A Token Economy Project with Chronic Schizophrenic Patients

By ROGER BAKER, JOHN N. HALL and KEITH HUTCHINSON

The token economy is an application of operant conditioning theory which is particularly relevant to the treatment of patients in institutions. The basic idea is to motivate the patient to behave more appropriately and constructively by giving him tokens whenever he behaves in such a way. The tokens have an incentive function, as the patient can cash them in for a number of different things, such as pleasant single room accommodation, cigarettes, magazines, TV viewing, drinks etc., depending upon what is programmed into the token economy. The theory of operant conditioning is that not only does the token motivate the patient to behave in certain ways, but Reinforces his behaviour. Such reinforced behaviour should eventually become an integral part of the 'way he is', continuing even when the token is removed.

Since the classic study of Ayllon and Azrin (1968) the number of token economy wards has multiplied considerably. Although a large number of these wards have been in the U.S.A., a survey carried out recently by the present authors showed that there are now about 30 ward operant programmes in hospitals in Britain (Hall, 1973). Many of the earlier studies were concerned with chronic psychiatric patients, and it is with these patients that the majority of token economies is still concerned. Indeed, it has been suggested that token economies are now the treatment of choice for such patients (McReynolds and Coleman, 1972).

What is the potential of the token economy approach and why does it appear to be of increasing popularity?

One important feature is that the key people in the therapy are those who have most contact with the patient—in hospital usually the nurses. Operant conditioning provides a simple, practical approach in which nurses can do something for patients. It is a structured approach with a unified theory behind it, namely that any patient can learn new modes of behaviour if the appropriate reinforcer can be found. Token economies specify the direction of learning and desirable behaviour and make sure that the patient is always reinforced for each instance of the behaviour. The nurse becomes a teacher and no longer a custodian, and there becomes hope for all patients instead of despair.

However, other therapies have been developed which are practical and offer hope for these patients. Russell Barton's Institutional Neurosis (1966) clearly lists what steps can be taken to help chronic patients, which if properly instituted could do a lot to improve long-term care in Britain. However there seems little evidence that it has been systematically applied. Why is this? While it lists some important techniques, the book possibly lacks the benefit of a unified theory. It tells in general what can be done, but offers little specific guidance on individual problems. The appeal of operant conditioning is that it offers an approach to each and every behaviour problem for all types of patient, and it is possibly this strength that lies behind its growing acceptance.

Many studies have described the use of tokens with schizophrenic patients in American hospitals. For instance, Schaefer and Martin (1966) dealt with the problem of apathy in a group of 20 chronic female schizophrenic patients. These patients were given tokens for personal hygiene, social interaction and adequate work performance, which were all thought to be incompatible with apathy. After three months the patients had become more outgoing than a comparison group of patients who had received custodial care.

* An article by Baker, Rosenthal, Ineson and Andrews (1972), describes this approach for nurses, and several 'cookbooks' are available, notably Schaefer and Martin's Behavioral Therapy (1969).
A TOKEN ECONOMY PROJECT WITH CHRONIC SCHIZOPHRENIC PATIENTS

care. Two reviews have been published in this area (Carlson, Herson and Eisler, 1972; and Davison, 1969). Some studies have been concerned with the acute schizophrenic patient. Henderson's (1969, 1970, 1971) community project involved token reinforcement for such things as making telephone calls to prospective employers, interacting with people in the community and completing a literacy programme.

Behavioural defects such as poor dressing (Heap, Boblitt, Moore and Hord, 1970), poor work performance (Ogburn, Fast and Tiffany, 1972), and lack of speech (Alumbaugh, 1971) have been modified, but perhaps more clinically relevant is the fact that symptomatic behaviour has been modified. For instance, in a study by Wincze, Leitenberg and Agras (1972), 'normal' speech was reinforced in patients with a delusional system. Out of 10 chronic schizophrenic patients, the delusional speech of 8 was diminished, 6 with varying degrees of success and 2 with total success. In addition, the design of the study makes it clear that tokens were responsible for the changes. Other symptoms modified have included thought disorder (Meichenbaum, 1969) and hallucinations (Rutner and Bugle, 1969), although it is important to recognize that there may be a limit to the amount of symptom change possible with these methods (Hutchinson and Baker, in press).

The research design of many token economies ignores one major fact; when token programmes are set up, other conditions are generated. For instance, the particular ward may receive more attention from professional staff, with consequent improvement in staff morale and attitude towards the patient. The nurses increase their efforts and provide a better standard of care for patients. They may now expect positive results. New activities and ward routines may be set up. In amongst all this the hitherto neglected patient receives far more stimulation and attention than usual. These factors have elsewhere been shown to influence behaviour (e.g. Suchotliff, Greaves, Stecker and Berke, 1970; Rosenthal, 1966; Cockburn, 1968). The relative contribution of these factors and the contribution of token reinforcement in the patients improvement have yet to be adequately assessed in the ward-wide situation.

The study here reported was a preliminary seven month study of token economy procedures with chronic schizophrenic patients in an English psychiatric hospital. The main aim of the study was to train the staff involved—a psychiatrist, two psychologists, three permanent trained nurses, and student nurses—in token economy methods, and to develop an appropriate set of assessment and therapeutic procedures. This was to prepare for a main controlled study which is still in progress. Some attempt was made to isolate factors other than token reinforcement in line with our criticism of other studies, although this study should be construed primarily as a demonstration and not as a controlled experiment.

**METHODOLOGY**

**Selection of patients**

The entire male population in the hospital was reviewed, and from them were selected all those patients who were under 50 years of age, had been in hospital over two years, and had a clear diagnosis of schizophrenia. Patients of known low intelligence and patients with known organic conditions were excluded. Patients who were clearly judged as paranoid by the assessing psychiatrist were not considered, as their response to incentives is known to be different to that of non-paranoid patients (Topping and O'Connor, 1960), and being put into an obviously manipulative system might have aggravated them. This gave a pool of 79 patients, who were then assessed on psychiatric symptoms by the Wing Psychiatric Rating Scale, and on degree of institutionalization by the Wing Ward Behaviour Rating Scale (Wing, 1961). From this pool, those 19 patients who were aged between 48 and 50 were considered for selection for the pilot study. Seven patients were chosen from the 19, whose range of symptoms and behavioural abnormalities were representative of the entire group. The remaining 60 patients left in the group served as the pool of patients for the main study.

The characteristics of the patients selected are shown in Table I. The Wing Ward Behaviour Scale Scores and Psychiatrist's Scale scores shown are those which formed the basis
of the selection of patients for the ward. These 7 chronic schizophrenic patients with varying behaviour problems were then removed from their original long-stay wards to the token economy unit.

**Layout of experiment and experimental procedure**

On arriving at the unit, the ward management and drug regime which had applied to each patient on his previous ward was, as much as possible, continued unchanged. This consisted of little more than custodial care, with some patients going off the ward daily to an occupational therapy or work task, together with most patients receiving a maintenance dose of phenothiazine drugs. In other words, no new treatment was introduced. As we had measured patients' behaviour in their previous setting, we hoped to be able to assess the effect of a change in wards, being on a smaller unit, having pleasanter material surroundings and a better nurse-patient ratio. This regime of unchanged treatment continued for six weeks.

As token economy programmes generally involve increased activation, an activity programme was introduced to try to assess this effect. Patients were exposed to far more stimulation than usual. For instance, each patient was now responsible for a task on the ward, such as washing up. An intensive occupational therapy programme was begun on the ward in the mornings. Trips to the cinema, swimming pool, shopping in Wakefield and sporting activities were held during the afternoon, and social evenings were organized, where games such as darts, dominoes and cards were played. Newspapers, books and television were also made available. This activity programme was continued throughout the whole pilot study. During this stage cash and comforts in kind were available to the patients as on their previous wards.

After three weeks tokens appeared for the first time, but were not contingent upon the patients' behaviour. At this point cash, and comforts in kind, were discontinued. The basic token (value 1) was a red plastic disc about an inch in diameter. In addition tokens of values 5 and 10 were introduced. In this period each patient received a free issue of 45 tokens a day, from which he had to pay 9 altogether for his meals and his own single room. The remaining tokens could be spent in the ward shop at the rate of 1 token for one new penny. We tried to make the shop as attractive as possible, with several brands of cigarettes and tobacco, sweets, biscuits, fruit drinks, magazines, items of clothing and many other goods. As we found that patients were only spending a limited amount of the tokens they had been given, various procedures were tried to extend spending, including pressure salesmanship to encourage extravagance, and the sampling of new goods. Altogether this period lasted for seven weeks.

At the close of the seventh week patients were
A TOKEN ECONOMY PROJECT WITH CHRONIC SCHIZOPHRENIC PATIENTS

told that they would have to earn tokens, and were instructed by means of a discussion and hand-outs how to do this. The contingent tokens were then introduced being earned by the satisfactory performing of various ward tasks which gradually increased in number and variety. The free issue of tokens was proportionately reduced as the ward tasks increased.

Altogether contingent tokens were in effect for 14 weeks. At the end of the 14th week we returned to the conditions that prevailed during the baseline testing, to act as a post-treatment measure. Patients then moved to other wards, and were re-assessed after two months.

Thus the layout of this pilot study was: 6 weeks of no treatment, 3 weeks of activity, 7 weeks of free tokens, 14 weeks of the token economy, 1 week post-treatment measurement, and a follow-up after 8 weeks.

**Description of the contingent token programme**

A description of some of the ways in which patients could earn tokens will help in understanding the system:

Every day before breakfast patients could earn tokens for shaving properly, making their bed and keeping their own room or dormitory area tidy, and for being at breakfast on time. After breakfast each patient had an allocated task for the day, such as doing the washing up, hoovering the day-room area etc. Tokens were given up to a fixed maximum, depending on the quality of the performance.

Apart from these tasks, there were 'floating tasks' for which patients could earn extra tokens. These were not allocated but were always available if patients spontaneously requested work. They included such things as making cups of coffee for the other patients, collecting the tokens from those who had requested drinks and handing them to the nurse, emptying ash trays, and helping to put away laundry. At the morning occupational therapy session on the ward the patients made basket-work trays, stools or rugs, and they could earn up to a fixed maximum for this work.

The dress of chronic schizophrenic patients is invariably dishevelled and untidy, and it was thought that this could be a factor operating against the ultimate discharge of the patients. Thus twice a day there was a spot check for neat appearance. On the basis of repeated observation during baseline testing we were able to select those items of dressing which were consistently poor for most patients, and one token was given for each of these items which had been correctly performed by the time of the spot check.

We did not consider these simple work tasks and dressing requirements to be the most important elements of the system. One activity upon which more importance was placed was social interaction; if a patient talked to other patients, helped them in some way, gave them things or communicated in any way, he could earn tokens. This was adjusted to the particular level of each patient, so that an extremely regressed patient could earn tokens for the spontaneous production of a monosyllable, whereas a less regressed patient would have to engage in a reasonably lengthy conversion. Nurses were aware that they were to continually raise their criterion for what they considered to be 'adequate' speech for each patient, so as to avoid the establishment of stereotyped patterns of social interaction.

We also considered it crucial to generate self-initiated activity. Many patients in chronic wards do not have to think for themselves—they merely go to meals when they are called, go to work when they are told, watch television when they are told, and so on. In our system, prompting was virtually abolished. Patients were not told 'time for O.T.' or 'clean your shoes'; if they performed any appropriate activity spontaneously they were given tokens, and if they did not they did not earn tokens.

Occasionally fines were used. These were generally only levied at dangerous acts such as patients lighting their cigarettes from the gas stove; but sometimes they were used as a component in an avoidance schedule. For instance, if the patients were sitting around doing nothing during an indoors games session, every 5 minutes they would be fined one token for this privilege. They were not told how to avoid the fine, namely by playing games, but were left to discover it for themselves, in their own way. Fines of this sort rarely had to be collected more than twice in a row, as patients
quickly discovered the more lucrative course of action.

Thus patients could earn tokens for simple activities such as ward tasks, O.T. performance, floating tasks and tidy dressing, but, more importantly, they could earn them for social interaction and self-initiated activity.

Occasionally programmes were initiated for individual behaviour problems, such as poor concentration or not wearing dentures, and an illustration of one such programme is given later.

**Assessment Methods**

The methods of measuring the changes in the patients are best subdivided into continuous measures and 'before and after' measures. Continuous measures included the measurement of specific behaviour and the use of more general psychiatric scales.

**Continuous Measures**

*Specific behaviour*

(a) **Dressing.** A daily rating scale for dressing was devised, consisting of 20 items which were rated as either present or not present. The rating was done by the nurses on three days each week. Examples of items are 'hair tidy', 'tie on', 'tie neat', 'shirt buttoned', and 'shoes laced': these were the five items mentioned earlier as being generally deficient. Therefore, they were placed among the actions for which tokens could be earned.

(b) **Initiative.** As mentioned previously, loss of initiative is a key factor contributing to the chronicity of schizophrenia, and initiative was therefore assessed while the patients were on the ward. A rating scale of 25 items of simple behaviour which normally occurred in the daily routine of each patient was constructed. Examples of the items rated were 'washing', 'shaving', 'making bed', 'sitting at table', 'eating food', 'clearing utensils', 'taking medication'. The degree of assistance (i.e. loss of initiative) required for each item was rated on a 3-point scale (0–2). This rating was done by the nurses on three days each week.

*General psychiatric scales*

(a) **Wing Ward Behaviour Rating Scale.** This is a 12-item scale and was completed each week by the nurses, who rated behaviour observed on the ward during the preceding week. The 12 items are scored on a 3-point scale (0–2) according to the frequency and/or severity of the behaviour. Wing has shown by factor analysis that this ward rating scale can be sub-divided into two factors: social withdrawal (8 items), that is to say lack of leisure interests, solitary behaviour etc., and socially embarrassing behaviour (4 items), that is to say posturing, talking to self, violent behaviour and over-activity.

(b) **Hamilton-Lorr Psychiatric Rating Scale.** The Hamilton modification of the Lorr rating scale (Hamilton, Smith, Lapidus and Cadogan, 1960), had the advantage for us of having been standardized with patients of a very similar cultural background to that of our own patients. Assessment on this scale is based on an open-ended psychiatric interview. The scale is straightforward and simple to use, being an 18-item 4-point scale. Yet it includes most of the symptoms and signs of chronic schizophrenia, for instance auditory hallucinations, various categories of delusion, incoherence, and poverty of speech, flattening of affect, etc. This assessment was performed on each patient at monthly intervals by the psychiatrist (K.H.).

**Before and After Measures**

(a) The structured clinical interview (S.C.I.) of Burdock and Hardestry (1968) was given. The S.C.I. consists of both an interview protocol and an inventory of 179 behavioural items, judged as present or absent. Norms are available for various comparison groups.

(b) A simple verbal conditioning task was carried out in two pre-treatment and one post-treatment session. Each session consisted of measuring the total amount of speech to five stimulus pictures ('tell me what you see'), in a baseline no reinforcement condition and in two subsequent social reinforcement conditions. As speech was reinforced during the study, comparing the total speech scores of pre-treatment and post-treatment testing should reflect any changes in speech.

According to the notion of secondary reinforcement, social approval, by being paired with tokens, should acquire greater reinforcing
value. One would therefore expect that in post-treatment testing patients would show an increased responsiveness to social reinforcement.

(c) Thirdly, the Leiter (1948) IQ test was given. This test has the advantages of being non-verbal, and of having a low basal age.

RESULTS

Earlier token economy studies have emphasized overall changes in groups of patients rather than in the individual patients. Allen and Magaro (1971) have pointed out that the type of patient best able to benefit from a programme cannot be identified from a group analysis, an argument which Skinner (1966) and Sidman (1960) have more generally advanced. A further problem with group analysis arises from the wide dispersion of scores on various measures. Scores are in some cases so good that there is little room for improvement, so that much of the apparent group variation may be attributable to the variation in scores of the relatively small number of more disturbed patients. For these reasons we will generally present both individual and group findings.

In the following diagrams where only 5 points on the graph are shown these refer to mean scores for each of the following periods: Baseline; Activity; Free tokens; First seven weeks of contingent tokens; Second seven weeks of contingent tokens.

CONTINUOUS MEASURES

Specific behaviour

(a) Dressing. From the dressing checklist certain key items were selected in the contingent token phase, for which token reinforcement was given. These are referred to as ‘target items’. The remaining items are referred to as ‘non-target items’. Fig. 1a shows mean group scores on target and non-target items during each of the experimental periods. There are three main findings:

(i) Spontaneous change. Fig. 1b shows the individual scores which comprise the group figure. It is noteworthy that before a token contingency was introduced the score on target items had improved considerably for 6 out of 7 patients. In 5 patients the degree of change in the three initial stages exceeded the degree of change caused by the token contingency. However, it is possible that the plateau found in the second seven weeks of contingent tokens might have been a function of the reduction of the maximum number of tokens that patients could earn for good dressing, which occurred in the last 3 weeks of this phase.

(ii) Individual differences. Fig. 1b also demonstrates how large the individual differences were. On the total checklist for instance, in the first week on the ward, the scores ranged between 10 and 33 per cent deviant dressing.

(iii) Target and Non-target change. The target items show a greater degree of change during all phases, as shown in Fig. 1a. The non-target items show an initial improvement from the baseline to the activity phase, but virtually no change thereafter. This graph also indicates the efficiency of the target-item identification procedure, in that the percentage of deviant behaviour covered by the remaining 15 items over the duration of the study remains at a low level.

Discussion. This particular target behaviour had improved so much before the contingent token phase, that a token reinforcement effect is overshadowed by other factors. These include better general care, more stimulation, and being on a smaller ward with better staff-patient ratios.
These may be sufficiently therapeutic to render a reinforcement programme for dressing unnecessary.

(b) Initiative. Fig. 2a shows the group changes in target and non-target items chosen from the initiative checklist, throughout the various experimental phases (the contingent token phase again being subdivided). The findings are as follows:

(i) Spontaneous change. Fig. 2b shows the individual scores on initiative, comprising both target and non-target items. In four cases where the patients initially required a marked degree of prompting the improvement due to the initial three phases far exceeded any change due to the token contingency, essentially replicating the findings with the patients’ dressing.

(ii) Individual differences. The individual differences as shown in Fig. 2b are vast. Two patients required 90 per cent prompting initially, while three patients initially required less than 10 per cent prompting.
(iii) Target and Non-target change. Fig. 2a indicates that the selection of target items for initiative is less efficient than for appearance. It also suggests a deterioration in non-target items, to a level worse than baseline, in the second part of the contingent token phase. This is a finding contrary to operant principles, since positive gains in target behaviour should have generalized to non-target behaviour which in this case was closely related in form.

Discussion. The deterioration in non-target behaviour during the contingent token period raises the question of whether there is a reinforcement effect at work. A model that fits the data better is that when tokens are given for initiative they are directing the patients attention towards these 'target behaviours' and away from their non-target behaviour. Consequently the non-target behaviour deteriorates. In other words a selective attention effect might be superseding a reinforcement effect.

(c) Individual sessions. To illustrate a type of behaviour which might receive individual attention we quote the case of our most regressed patient. One fundamental problem with him was that he could not sustain any activity. On even a simple work task, such as sandpapering a tray base, he would work sporadically, be easily distracted, and spend most of his time staring around him. Nurses had evolved a habit of telling him what to do, whereupon he would
start the activity, but in the absence of continued instructions would soon cease. Fig. 3a illustrates the percentage of time the patient worked at sanding a tray. At the start of the first minute the nurse prompted him to work and left him for ten minutes. It can be seen that in those ten minutes his working performance steadily declined. A programme was then started to encourage sustained performance at this task. This involved reinforcing the patient with a blue token and approval for 15 seconds continuous work after a prompt. When such an interval of work could be maintained the patient would be reinforced for 30 seconds continuous work, and so on, working up to greater intervals. In addition, since the patient would only sustain work when a nurse was in the vicinity, nurses gradually increased the distance they were away from the patient, until they were able to leave him entirely. In other words there was ‘shaping’ of the amount of time this patient worked, and ‘fading’ of the nurses presence. If the patient was distracted—e.g. stopped working when there was a noise outside—he would be fined one token, with nurses expressing their disapproval of this. As blue tokens were given very frequently, we established an exchange rate whereby six blue tokens for 15 - 90 secs. work

5 sessions of tokens for 5 mins. work

5 sessions of tokens for 7 mins. work

FIG. 3.—Sustained attention of Patient B.
tokens could be cashed in for one red token, the basic unit in the general ward programme. Figs. 3b, 3c, 3d, and 3e, respectively represent measures of his performance as the intervals of time the patient had to work increased. Measures were taken under conditions of no reinforcement as during the baseline measure. After 21 sessions it can be seen that this patient was able to sustain his attention for ten minutes unprompted. A different task (sorting four different coloured toys into piles), was then given to him, and he was also able to sustain his attention to this, demonstrating that the learning was not specific to one task only.

General psychiatric scales

(a) Wing Ward Behaviour Rating Scale. As was mentioned earlier, this scale is comprised of two factors, social withdrawal and socially embarrassing behaviour. The social withdrawal factor consists of 8 items with a maximum disturbance score of 16, and the socially embarrassing factor consists of 4 items with a maximum disturbance score of 8.

Fig. 4 presents the weekly mean scores for the group of 7 patients on both social withdrawal and socially embarrassing behaviour, and Table IIA and IIB presents statistical comparisons between each phase. As can be seen, from the figure, social withdrawal improved throughout the first three phases, the difference between the phases being statistically significant. There was no further improvement when contingent tokens were introduced, if anything a slight initial deterioration in the first seven weeks (this was insignificant). One important feature was that at the two month follow-up withdrawal had deteriorated to the original baseline level. This apparent deterioration may be due to rater unreliability, since the follow-up assessment was carried out on a different ward by a nurse relatively unfamiliar with the use of the Scale. Socially embarrassing behaviour, on the other hand, appears from the figures to have deteriorated during the pre-token phases (although this proved insignificant) and continued to deteriorate during the contingent token phase, reverting to the former low level of disturbance during follow-up. The deteriorated socially embarrassing scores during contingent tokens
were significantly worse than those of any other phase, at high levels of statistical confidence. After the eighteenth week measures were taken to improve the reliability of the socially embarrassing items,* and such an elevation in scores might possibly be accounted for by this changed assessment method rather than by the contingent tokens. However, the associated Hamilton-Lorr scores would suggest that this is not the case.

(b) Hamilton-Lorr Psychiatric Rating Scale. To assist in interpreting results, the individual items in the Hamilton-Lorr rating scale have each been allocated to an accepted area of

* This involved keeping a frequency count of each instance of 'embarrassing behaviour' rather than nurses estimating the frequency from their memory of the past week, as occurs in Wing's original version.
clinical disturbance in chronic schizophrenia. Thus, in addition to the total scores, we can consider results in terms of thought disorder, delusions, hallucinations, catatonic symptoms and emotional disturbance.

The mean total scores for the group at each stage of the experiment show a slight improvement in each phase up to the introduction of contingent tokens, and then a slight deterioration in the contingent token phase, with none of these changes being statistically significant. There is then a highly significant deterioration from the beginning of contingent tokens to the two monthly follow-up point ($p < 0.001$). In general the total Hamilton-Lorr scores for the group move in the same direction as social withdrawal on the Wing Ward Behaviour Rating Scale, i.e. an initial improvement up to the stage of contingent tokens, then a gradual deterioration during contingent tokens and finally a marked deterioration in the follow-up. The individual patients who contributed most to these changes were the ones who were originally most withdrawn clinically.

Fig. 5 shows the individual symptom areas which make up the total scores. It can be seen that the deterioration in the contingent token phase is accounted for by the areas of thought

---

**Fig. 5.—Hamilton-Lorr Psychiatric Rating Scale.**
disorder, emotional disturbance and catatonic symptoms. This deterioration also persisted during the two months follow-up period.

**Before and After Measures**

*Structured clinical interview*

The S.C.I. was given to each patient during the baseline six weeks, and later in the post-treatment week, by one of the authors. Our findings were as follows:

(i) In terms of the total number of items rated as deviant there was a significant improvement of patients from pre-testing to post-testing ($t = 4.05; p < 0.01$).

(ii) Of the ten sub-scales of the S.C.I. two (conceptual dysfunction (28 items) and lethargy-dejection (17 items)) accounted for 67 per cent of all the items rated as deviant in both pre-testing and post-testing. Two further sub-scales (incongruous behaviour (25 items) and incongruous ideation (28 items)) accounted for a further 23 per cent of the deviant items. In other words, for our patients, only four sub-scales were relevant in describing their aberrant behaviour.

(iii) In each of these four sub-scales we looked for changes of two items or more from pre-testing to post-testing.

In conceptual dysfunction, 5 patients out of the 7 changed, 4 for the better.

In lethargy-dejection, 4 patients changed, all for the better.

In incongruous behaviour, 3 patients changed, all for the better.

In incongruous ideation, 1 patient changed for the better.

*Verbal conditioning test*

Each of the three sessions involved 5 baseline trials followed by 10 reinforcement trials. Two of these sessions were held in the baseline phase of the experiment and one after the contingent token phase. This procedure gives information on: (a) Change in the total amount of speech from pre- to post-experiment testing; and (b) Change in the rate of learning from pre-testing to post-testing. Table IIIA indicates that patients C and F showed the most marked improvement in amount of speech from sessions 1 and 2, to 3. For these patients the total seven months of the ward programme had the greatest general effect in increasing amount of speech. The improvement for the whole group from pre-testing to post-testing was not significant. Table IIIB shows the amount of learning occurring...
within each session. It can be seen that for all except patients B and E there is no appreciable change in the rate of learning before and after the experiment. The two patients who changed both showed less learning within the session after the experiment, possibly suggesting less sensitivity to social reinforcement. These patients were the lowest of the group in the measurement trials of all 3 sessions.

**IQ**

The Leiter IQ test was given before and after the experiment. The average mental age at the beginning was 6.3 years, ranging from 4.1 to 9.5 years. After the experiment the mental ages of 5 patients had improved, and 2 had shown no change. The whole group showed a mean improvement of about 0.4 of a year, and this was statistically significant (t = 2.68; p < 0.05). Table IV shows the before, after, and change scores.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Pre-</th>
<th>Post-</th>
<th>Change (in years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6.5</td>
<td>7.5</td>
<td>1.0</td>
</tr>
<tr>
<td>B</td>
<td>4.08</td>
<td>4.33</td>
<td>0.25</td>
</tr>
<tr>
<td>C</td>
<td>9.5</td>
<td>9.5</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>8.25</td>
<td>8.25</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>5.5</td>
<td>5.75</td>
<td>0.25</td>
</tr>
<tr>
<td>F</td>
<td>4.33</td>
<td>4.75</td>
<td>0.42</td>
</tr>
<tr>
<td>G</td>
<td>6.25</td>
<td>7.0</td>
<td>0.75</td>
</tr>
</tbody>
</table>

\[ t = 2.68 \]

**TABLE IV**

Mental ages (Leiter)

As a training exercise the experiment was a considerable success. By the end of the study the nursing staff were confident of their ability to carry out the programme, and the devolution of responsibility for major aspects of the programme has been well received. They now have a consistent approach to the patients, and have shown themselves fully capable of making on-the-spot decisions which have been in good accord with the theory of operant conditioning.

In terms of organizing a token economy the results indicate some guidelines:

1. The great amount of individual variation of scores indicates the difficulty in dealing with a heterogeneous group, such as ours. If patients are reinforced at different levels of behaviour the nurses have the difficult task of remembering a large set of different contingencies. On this ground homogeneous groups are to be preferred, so that similar contingencies apply to all patients.

2. Also, it is redundant, even harmful, to bring under token control some activity that was already under the patients' own control, since internal control is one of the ultimate aims of therapy. A blanket reinforcement programme, where for instance all patients are reinforced for dressing, is for this reason not likely to help a heterogeneous group of patients, but obviously can help a homogeneous group.

3. As social withdrawal had deteriorated considerably after the patients had been off the ward for two months, a follow-up is necessary. Carlson, Hersen and Eisler (1972) have pointed out that follow-up has not in fact been observed in many token programmes. Our result points to the further need to develop a prophylactic post-treatment environment, where improvements can be maintained, and to the need for adequate preparation for return to other settings. If this is not done, old contingencies—such as nurses doing jobs for patients—may reassert themselves, to the patients' detriment.

Patient B was in fact very rapidly brought under the control of old contingencies by ward staff who remembered him as he was and who responded to him accordingly.

**DISCUSSION**

The patients have undoubtedly improved as a consequence of this programme, and over the course of the programme some patients showed considerable changes in specific areas. Clinically, the experiment was therefore justified.

From the nursing point of view there was also change during the programme. After initial staffing difficulties, the permanent nursing staff settled down, and have all continued on the ward during the ensuing main experiment.
for adequately measuring psychiatric state while patients are on a token economy. Although it transgresses the theoretical framework of operant conditioning to talk of an 'illness' or to consider psychiatric symptoms as anything other than operant behaviour, the results suggest that the particular behaviour referred to as a 'symptom' should be adequately monitored, even though this behaviour may not be directly reinforced.

(4) There was little evidence that a specific token contingency was the main factor in changing the patients' target behaviour. The greatest change for most patients was generally during the early stages of the experiment. This would indicate the necessity to integrate other approaches—such as stimulation, structured activity, personal possessions and responsibility, high nurse morale etc.—with any token procedures. A token system does not absolve nurses from the need of consistent care, material provision of clothes etc. Such things as availability of laundering services, and clothing that fits, may be more important in improving the patients appearance, than token reinforcement of neat dressing behaviour.

To consider the results in perspective, a cautionary word is necessary about the experimental design. The layout of the experiment has the implication that each treatment phase can be assessed before the introduction of the succeeding phase. In the present study stability within each phase was not achieved before the next phase was begun, because of limitations in time available. Some of the results show a linear trend, so that the results could be attributable to a single factor effective from the beginning of the transfer of the patient to the unit. The design of the study does not therefore preclude the possibility that some of the changes, even if significant, could be due to time alone. It is important that the length of each experimental phase should be adequate, since there is suggestive evidence (Appleby, 1963) that a period of at least eight months is necessary for treatment effects to stabilize with chronic patients. Serious flaws have also been pointed out in this type of experimental design when different treatment conditions succeed each other (Poulton and Freeman, 1966; Johnson and Lubin, 1972). Interactive or inhibitory carry-over from one treatment to another may well occur. Where the time intervals of treatment are short, as in this study, the likelihood of this type of flaw is further maximized.

Concerning the 'economy' side, Winkler (1971) has suggested that token economies resemble the theory of technology of economies at large. He quotes Engels' law as stating that with changes in income it is luxury spending that varies, whereas the amount spent on basic necessities remains the same. Our results confirm this—spending did not change on the necessities, which in our case was meals and sleeping accommodation. Patients rarely missed meals through lack of tokens. It was the luxury spending that varied according to token earnings, as can be seen from Fig. 6. When the free issue of tokens decreased during weeks 18, 19 and 20 it was luxury spending that fell.

Winkler has also stated as an economic principle that the more patients get, the more they spend. While this may be generally true, our sample of 7 chronic schizophrenic men certainly included three who were satisfied with a very few luxuries and would not spend more when having more tokens. As this failure to want more reinforcers is as much a part of apathy as other behaviours, it is obviously a desirable target of therapy to extend the amount and variety of goods or events that will act as reinforcers. Ayllon and Azrin (1968b) have suggested a method for doing this, but we found it very difficult to achieve. In other words we could not make these three patients want more than a meagre amount. It is of interest that these patients were those with the highest scores on the Wing withdrawal factor at the start of the study. Cohen, Florin, Grusche, Meyer-Osterkamp and Sell (1972) in discussing their tokens programme have referred to Meehl's formulation of anhedonia (1962), which suggests that one of the central deficits of schizophrenia is the patient's lack of wants. In operant terms the failure to find adequate reinforcers is a key problem, since behaviour change is said to be a function of the reinforcer. It has been suggested (Dovey, 1973) that this limitation in scope of reinforcers is a characteristic of a male rather than a female population.
A TOKEN ECONOMY PROJECT WITH CHRONIC SCHIZOPHRENIC PATIENTS

In summary, we have generally found encouraging results in running a token economy, although there are a few disturbing problems which certainly demand a more thorough investigation. Very careful experimental design and assessment procedures are necessary in order to determine the exact factors responsible for changes in patients on token economy wards.

**Summary**

A group of 7 schizophrenic patients with varied symptoms were moved to a ward where different aspects of a token economy system were sequentially introduced. At first patients were just exposed to a smaller ward with a better standard of care (Baseline). Then a stimulating programme of daily events was started (Activity). Next free tokens were given (Non-contingent) and lastly the full scale token reinforcement procedure was instituted, where patients had to earn tokens for appropriate behaviour. The reason for this design was to try to understand more about the therapeutic ingredients at work in a token economy system. Theoretically the reinforcement by tokens is the key agent of therapy and so the last phase should have caused the most improvement.

Patients improved in several areas, notably on dressing and initiative, and on withdrawal and thought disorder, although most of the improvement occurred before the token reinforcement phase was instituted. Also, psychiatric deteriora-
tion was associated with the token reinforcement phase. In all, token reinforcement did not emerge as the critical therapeutic agent.

Acknowledgements

This study was carried out at Stanley Royd Hospital, Wakefield.

We gratefully acknowledge the support of the Mental Health Research Fund, The Medical Research Council, and the Department of Health and Social Security, and the help and advice given us by Prof. Max Hamilton and Prof. Gwynne Jones in organizing the research. The study could not have been carried out without the enthusiastic co-operation of the ward nursing staff—Mr. G. Rosenthal, Mr. H. Ineson, and Mr. R. Agbah—and the unit nursing officer—Mr. P. Andrews. We are also grateful for the help given to us by the medical, nursing, and administrative staff of Stanley Royd Hospital.

References


Hutchinson, K., and Baker, R. 'The effect of token economy regimes on the symptoms of schizophrenic patients.' (In preparation).


A TOKEN ECONOMY PROJECT WITH CHRONIC SCHIZOPHRENIC PATIENTS


A synopsis of this paper was published in the August 1973 Journal.

Roger Baker, B.A., Ph.D., Mental Health Research Fund Bates Fellow in Schizophrenia, Department of Psychiatry, University of Leeds

John N. Hall, B.A., M.Sc., A.B.Ps.S., Medical Research Council Research Fellow, Departments of Psychiatry and Psychology, University of Leeds*

Keith Hutchinson, M.B., Ch.B., D.P.M., Medical Research Council Research Fellow, Departments of Psychiatry and Psychology, University of Leeds

(Received 9 April 1973)

* Reprints are available from John Hall, Department of Psychiatry, 15 Hyde Terrace, Leeds LS2 9LT.
A Token Economy Project with Chronic Schizophrenic Patients

ROGER BAKER, JOHN N. HALL and KEITH HUTCHINSON

Access the most recent version at DOI: 10.1192/bjp.124.4.367

References

This article cites 0 articles, 0 of which you can access for free at:
http://bjp.rcpsych.org/content/124/581/367#BIBL

Reprints/permissions

To obtain reprints or permission to reproduce material from this paper, please write to permissions@rcpsych.ac.uk

You can respond to this article at

/letters/submit/bjprcpsych;124/581/367

Downloaded from

http://bjp.rcpsych.org/ on June 25, 2017
Published by The Royal College of Psychiatrists

To subscribe to The British Journal of Psychiatry go to:
http://bjp.rcpsych.org/site/subscriptions/