Means, standard deviations, and sample sizes as well as the resulting confidence intervals for each variable are shown in table 4. Intervals excluding the value 0 imply a significant treatment effect: the WCST, Six Elements, and digit span. The advantage in each case is in favor of neurocognitive remediation. Other variables with confidence intervals that did include 0 appear to considerably favor neurocognitive remediation, such as response inhibition, Stroop, and working memory in the dual span task.

The analyses reported in table 4 do not take into account multiplicity (i.e., multiple testing; see Senn 1997). This problem can be addressed in several ways, the most usual being the Bonferroni correction. This correction is, however, known to be extremely conservative so there is a strong possibility that interesting effects could
Table 4. Changes in cognitive performance between the two therapy groups using analysis of covariance with baseline as the covariate

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Baseline</th>
<th>Posttreatment</th>
<th>Confidence Intervals for adjusted treatment differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>n</td>
</tr>
<tr>
<td>Cognitive flexibility tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>NCR</td>
<td>30.3 (13.1)</td>
<td>35.4 (13.2)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>33.1 (12.8)</td>
<td>38.7 (13)</td>
<td>15</td>
</tr>
<tr>
<td>Hayling</td>
<td>NCR</td>
<td>14.1 (8.9)</td>
<td>6.3 (4.8)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>9.2 (9.1)</td>
<td>5.4 (6.3)</td>
<td>15</td>
</tr>
<tr>
<td>Trails</td>
<td>NCR</td>
<td>30.7 (24.6)</td>
<td>20.9 (14.4)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>27.6 (17)</td>
<td>20.1 (14.4)</td>
<td>15</td>
</tr>
<tr>
<td>WCST</td>
<td>NCR</td>
<td>2.6 (1.1)</td>
<td>2.9 (1.2)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>2.7 (1.5)</td>
<td>2.2 (1.5)</td>
<td>15</td>
</tr>
<tr>
<td>Response inhibition</td>
<td>NCR</td>
<td>668 (801)</td>
<td>555 (375)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>470 (226)</td>
<td>380 (175)</td>
<td>15</td>
</tr>
<tr>
<td>Stroop</td>
<td>NCR</td>
<td>68.2 (23.5)</td>
<td>83.5 (18.2)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>72.7 (24.0)</td>
<td>78.7 (22.6)</td>
<td>15</td>
</tr>
<tr>
<td>Planning tests (NS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower of London</td>
<td>NCR</td>
<td>8.3 (5.7)</td>
<td>8.2 (4.7)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>7.0 (6.2)</td>
<td>7.0 (6.2)</td>
<td>15</td>
</tr>
<tr>
<td>Modified Six Elements</td>
<td>NCR</td>
<td>3.0 (3.6)</td>
<td>4.8 (1.3)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>3.3 (1.5)</td>
<td>3.8 (1.5)</td>
<td>15</td>
</tr>
<tr>
<td>Memory tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit span</td>
<td>NCR</td>
<td>7.5 (2.6)</td>
<td>8.9 (3.6)</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>8.5 (3.2)</td>
<td>8.5 (2.6)</td>
<td>15</td>
</tr>
<tr>
<td>Sentence span</td>
<td>NCR</td>
<td>2.0 (0.7)</td>
<td>2.3 (1.0)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>1.7 (1.2)</td>
<td>2.2 (1.1)</td>
<td>12</td>
</tr>
<tr>
<td>Visual span</td>
<td>NCR</td>
<td>4.9 (1.5)</td>
<td>5.9 (1.5)</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>5.2 (1.9)</td>
<td>6.0 (2.0)</td>
<td>12</td>
</tr>
<tr>
<td>Dual span</td>
<td>NCR</td>
<td>86.9 (13.4)</td>
<td>91.8 (15.3)</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>85.2 (13.5)</td>
<td>81.9 (17.5)</td>
<td>13</td>
</tr>
</tbody>
</table>

<sup>Note.—SD = standard deviation; NCR = neurocognitive remediation; IOT = intensive occupational therapy; WCST = Wisconsin Card Sorting Test; NS = not significant.</sup>

<sup>1</sup>Statistically significant treatment effect.

be missed, particularly as this was primarily a hypothesis-generating study.

**Functioning and symptom measures.** Table 5 shows the means, standard deviations, and sample sizes for the four functioning measures. Of the confidence intervals, only self-esteem showed a significant direct treatment effect, with the advantage being to the neurocognitive remediation group.

**Clinical Improvement.**

*Are there generalized effects within cognitive domains?* All improvements in test performance using the proxy effect size measure are shown in figures 1 through 3 for individual tests in each of the three separate domains (cognitive flexibility, planning, and memory, respectively). Analyses of the threshold index showed that more people in the neurocognitive remediation group achieved the status of "improved" (more than half the tests showed improvement) on the cognitive flexibility tests (63% vs. 24%, \(X^2 = 5.43, df = 1, p < 0.02\)) and the memory tests (53% vs. 13%, \(X^2 = 5.4, df = 1, p < 0.02\)) but not on the planning tests. For two of the three subgroups of tests, these analyses indicate some generalization across the domain in favor of the neurocognitive remediation group. However, there was little concordance between the three indexes (percentage agreement: cognitive flexibility/planning = 53%; memory/planning = 27%; memory/cognitive flexibility = 52%), suggesting that individuals differ on the gains made in treatment.

The level of cognitive improvement seems to be related to the type of drug prescribed: people who take atypical medications show more consistent improvements.
in performance on the cognitive flexibility tests, but this difference is not significant (proportions achieving more than 50% cognitive flexibility change; typical = 38%, atypical = 53%, $\chi^2 = 0.78$, ns). However, there seemed to be an interaction between therapy and type of drug prescribed. There was little difference between the proportions of people who improved on the typical antipsychotic medication (43% cognitive remediation vs. 33% control therapy), but in the atypical antipsychotic medication condition the figures were 78% versus 13%. This interaction was tested by comparing two logistic regression analyses. In the first analysis, level of improvement (< or > 50%) was the dependent variable and both drug type and therapy were entered as predictor variables. In the second

**Figure 1.** Changes in tests of cognitive flexibility over the trial

![Chart showing changes in cognitive flexibility tests over the trial](image)

**Table 5.** Changes in functioning measures in the two therapy groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Baseline</th>
<th>Posttreatment</th>
<th>Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>NCR</td>
<td>32.4 (7.5)</td>
<td>17</td>
<td>35.9 (5.8)</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>29.3 (8.5)</td>
<td>16</td>
<td>29.6 (8.4)</td>
</tr>
<tr>
<td>PSE total score</td>
<td>NCR</td>
<td>23.6 (16.1)</td>
<td>17</td>
<td>14.2 (10.5)</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>26.8 (17)</td>
<td>16</td>
<td>23.1 (15.2)</td>
</tr>
<tr>
<td>BPRS total score</td>
<td>NCR</td>
<td>43.2 (11.7)</td>
<td>17</td>
<td>39.8 (12.1)</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>46.5 (10)</td>
<td>16</td>
<td>42.3 (10.6)</td>
</tr>
<tr>
<td>Social behavior total score</td>
<td>NCR</td>
<td>12.9 (7.7)</td>
<td>17</td>
<td>12.4 (8.2)</td>
</tr>
<tr>
<td></td>
<td>IOT</td>
<td>11.1 (8.1)</td>
<td>16</td>
<td>12.9 (9.5)</td>
</tr>
</tbody>
</table>

Note.— SD = standard deviation; NCR = neurocognitive remediation; IOT = intensive occupational therapy; PSE = Present State Examination; BPRS = Brief Psychiatric Rating Scale.

^1Statistically significant effect.
analysis, an additional variable, the interaction between type of drug and therapy group, was also entered. Both models were significant, but the difference between two $\chi^2$ statistics was 3.767, which just fails to reach significance (5% level, 3.84). There is therefore a suggestion that the interaction between drugs and level of improvement is important in attaining consistent cognitive improvement.

In order to investigate what predicts improvement, each of the threshold indexes was entered as a dependent variable in a logistic regression using a forward stepwise method. The explanatory variables considered were initial performance on the tests making up each index, age, gender, initial functioning and symptom measures, therapy group, and initial drug type. Therapy group was a significant predictor for both the cognitive flexibility and memory groups of tests but not the planning group of tests (for cognitive flexibility, therapy group 95% confidence interval (CI) for odds ratio = 1.22, 29.47; for memory, therapy group 95% CI for odds ratio = 0.022, 0.810). No other variables were significant predictors.

Do changes in cognition affect functioning? Although there was no generalized effect of neurocognitive remediation on social functioning, changes in cognitive performance might have to reach a threshold before change in social behavior is detectable. Figure 4 shows the social behavior problem scores at baseline and posttreatment (3 months later) when cognitive performance falls above or below the threshold level for each of the three domains of tests. The pattern of performance is similar in each group (i.e., patients who improve their test performance above the criterion show decreases in social behavior problems). A stepwise regression analysis was used to explore putative predictors of posttreatment social behavior as the dependent variable. Independent variables included baseline social behavior, which was forced into the analysis at step one, as well as improvement on cognitive tests, initial levels of cognitive performance, symptoms, age, and gender. The final model included baseline social behavior but also improvement on cognitive flexibility tests and initial levels of verbal fluency as predictors ($F = 5.4, p < 0.006$). This model explained 34 percent of
Figure 3. Changes in memory tests over the trial

Figure 4. Social functioning change related to domain of cognitive improvement

Note.—A decrease in score denotes improvement.
the variance, which was 29 percent more than baseline social behavior measures alone.

Discussion

The participants, who were randomly allocated to the two treatment groups, did not differ initially on any of the outcome measures and were well matched on all sociodemographic and clinical variables. The study was designed to test whether there were any specific effects of cognitive remediation over and above the nonspecific effect of intensive contact with a therapist. Unlike many other studies in which the effects of therapy are compared with a control group who received standard care, in this study the key analyses relate to the differential effects of two treatments. Several different domains were examined to test the generality of improvements. These included neuropsychological tests, social functioning, symptoms, and self-esteem.

In summary, there was some evidence for significant differential improvements in individual test performance in favor of the neurocognitive remediation group, even when using an intention-to-treat protocol and covariance analyses. This form of analysis is rarely used in longitudinal assessments of psychological interventions, particularly cognitive remediation. All the mean changes were in favor of the active therapy condition. In addition, more patients receiving neurocognitive remediation achieved a threshold of improvement (at least half the tests improved) in both the memory and cognitive flexibility domains. These differential improvements in performance are unlikely to be a direct result of practice because the cognitive remediation program contained many subtasks that were dissimilar to the content of the neuropsychological tests. Participants seem to be learning to engage information-processing strategies, which can be used in a variety of tasks.

Patients with schizophrenia want to improve their quality of life and their functioning. No changes in symptoms or social functioning were directly attributable to the therapy; however, self-esteem improved differentially, with little change in the intensive occupational therapy group but improvements in the remediation group. This may be because the participants, who by definition were high users of services, were able to perceive their increased performance each day with the program’s errorless learning procedures. For many of these patients, this may have been the first time they experienced such success in any rehabilitation program. This effect is important, as it may contribute to future changes, such as adherence to treatment and acceptance of further rehabilitation efforts. The data from the followup will allow us to test this hypothesis.

There were no significant correlations between any of the functioning measures (apart from between the two symptom measures), so it seems unlikely that symptoms were an explanatory variable in these results. Cognitive performance change was related to changes in social functioning, but only if the change in performance reached a threshold. Most, but not all, of the people who achieved this threshold were in the remediation group. Two-thirds scored more than 50 percent improvement; in the control therapy, only one-quarter of the group attained the same level.

Practice effect presents another possible explanation for changes in cognition. But several studies have suggested that there are no practice effects on many of the tests employed. For instance, Nopoulous et al. (1994) showed stability of performance early in the course of the disorder on many tests, including memory, fluency, and some attention tasks over a 1- to 2-year period, even when the initial tests were carried out at index hospitalization. Other research groups (Sweeney et al. 1991; Hoff et al. 1992; Rund et al. 1997) have also found this result. Although there may be an effect of practice on Stroop and Trails, other tests improved as well, including those that were stable in other studies (e.g., fluency and memory tests).

Significant cognitive changes could also be attributable to therapeutic contact. Both treatments were designed to be active (i.e., they had face validity to help patients improve functioning). The intensive occupational therapy offered the opportunity to practice both attention and concentration skills on ecologically valid tasks, as well as to learn some new skills and to develop a role. These opportunities occurred in intensive, one-to-one social contact where conversational skills might develop because of the practice of massed sessions within each week. Both therapies also provided opportunities for positive feedback. These nonspecific effects clearly have an effect on cognitive performance, although not on self-esteem. They might also have affected symptom measures: Two of our patients commented that during treatment they had so much to think about that their symptoms, particularly hallucinations, were not so apparent. This anecdote was supported by our data (mean BPRS hallucinations: baseline = 3.5, posttreatment = 3.0, t = 1.822, p = 0.04 one-tailed). This trend implies that benefits of intensive individual occupational programs may have an immediate effect on some symptoms.

It is always difficult to disentangle the effects of medication and those symptomatic changes. However, there was no evidence of any relationship between dose changes and changes in symptom measures. One of our analyses did suggest that atypical neuroleptics, mainly clozapine, were related to improvements in cognitive per-
performance, particularly cognitive flexibility, but only in the neurocognitive remediation group. This interaction may occur for a number of different reasons. Patients prescribed clozapine were not randomly assigned to that drug. Although the medication groups did not differ initially on any cognitive measures, people prescribed atypical medication are likely to be more cooperative and trustworthy than people prescribed depot medication. They are therefore likely to attend any treatment and complete homework assignments, which may account for differential effects. Alternatively, the trend may be due to the effect of atypical medication on a component of test performance that was also improved by remediation, possibly the practice derived on timed components. This conjecture would fit with other data (Buchanan et al. 1994) and the speculation of others (Goldberg and Weinberger 1994).

Although neurocognitive therapy produced no direct effects on social functioning, evidence suggests that if cognitive performance is improved enough, social functioning is affected. If it is the amount of cognitive change that is relevant to generalization to social functioning, then the aspects of neurocognitive remediation that will be predictive of the most change must be delineated. In this study, improvement did not seem to depend on characteristics present in the sample prior to the trial. The main predictor was therapy group, not IQ according to NART, current medication, vocabulary level, or the initial values of tests, even though the control therapy had good face validity and, in other studies, had been associated with improvements in cognitive performance (Brown et al. 1993).

This study was a strict test of a remediation program because the tests were dissimilar to the content of the program and the effects could not be the result of practice. The patients were both cognitively and socially handicapped and were suffering from a chronic form of the disorder. Despite a relatively small sample size, there were differences in outcome between therapy groups. So what was the key to this success? Perhaps it lies in the sophistication of the program itself, which not only targets specific aspects of executive functioning but also depends on procedural and errorless learning. The individual sessions themselves also contain many different types of tasks, so generalization naturally occurs within the program itself.

The threshold effects of changes in cognitive performance support the putative link between cognitive and social functioning, indicated not only by empirical data (some of which is presented in this issue of the Bulletin) but also in theories suggested by Brenner et al. (1994) and more recently by Green (1998). Both these theories suggest a way in which these links might be made. Structural equation modeling to estimate treatment differences on the assumed cognitive flexibility, planning, and memory latent variables would allow more thorough explanations of the data. But these sorts of analyses require considerably larger sample sizes to be viable.

Our positive results must be built upon by exploring the predictors of change in more detail. The largest improvements were made by patients who were taking atypical medication and who were given a complete neurocognitive remediation program. These improvements seem to have an effect on social functioning. To the extent that these improvements may offer protection from information processing overload and perhaps increased positive symptoms, they could also reduce the need for future services (Wykes 1994).

The program itself might be further refined following case study investigations of rates of change during therapy. In addition, the followup data currently being collected will indicate whether the effects of the therapy are durable. It is quite possible that none of the effects found at posttreatment will continue when they are not supported by a program. This possibility does not negate the positive finding of this and other remediation studies; rather, it poses another question: How do we support the improvements after the intensive period? Continuing, but less intensive, support for cognitive functioning may be required for many years. The costs of such treatment could be offset by improvements in functioning and reduced dependence on more expensive services such as inpatient treatment. After all, no one is suggesting that other treatment programs, either pharmacological or occupational, should produce gains after they have been withdrawn, so why should this be a criterion for psychological interventions?

References


Morris, R.G.; Rushe, T.; Woodruffe, P.W.R.; and Murray, R.M. Problem solving in schizophrenia: A specific deficit


**Acknowledgment**

The study was supported by the Stanley Foundation and the Medical Research Council (UK).

**The Authors**

Til Wykes, M.Phil., Ph.D., is a Reader in Clinical Psychology at the Institute of Psychiatry in London and an Honorary Consultant Clinical Psychologist at the Maudsley Hospital. Clare Reeder, B.A., Julia Corner, B.Sc., and Clare Williams, B.A., are trainee clinical psychologists. Brian Everitt, B.Sc., is Head of Biostatistics and Computing at the Institute of Psychiatry, London, UK.