

*Study of the Altro Health and Rehabilitation Services' program for rehabilitation of posthospitalized mental patients illustrates some requirements for evaluative research in mental health. Limitations imposed by control group designs and selective biases arising from the operation of the program are taken into account.*

## Evaluating a Rehabilitation Program for Post-Hospital Mental Patients

HENRY J. MEYER, Ph.D., and EDGAR F. BORGATTA, Ph.D.

INTEREST in rehabilitation services for posthospitalized psychiatric patients has been stimulated by an increase in release rates from hospitals, following introduction of drug therapies, and by wider recognition that family, employment, and other social conditions appear to affect successful adjustment in the community (1-4). This interest may be thought of as a counterpart to the even greater attention paid to preventive programs (5).

Whether programs are directed toward prevention or rehabilitation, claims for their efficacy inevitably raise the question of evaluation. In the interest of efficiency as well as economy, sound assessment of effectiveness is obligatory for those who promote these programs as contributions to mental health. The need for evaluative research is widely acknowledged (6). But evaluative research on a service program in the field of mental health is, as

Clausen has said, "so appealing and yet so hazardous" (7).

Because of the pressing need to evaluate programs in which much time, effort, and money have been invested, most studies fall short of the rigor expected of scientific research. This is understandable in view of the difficulties of evaluating complex, changing, and often experimental efforts. Studies of the staff's or clients' satisfaction with a program, subjective judgments about progress or change, case studies to illustrate success, and similar reports may contribute useful knowledge about the operations of a program and may be valid sources of insight and understanding. But it is of no help to the orderly development of scientific knowledge to accept these studies as demonstrations of success or failure when it is possible to attempt more rigorous research. The state of our ignorance and the means of overcoming it should be accepted so that we may proceed slowly, and often painfully, to gain secure knowledge of what is being accomplished. The conclusion that this is the way of progress in evaluative research has been recognized not only for programs directly concerned with mental health but also for related fields, such as social casework, prevention of delinquency, and education of parents (8-10).

---

*Dr. Meyer is professor of sociology and social work at the University of Michigan, Ann Arbor. Dr. Borgatta is associate social psychologist at the Russell Sage Foundation and adjunct professor of sociology at New York University. This paper was presented in somewhat different form at the annual meeting of the American Sociological Society in Washington, D. C., on August 27, 1957.*

Many technical and practical problems are encountered when a rigorous evaluative study is attempted. This paper will not discuss them. Instead, we shall consider how certain decisions about the research design and how certain problems that arise in executing the research within an operating program restrict the interpretation of the findings. Using a recent effort to evaluate the effectiveness of a rehabilitation program for posthospitalized mental patients as an illustration, we shall examine in particular the limitations imposed on the definition of what is being evaluated.

### The Altro Project

The specific rehabilitation program that was subject to our evaluative investigation was that of the Altro Health and Rehabilitation Services, Inc., located in the Bronx, N. Y.

For more than 40 years Altro has operated a sheltered workshop for tuberculous patients, and it has served cardiac patients for more than 6 years (11, 12). After extended periods of hospitalization, these patients come to the Altro workshop where they remain, on the average, about a year before they graduate. They work in a factory environment but with work pressures adjusted to their health rather than to the demands of production. The workshop is, indeed, a modern garment factory manufacturing uniforms for nurses and hospital gowns that are sold by a sales force on the competitive market. Patients are paid union scale wages at piece rates for goods produced. They are subsidized when necessary to permit them to remain in the rehabilitation program.

The patients are under close but unobtrusive medical supervision; they are provided with periodic health examinations and continuous convalescent treatment when required. All patients put in a full day, but the actual time each works is determined by his particular needs. Work and rest periods alternate in recognition of disability and the development of work tolerance. Supervisors in the factory and instructors in the office-training program are practiced in dealing with the special problems of posthospitalized patients and many are themselves graduates of Altro. Each patient is assigned a caseworker and other social services

are available such as vocational counseling, job placement, and educational assistance.

The program is intended to harden the patient so that his transition from hospital to demands of normal living will not be so abrupt as to threaten his recovery and precipitate rehospitalization. By letting patients test the limits of their capacities in its workshop, Altro hopes that its clients will learn to function independently in the community.

From retrospective studies there is evidence that Altro makes an effective contribution to the rehabilitation of tuberculous and heart patients (13, 14).

With the need of tuberculous patients for rehabilitation services decreasing, and with the cardiac program well established, Altro turned its interest to the inclusion of psychiatric patients. In 1953 Altro began a year of experimental collaboration with Hillside Hospital, a private mental hospital in New York City (15). Ten patients were referred to Altro by Hillside, and seven participated in the workshop. With this experience added to the years of serving other types of patients, Altro felt ready to extend its services tentatively to certain types of hospitalized mental patients from the more heterogeneous populations of State hospitals (16). It did so with the widespread encouragement of psychiatrists and other professional persons concerned with mental health. Indeed, throughout discussions of potential sources for the limited number of patients that Altro felt able to serve, the director of Altro was repeatedly assured that this sort of rehabilitation service was greatly needed in the psychiatric field and that there would be strong demand for it.

With such assurances, arrangements were concluded with the New York State Department of Mental Hygiene for exploring the value of this service to patients of its Bronx aftercare clinic. Financial support was obtained to extend Altro's program, and the Russell Sage Foundation gave financial assistance for an evaluation of this effort. In accepting such assistance, Altro committed itself to a control-group design as an evaluative requirement. The patients were to be assigned at random to the experimental and control groups. The research question could therefore be stated at the outset as fol-

lows: Given matched groups of patients, will those who receive Altro services show greater progress, by some criterion of successful rehabilitation, than those who do not receive such services?

Altro planned to accept about 80 psychiatric patients during the 2 years of the study period. These patients would at any given time constitute about 20 percent of the normal caseload of approximately 200 patients at the workshop.

### Defining Experimental and Control Groups

A service program, such as Altro's, is in general oriented to practice rather than to scientific research, and the canons of scientific methodology frequently appear to contradict the canons of practice. Altro's acceptance of a control-group design was therefore a bold step toward rigorous evaluation. As decisions were made to further this design, however, it was necessary to make some of the implications explicit.

With a control-group design one is in a position to say whether or not a program makes a difference to some population, on the basis of some criterion, such as return to the hospital. A program of rehabilitation includes many factors that might affect the client's welfare. The Altro program offers medical care, casework, vocational training, and other services as well as its distinctive workshop. Unless these factors can be specifically identified so that their counterparts can be examined in the experience of persons in the control group, our conclusion must be framed in general terms. We could say that Altro's patients in general fare better than non-Altro patients. From the standpoint of the community, this might be called the potential impact of Altro. It leads to statements about what the consequences would be if all patients were provided the same services as Altro's patients.

This kind of conclusion is very useful. It would require, however, further qualification in terms of the expressly defined population of patients to which it could be applied. If experimental and control groups were drawn from the population of all ex-hospitalized patients, our conclusion would mean one thing. If they were drawn from a population of patients limited by

interest in Altro, by stated characteristics, or by other definitions, our conclusion would thereby be restricted. This may be illustrated by considering a number of possible definitions of the subject population that might (ignoring practical considerations) be applied in this study.

Patients who were released from mental hospitals could respond in at least the following four ways to invitations to enter the Altro program:

1. They could be uninterested and decline the invitation.
2. They could be interested, explore the possibility, and decline to enter.
3. They could be interested, explore the possibility, enter, and withdraw at some stage in the program before it had been completed.
4. They could be interested, explore the possibility, enter, and graduate.

Furthermore, patients responding at any of these levels would constitute a population all of whom had had at least some information about Altro, ranging from mere knowledge of its existence to full participation in it. Therefore a no-contact category is logically required. Of such, at least two subcategories must be recognized since the distinction between them might be relevant to rehabilitation: (*a*) those ignorant of Altro altogether, and (*b*) those who know of Altro's existence but have had no further contact with it. Knowing about Altro could result in such meaningful reactions as: Is Altro the kind of place for people like me? Is it good or bad that people want to help former patients?

Similarly, each of the successive levels of contact noted above might reasonably be expected to have some effect on the patient's rehabilitation. Would the patient view an invitation to Altro as supportive or threatening? If a patient declined an invitation to Altro would he be punished or rewarded? And so on.

Generalization would be limited to that level selected as the population from which experimental and control groups were chosen at random. If a no-contact comparison group were used, statements about the impact of knowledge of Altro plus subsequent contact would be permissible. If the population were defined as patients informed about Altro but with various levels of contact with it, generalization would be appropriate only at the specified level. If

we want to know what Altro achieves only through its treatment services, we would have to compare those who went through the program with patients who might have gone through if they had had the opportunity. Still undefined would be which of the many things included in the rehabilitation program produced given results for experimental cases as compared with control cases.

It is conceivable that the treatment experiences for the experimental group could be recorded in detail at all points after their selection. In this manner specific treatment efforts might be related to variations in the degree of success or failure observed after treatment. It would be necessary, however, to have similar detailed knowledge of the experiences of the comparison group. Some of the features of the rehabilitation program at Altro, sympathetic work supervision, for example, might well be present in the experiences of patients in the control group.

Similarly, the effectiveness of any one of the specific kinds of treatment services could be stated only if the method of assignment to receive different services was governed by a random rather than a selective procedure.

For this particular evaluation the population was defined as patients who accepted the invitation to enter the Altro program. Among such, the treatment and comparison groups were selected at random and our generalizations must apply only to this population. The point here is that the design limited generalization to a segment of the population that might provide the answer to the question: Is Altro effective in rehabilitating posthospitalized mental patients?

### **Defining Treatable Patients**

Restrictions from another side limit generalization in evaluative research. Any given treatment or service program assumes a part of the answer to the question of its own effectiveness by directing itself to predetermined categories of patients. Thus it says, in effect, we will (or can) work better with one type of client than with another.

The "community function" of the agency draws the broadest boundaries; for example, to

serve the aged, the adolescent, the tuberculous, or the ex-mental hospital patient. Within these boundaries further criteria of inclusion or exclusion are explicitly or implicitly accepted. Adjectives modify the clientele: healthy aged, female adolescents, arrested tuberculosis cases, or ex-hospitalized schizophrenic patients, are acceptable for treatment.

Still further exclusions are made by defining the type of services available; for example, healthy aged who need a home, adolescent girls who are going to have out-of-wedlock babies, arrested tuberculosis patients who can be expected to return to work or housekeeping, or ex-hospitalized schizophrenic patients who need rehabilitation.

Finally, among those potential clients for whom the services are intended, agencies develop conceptions of clients who can be best served by their skills. The rationale for these conceptions rests on estimates of competence in practice based on professional training and experience. It accepts, indeed, an evaluative conclusion before the question of evaluation is posed for research.

Thus the question asked of evaluative research might be restated as follows: How successful are we with those clients we want to serve and think we can help? This is an entirely legitimate question but it is much more restricted than the question: How effective is our program?

Research operations that come to grips with this restriction must seek explicit criteria to define the subject population from the viewpoint of the service agency. In the Altro project, staff and consultants formulated these general criteria: Bronx residents, 20-40 years of age, admitted only once to a mental hospital for 3-24 months, having formal psychiatric diagnosis of dementia praecox. In addition, to protect the going workshop program and in the interests of the patient, clinical criteria were to be applied in interviews by a psychiatrist to eliminate those who were "too sick" (revealing disturbing or dangerous psychotic symptoms, physically incapacitated, addicted to drugs), those who were "too well" (not in need of rehabilitation), and those having "alternative plans" (employment, household duties). Cases that survived all these criteria would be avail-

able to Altro and hence the subjects of the evaluative research.

A conflict between voluntary and authoritarian programs of treatment becomes pertinent at this point. Most private agencies, Altro included, adopt a treatment philosophy that expects the client to "accept" or "want" help in contrast to a treatment philosophy that asserts what the patient "should" or "must" do or have done to him. For example, State mental hospitals get "permission" to administer shock therapies; private agencies seek "cooperation" and "acceptance" of casework by their clients. Without prejudging the success of these approaches, we point out that the compromise of these conflicting philosophies of professional responsibility injects an additional selective definition of the subject population in evaluative research. We can describe some of the effects of this selective process on the Altro project.

Of the total caseload of ex-mental hospital patients at the Bronx after-care clinic, about 18 percent qualified under the general criteria enumerated above. Of these, about 28 percent survived the clinical screening, constituting only about 5 percent of the general population of posthospitalized mental patients. This is the point at which treatment and control groups were selected and evaluation could be said to apply only to these patients. Analysis of these groups, selected at random, indicates that the two samples do not differ significantly at the time of assignment with respect to any known characteristics.

The pool of patients from which both experimental and control groups were drawn does differ from the population of ex-hospitalized patients meeting the general criteria. The clinical screening process tended to anticipate the kind of clients the agency might favor. Specifically, the group of "available" patients includes a disproportionate number of single, white, better educated persons who have a higher occupational status, and who live with their parental families.

Because the selection just described occurs before experimental and comparison groups are chosen, it may be said to be controlled, if follow-up and before and after measurements are applied to these samples. However, it cannot be

assumed that the process of selective bias will cease at this point.

The professional approach of those providing service tends to select patients of unknown characteristics among those available to receive treatment. The logic of evaluative research would require, without compromise, that those designated for treatment be given treatment; the values of social casework and rehabilitative practice tend to assume that treatment should only be given to those who would accept it. Practitioners often argue that they can be effective only with such patients although this is one of the key questions that evaluative research is supposed to answer.

Of the pool of patients available as experimental cases for the Altro project, about a third had no contact whatsoever with Altro, another third had limited contact with members of the staff, and the final third actually entered the workshop. If the latter group is considered to be the only one composed of cases subject to the rehabilitation program, it constitutes a selected population for which the principles of selection might be explored retroactively. If the individuals in that group differ from the control group, the differences may be attributable to selection and not to treatment. Comparing patients (*a*) who received treatment, (*b*) who had lesser contact with Altro, and (*c*) who were offered services but received none with the control group can provide a picture of how this selective process operated. We can attempt to match each of these sets of experimental patients with patients in the control group. But the rigor of random matching cannot be claimed except to assess the effect of being offered the opportunity to enter Altro.

Inspection of the consequences of the selective process suggests that those patients reaching the rehabilitation program proper tend to be of two types: (*a*) those rather highly motivated to accept the kind of casework help and vocational training Altro offers, and (*b*) those who accepted Altro largely because they seemed unable to make any other sort of adjustment to the world outside of the hospital. Naturally enough, Altro is likely to feel that it is successful with the former and unsuccessful with the latter.

Were it possible to draw experimental and

control samples from those patients who actually enter the workshop and who are therefore subject to the rehabilitation program, we could evaluate the program. Because this would require arbitrary denial of service to some of those who have already accepted it, such a design is difficult to execute even for the most willing operating agency. A more feasible alternative would be to require the agency to extend maximum effort to see that patients selected for the experimental group were brought into the program.

### **Defining Successful Rehabilitation**

Whatever the experimental design, the definition of what constitutes successful treatment poses additional issues. These will not be considered in detail in this paper. Criteria for successful rehabilitation can range from "objective" measurements, such as permanent avoidance of readmission to a hospital, to clinical judgments about the level of mental health of the patients, with various indicators of adjustment in between.

Whatever criteria are used, the value of a study is enhanced if the experimental design requires impartial application of the criteria to experimental and control groups alike. It will not, to be sure, support general statements about effectiveness of a program unless a fairly inclusive range of criteria is offered. It is at this point that some theory of what the rehabilitation program is supposed to do for patients becomes indispensable. Theory about treatment of the mentally ill is in a state of flux today. Therefore evaluation studies should, in our opinion, always include indicators such as rate of readmission along with other criteria. Attention should be given as well to the duration of whatever effects are observed.

In the Altro study, hospital status 1 year after release is the primary criterion but, in addition, employment adjustment, presence and type of psychiatric care, and a judgment of the patient's current competence for managing are included in the followup interview. Brief before-and-after attitude tests have also been used.

### **Conclusion**

If a research plan as promising as that reported here encounters such serious obstacles, is

it futile to attempt rigorous evaluative research that is so badly needed? We think not. Several kinds of contributions can be made, and their importance will be directly related to the rigor with which the evaluation was planned and executed.

In the first place, a description of the selective process is of crucial importance to the interpretation of evaluation studies, and this is seldom reported in the literature of mental health. Second, analysis of the results of well-designed evaluation studies, even when they fall short of full success, will contribute to an understanding of the service program with greater certainty and more appropriate caution. Furthermore, each carefully conceived effort to meet the requirements of valid evaluation will provide experience to enhance the next attempt.

We sought to make explicit some of the inherent requirements of evaluative research. The difficulties encountered in the Altro project should not discourage evaluative research; their identification is the first step toward overcoming them. But this report should encourage modesty in making claims for service programs in mental health. As more certain knowledge accumulates, we may expect these programs to become more firmly based and demonstrably more effective.

### **REFERENCES**

- (1) Brill, H., and Patton, R. E.: Hospital population trend and drug therapy. *Pub. Health Rep.* 72: 629-632, July 1957.
- (2) Greenblatt, M., Levinson, D. J., and Williams, R. H.: *The patient and the mental hospital.* Glencoe, Ill., Free Press, 1957 (especially Part IV).
- (3) Schwartz, C. G.: *Rehabilitation of mental hospital patients.* Public Health Service Pub. No. 297. *Public Health Monogr.* No. 17. Washington, D. C., U. S. Government Printing Office, 1953.
- (4) Rennie, T. A. C., Burling, T., and Woodward, L. E.: *Vocational rehabilitation of psychiatric patients.* New York, Commonwealth Fund, 1950.
- (5) Kotinsky, R., and Witmer, H. L. (eds.): *Community programs for mental health.* Cambridge, Mass., Harvard University Press, 1955.
- (6) National Institute of Mental Health: *Evaluation in mental health, a review of the problem of evaluating mental health activities.* Public Health Service Pub. No. 413. Washington, D. C., U. S. Government Printing Office, 1955.

- (7) Clausen, J. A.: *Sociology and the field of mental health*. New York, N. Y., Russell Sage Foundation, 1956, p. 41.
- (8) Kogan, L. S.: Evaluative techniques in social case work. *Soc. Serv. Rev.* 26: 305-309 (1952).
- (9) Witmer, H. L., and Tufts, E.: The effectiveness of delinquency prevention programs. *Children's Bureau Pub. No. 350* (1954).
- (10) Brim, O. G.: Evaluating the effects of parent education. *Marriage & Family Living* 29: 54-60, February 1957.
- (11) Hochhauser, E., and Hentel, C.: History, philosophy, and operation of the Altro Work Shops, Inc. *Brit. J. Phys. Med.* 17: 3-7 (1954).
- (12) Hochhauser, E.: Industrial convalescence in the sheltered workshop. *J. Rehabil.* 19: 8-13, March-April, 1953.
- (13) Siltsbach, L. E.: *Clinical evaluation of the rehabilitation of the tuberculous experience at Altro Work Shops 1915-1939*. New York, National Tuberculosis Association, 1944.
- (14) Black, B. J., Jezer, A., and Adelson, D.: *Evaluation study of the Altro cardiac project*. New York, Altro Health and Rehabilitation Services, Inc., 1957. Mimeographed.
- (15) Bellak, L., Black, B. J., and others: Rehabilitation of the mentally ill through controlled transitional employment. *Am. J. Orthopsychiat.* 26: 285-296 (1956).
- (16) Black, B. J.: *The workaday world: Some problems in return of mental patients to the community*. In *The patient and the mental hospital*. Glencoe, Ill., Free Press, 1957.

## technique

# Use of Plastic Trays in the CF Test

The bureau of laboratories of the New York City Department of Health has found that transparent plastic trays are a practical and reliable substitute for test tubes in the Kolmer complement fixation test.

The bureau explored the use of plastic trays in an attempt to reduce the costs of syphilis serology without diminishing the quality of its work. Experiments determined that the Kolmer complement fixation test performed in one-half quantity in the wells of the plastic trays matches the results of the test tube procedure.

In a comparative study, 578 serum specimens were tested simultaneously in plastic trays and in test tubes, each test in duplicate. The readings on the two sets of tests were in agreement on 551 of the serums and in disagreement on 27.

Readings in agreement on the duplicate tests performed by each of the two techniques were:

<i>Tube and tray tests</i>	<i>Serums</i>
R, R.....	369
R, WR.....	1
WR, WR.....	3
NR, NR.....	169
A/C, A/C.....	9
Total in agreement...	551

Readings in disagreement on the duplicate tests performed by each of the two techniques were:

<i>Tube test</i>	<i>Tray test</i>	<i>Serums</i>
R, R....	R, A/C.....	2
R, R....	WR, WR.....	7
R, WR..	WR, WR.....	1
NR, NR..	WR, WR.....	6
NR, NR..	WR, NR.....	4
NR, NR..	NR, NR.....	1
R, R....	NR, NR.....	1
WR, WR..	NR, NR.....	1
WR, NR..	NR, NR.....	2
R, R....	A/C, A/C.....	1
NR, NR..	A/C, A/C.....	1
Total in disagreement..		27

The degree of reactivity was determined according to the following scale:

- Reactive (R)=10 to 100 percent complement fixation.
- Weakly reactive (WR)=5 to 10 percent complement fixation.
- Nonreactive (NR)=0 to 5 percent complement fixation.
- A/C=test and control show equal or nearly equal inhibition of hemolysis.

In performing qualitative complement fixation tests in plastic trays, now routine in the bureau's syphilis serology laboratory, serum is put into the wells with a 0.2 ml. pipette,

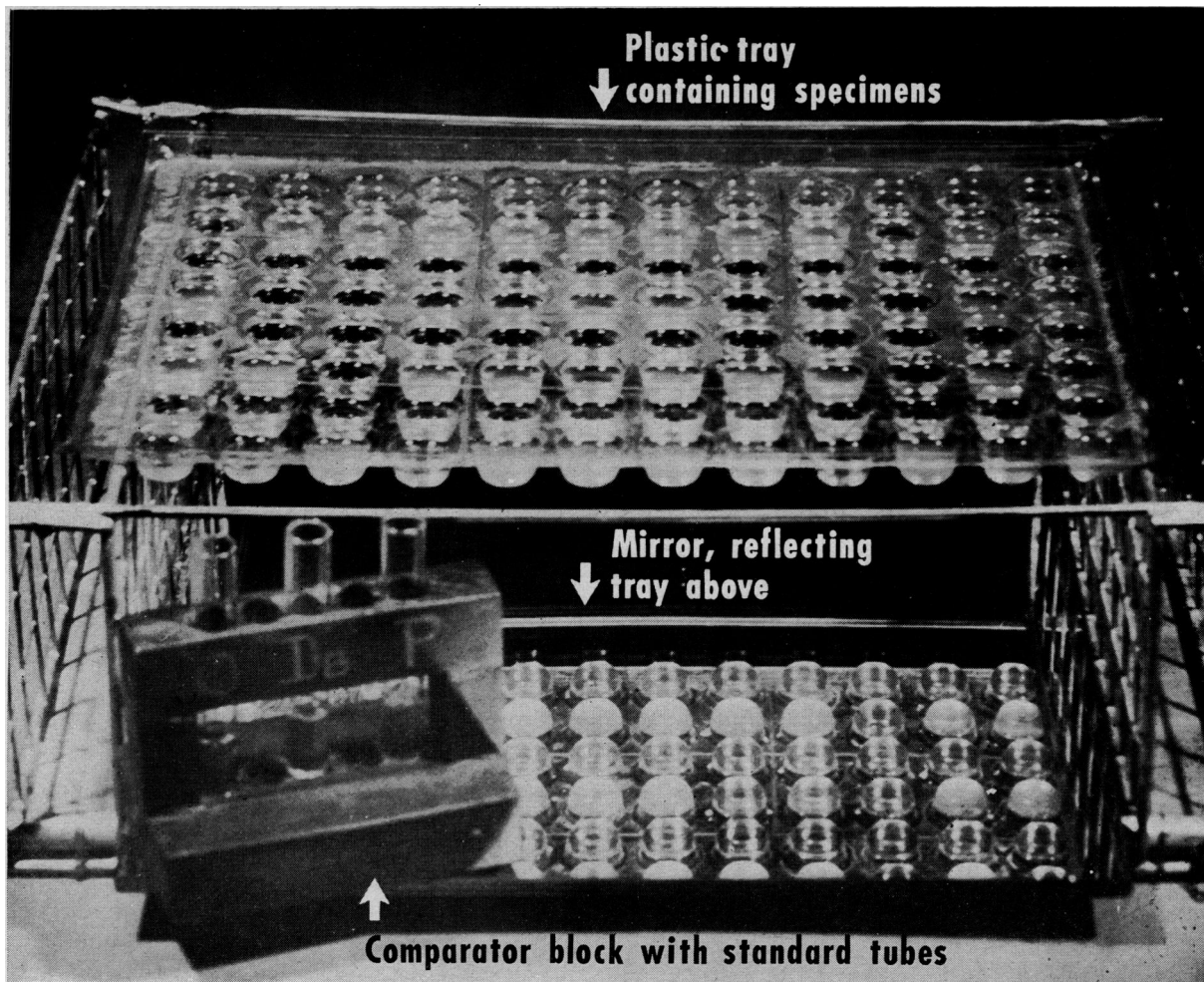
and reagents prepared by Kolmer's procedure are added with automatic pipetting machines calibrated to deliver 0.25 ml. and 0.5 ml.

Reagents and serums are mixed by holding the tip of the pipette sufficiently high so that the force of ejection agitates the contents of the well.

After the trays are filled, they are stacked on top of one another, with sheets of cardboard in between, and refrigerated. The contents are given primary incubation by floating the trays on the surface of a 37° C. waterbath; reagents are added; trays are again placed in the waterbath and are then placed on a rack to be read.

The rack supports the tray about 3½ inches above a plate glass mirror reflecting the bottom of the tray to facilitate reading. A fluorescent bulb desk lamp, placed about 8 inches above the tray, shines directly through it.

Reading controls are prepared according to Kolmer's procedure. Reading standards are made by adding 1.5 ml. of 0, 5, and 10 percent standards to 75- by 12-mm. Kahn tubes. Serums with questionable reaction are transferred to Kahn tubes and compared directly with the standards. The comparator block



facilitates direct comparison (see figure).

When the test is finished, the trays are easily cleaned by flushing them under a tap, rinsing them in distilled water, and air drying.

One thousand specimens, formerly requiring 2,000 Kolmer test tubes, are now tested in 21 trays, each costing 47 cents. Each tray can be used at least 6 times, representing a cost of about 8 cents for 48 specimens.

Use of one tray is equivalent to the handling of 96 test tubes for refrigeration, incubation, and reading. In our laboratory, 500 complement fixation tests were performed by one technician utilizing the trays.

A tray, containing 48 specimens, can be read at a glance. The technician's hands are thus free while reading, enabling him to record results without clerical assistance.

Without sacrificing quantity, the

plastic tray technique permits use of one-half the quantity of reagent needed by the test tube method and effects a substantial reduction in the cost of personnel, refrigerator space, and cleaning facilities.

—DANIEL WIDELOCK, Ph.D., ANNA D. REYNOLDS, JOHN TRUELOVE, and EVELYN V. ENGELKE, *bureau of laboratories, New York City Department of Health.*