The Development and Use of Measurement Models: A Short Feasibility Study Based Upon Four Completed Phase I Reports

ANNEX D

to

THE NATIONAL INSTITUTE'S INFORMATION MACHINE
A Case Study of the
National Evaluation Program

THE URBAN INSTITUTE
2100 M Street, N.W., Washington, D. C. 20037

Prepared under Grant Number 76NI-99-0110 from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice.

9-5070-04

THE DEVELOPMENT AND USE OF MEASUREMENT MODELS: A SHORT FEASIBILITY STUDY BASED UPON FOUR COMPLETED NEP PHASE I REPORTS

TABLE OF CONTENTS

	Page
Preface	v
I. INTRODUCTION	1
II. MODELING TECHNIQUE, CONVENTIONS, AND AN EXAMPLE	4
A. WHAT IS BEING MODELED	4
1. THE INTERVENTION AND ITS ENVIRONMENT LEAD TO MEASUREMENT MODELS	5
2. QUESTIONS, ISSUES, EXPECTATIONS -- LOGIC MODELS	9
B. MODELING CONVENTION	10
1. LOGIC MODEL	11
2. MEASUREMENT MODEL	11
3. EXAMPLES OF THE MODELS	14
III. HOW THIS APPROACH APPLIES TO A PHASE I STUDY	19
A. HOW TO USE THE MODELS IN A PHASE I	19
1. REPRESENTING INTERVENTIONS AND ISSUES	19
2. ORGANIZING MATERIALS AND FILES	20
3. SYNTHESIS OF AVAILABLE INFORMATION	20
4. DEVELOPING TYPOLOGIES FOR THE TOPIC AREA	21
5. DESIGN OF EVALUATION	21
B. GUIDELINE FOR BUILDING AND USING MODELS IN AN NEP PHASE I	22
IV. REPORTING ON ATTEMPTS TO CONSTRUCT BOTH LOGIC AND MEASUREMENT MODELS FROM SELECTED NEP TOPIC AREAS	30

V. SOME FURTHER OBSERVATIONS FROM THE FEASIBILITY EXERCISE 59

A. THEORETICAL ISSUES NEED TO BE REDUCED TO TESTABLE QUESTIONS TO FOCUS THE WORK 59

B. GENERAL MEASUREMENT MODELS COULD BECOME QUITE COMPLEX UNLESS TESTABLE QUESTIONS APPLICABLE TO THE ACTUAL OPERATIONS OF SETS OF SIMILAR PROJECTS ARE ADDRESSED 61

C. PRESENCE AND AVAILABILITY OF SPECIFIC DATA 64

VI. CONCLUSIONS AND RECOMMENDATIONS 66

LIST OF EXHIBITS

Exhibit		Page
1	Measures Which Describe A Direct Intervention	8
2	Measurement Model For Methadone Treatment Program (Type A) In City X	15 & 16
3	Logic Model For Methadone Treatment Program (Type A)	17
4	Building and Using Models in NEP	23
5	Intensive Special Probation Project Logic Flow Diagram (from Phase I report)	34
6	Juvenile Diversion Program Flow Diagram (from Phase I report)	35
7	Intensive Special Probation Simplified Case Flow Diagram for a Project Intervention (constructed from information in the report)	36
8	Example A - Simplified Measurement Model of an Intensive Special Probation (ISP) Project - Client Flow Diagram	38
9	Simplified Measurement Model of an Intensive Special Probation Project - Client Flow Diagram	39
10	Intensive Special Probation Generalized Client Flow Diagram	42
11	General Measurement Flow Model of Juvenile Diversion Within and Outside of JJS, including Measurement Points and States	44
12	Residential Inmate Aftercare Client Flow Model with Some Measurement Points and Measures	45
13	Measurement Model of Pre-Trial Release Program Showing Measurement Points and Measures	46
14	Intensive Special Probation Generalized Client Flow Diagram Showing Sections Answering A Particular Expectation	51
15	General Measurement Flow Model Of Juvenile Diversion Within And Outside Of JJS, Indicating Measurement Points And States For A Particular Testable Logic Model	55

LIST OF TABLES

1	TEST OF STATES AND MEDIA	26
2	TABULATION OF MEASUREMENT POINTS AND MEASURES	41
3	SYNTHESIS OF MEASUREMENT POINTS AND MEASURES OF GENERAL INTEREST AS DERIVED FROM EXHIBITS 5 AND 6	43
4	MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES CONTAINED ON ISP GENERAL MEASUREMENT MODEL	52 & 53
5	MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES CONTAINED ON JUVENILE DIVERSION GENERAL MEASUREMENT FLOW MODEL	57 & 58

Preface

This document reports on a feasibility study into the development of usable measurement models conducted using the full set of products from each of four National Evaluation Program Phase I studies:

Inmate Aftercare (Halfway Houses),
Pretrial Release,
Intensive Special Probation, and
Juvenile Diversion

It presents further development of approaches for collecting, organizing, and analyzing the data during a Phase I grant under the National Evaluation Program.

This feasibility effort was accomplished with approximately four man-months of effort to explore the problems and difficulties of developing examples of measurement models from four existing NEP Phase I study reports. Deciding what to report and how to report it consumed perhaps another three man-months.

The general rules were: (a) an experienced analyst (Elliot Ratner) performed the entire exercise from his desk top using only the LEAA Work Description for an NEP Phase I Study, the completed study products (including site visit material), and our own research guides and (b) speculations upon factual material beyond the scope of the material reported in the studies were excluded.

The conclusion was that structuring such models does appear feasible within limited periods of time, subject to constraints reported herein. Such an effort would, however, be considerably more fruitful during the course of each grant when a wealth of additional material is available to the grantee. Some indications for future LEAA policy considerations are also noted.

The document includes in general form approaches developed during the course of the work, indications of how to use the approaches in a Phase I study, and some of the examples drawn up during the exercise. This material should be of use in the further development of the National Evaluation Program.

I. INTRODUCTION

Under the rubric of the National Evaluation Program (NEP), the National Institute of Law Enforcement and Criminal Justice (NILECJ) of the Law Enforcement Assistance Administration (LEAA) is conducting research and evaluation on groupings of similar operating projects. Each grouping, called a topic area, has been identified for study because of expressed interest throughout the criminal justice system. An NEP study of a topic area is carried out by awarding grants in two distinct and formal phases. Phase I is a grant covering the development of an operational definition of the topic area. Phase II covers an actual field evaluation for those topic areas where it is shown that the implementation of an evaluation is warranted on the merits of both cost and importance of the information/knowledge gaps identified in Phase I.

Using fiscal year 1975 and 1976 funding, 27 Phase I studies have been conducted. Each awarded study was guided by an identical Phase I "work description."[1] The work description set forth and detailed a series of tasks to (a) develop an operational definition of the topic area in terms of how projects work in practice, (b) accumulate existing knowledge about a topic area, (c) identify the "knowledge gaps," and (d) design alternative approaches to fill the gaps identified on the basis of specific criteria.

1. "Work Description for a Phase I Study, under the National Evaluation Program," November 1974, was prepared jointly by NILECJ and by The Urban Institute. A revised work description dated February 1977 has been prepared based on NEP experience to date and is now in use. All of this work is based upon the original work description.

An NEP Phase I study was to consist of the following six deliverable working products:

<u>Product</u>

Product	
1	-- Issue paper drawn from general knowledge and past findings.
2	-- Flow diagrams of existing project intervention activities and accompanying descriptions (not required to be in publishable form).
3	-- A framework(s) developed from the information collected, for use in analyzing existing activities in the topic area.
4	-- An assessment--in terms of measures and comparisons identified on the framework--of what is presently known about interventions in the topic area.
5	-- An evaluation design based upon the framework and all prior knowledge gathered. The design should cover what is not known that could be authoritatively evaluated, how such evaluation(s) would be performed, and probable cost(s).
6	-- The design of a project evaluation in the form useful for a single local project.

The NEP Phase I Work Description called for the presentation of findings in the form of general flow diagrams. For products (2) and (3), a flow diagram was to be produced that displayed points at which measurements could be taken on field projects to answer evaluation questions identified in the issue paper (Product 1). The flow diagram was to be used to organize and synthesize existing data and knowledge.

A review of the Phase I final reports that was conducted in the first quarter of 1976, and subsequent reviews of more recent (final report) submissions, found that the grantees used quite different flow modeling techniques to organize and analyze the information that they had collected. In those cases where flow modeling techniques were used least, the synthesizing and evaluation design products are also weakest. The result is that there may be information available to the NEP grantees, and in their files, that has not been extracted for use by others.

In view of this situation, The Urban Institute, as part of their NEP development and assistance activities, recommended a small task to explore the feasibility of drawing flow models from a few of the working products that have been submitted. The results of that feasibility study are reported in the following five chapters.

- Chapter II: The modeling techniques and conventions used in this study and an example of them.
- Chapter III: A discussion of how the approach developed applies to a Phase I study.
- Chapter IV: The actual attempt to use the approach on NEP Phase I reports.
- Chapter V: Further observations from the feasibility exercise.
- Chapter VI: Conclusions and recommendations.

The descriptive material in Chapter IV was, of course, developed first. Chapters II and III represent attempts to proceduralize some of the work that was done--in the NEP context in Chapter III and more generally in Chapter II. These two chapters represent a more specific expansion of some of the material used in the UI training sessions for NILECJ grant monitors.

II. MODELING TECHNIQUE, CONVENTIONS, AND AN EXAMPLE

Based upon the first two years of experience with the NEP, there is a need for:

- a technique for describing the issues or questions people are interested in having an evaluation answer,
- a technique for determining which questions can be (or have been) answered with existing data, and
- a technique for describing what questions can be answered using measurements from operating projects.

Obviously, the three techniques must be compatible (i.e., have some common language and format). Here we describe a modeling approach that may meet these needs. It can be followed by grantees in doing a Phase I and by NILECJ in reanalyzing data collected and reported in completed Phase I's.

The first section describes what is being modeled--the intervention being made by topic area projects, as captured by a measurement model, and the beliefs and expectations for those projects as captured by logic models. The second section describes the modeling convention that will be followed. The third section provides an overview of how the models are used in organizing data, synthesizing information and carrying out the analyses. The last section describes a process for constructing measurement and logic models from Phase I products.

A. WHAT IS BEING MODELED

In the end we need to model (1) the portions needed of the Topic Area projects that represent the measurable intervention being made into society and (2) the expections, questions or issues people are interested in raising about the intervention. This section describes both.

1. <u>THE INTERVENTION AND ITS ENVIRONMENT LEAD TO MEASUREMENT MODELS</u>

The NEP requires grantees to examine and describe the projects in a topic area as they operate in the field. The description is to cover the project characteristics and effects of interest to the national and local criminal justice decision makers. The flow model used to describe a related set of measurements from a project or projects is called a <u>measurement model</u>. The <u>measurement model</u> is a flow model representation of related measurements from the operating project(s) which defines the project operations in terms of input/outcome, process and impact measures. The flow model identifies points in the operating projects and its environment where measurements of interest to an evaluation can be or are being taken and important interrelationships between these points. In what follows we provide a general description of the types of measures needed to describe a project. The scheme was used to guide the NEP flow modeling. Section III.B. describes how the grantee, guided by the logic model (II.B below), constructs measurement models from the detailed project site descriptions.[1]

An NEP topic area consists of projects distributed nationally that are intervening in the criminal justice system or society for some purpose. These <u>interventions</u> can be described in terms of measures defined on the project and measures defined on the projects environment. We will distinguish here the four types of measures:

<u>process</u> - measures describing the project activity and operations.

<u>input and outcome</u> - measures defined on that part of the environment that goes through the project or is directly affected by it.

<u>impact</u> - measures defined on that part of the environment which the project may affect but does not directly serve or treat.

---

1. The "detailed project descriptions" might be referred to as project "equivalency models." The terms used in this paper are adaptations for the NEP of a basic approach to analysis described in more detail in Urban Institute Working Paper 783-0[illegible], "Representation of Reality: Measurement Models in Evaluation," by Joe N. Nay, John D. Waller, John W. Scanlon and Peg Kay, and Working Paper 783-09, "Evaluability Assessment," by Joe N. Nay and Peg Kay.

An illustration of the relationship between the intervention activities of the project and the environmental setting in which it operates is shown in Exhibit 1. The outer cloud-shaped boundary encloses that portion of society that is expected to influence or be influenced by the intervention. The smaller or inner cloud shaped boundary encloses the immediate environment of the intervention being made. The intervention in shown sitting in its "immediate" environment. Flows are shown as dark arrows, their tails being the source and their heads being their destination. Different measurement points are shown for the three types of measures.

The interventions themselves--training people, treating people, fostering improved relationships among people--are often processes and their measurements are called "process measures." There are specific "inputs to the process" of the intervention, such as funding, manpower, guidelines and policies, that are needed in order for the intervention to be established as a process on a continuing basis. They flow into the project from some environment (not shown in Exhibit 1).

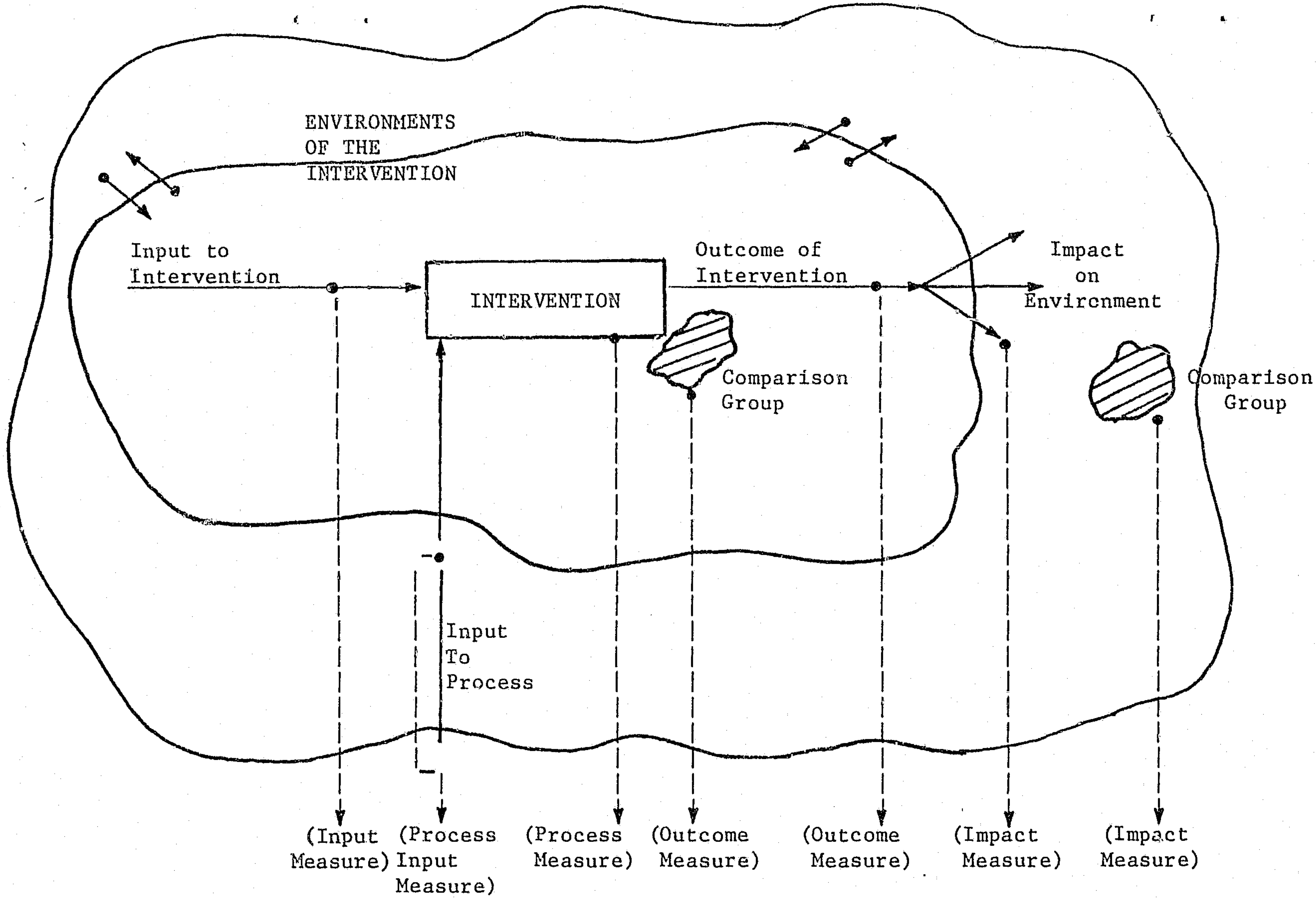
The intervention providing services will also receive a flow from the environment which it will treat or serve or process or regulate. These flows, referred to as "input to the intervention," might be the unemployed worker, the criminal adult inmate referred to a halfway house, or the individual arrested for a criminal offense who after adjudication is released to an intensive special probation program. They will be treated as though they appeared at the project, received the services provided and went on.

The input to the intervention will flow through and out of the intervention. The intervention will have some immediate effect on some characteristic of the flow, e.g., people are trained to acquire or enhance some sort of skill, people are given therapeutic services to become healthier or sicker, runaways are housed or returned home. Measures of these direct effects on the immediate environment are called "outcome measures."

Beyond the immediate "outcomes" are the more distant results of the intervention process. As Exhibit 1 illustrates, all of this intervention activity takes place within a wider universe that represents that part of society that is linked to the intervention through the immediate environment. For instance, the training of people to acquire a skill, such as welding, may have some effect on the overall employment situation of the area. The helping of ex-convicts to reintegrate into their community, an "outcome," may affect that society by reducing welfare payments to families and improving the stability of families. It may even eventually reduce crime. These results "further from the process" are generally called "impact measures." These measures usually have to be taken after some time has elapsed from the completion of the intervention activity itself. Control and/or comparison data are often taken also from similar or at least "nearby" parts of the larger society to determine whether the expectation or impact observed was contributed by the intervention or by something else.

In addition, it may be necessary to measure various societal factors that react with the intervention process and modify, enhance, or constrain its operation. Such factors may include the fact that the intervention made to train people to become welders is taking place in an environment where there is chronic unemployment or jobs for this skill category are controlled by labor unions. These interactions across the environment boundaries are shown as flows by the arrows in Exhibit 1.

Many things can be modeled from any project, either by equivalency models or measurement models. What to model must often be reduced (to make the work practical) by deciding what questions either are to be answered (or might easily be answered). This is considered in the next section.



KEY:

———►: Flows of some medium (people, cases, funds ...) in the environment and/or in the intervention.

---►(   ): Points of measurement and measurement type.

EXHIBIT 1: Measures Which Describe A Direct Intervention

2. QUESTIONS, ISSUES, EXPECTATIONS -- LOGIC MODELS

When a topic area is discussed by policy makers, practitioners and the literature, a lot of different types of expectations and information requests will arise. Different groups of people will want to know different things. A legislative body may only want to know if it works, practitioners may want to know how to make it work. One task in evaluation design (and in the NEP) is to decide which questions are being asked, which are answerable and, finally, in which to spend resources to obtain further answers.

The flow models used to describe the questions an evaluation is being asked to answer are called Logic Models. A logic model is a flow model representation of the questions, issues or expectations an evaluation audience has for a topic area and which are presented, to the extent possible, in terms of the input/outcome, process and impact measurements they would want to see taken to resolve the questions and issues identified to their satisfaction.

In the following sections we will describe the format for logic models. Here we discuss three important considerations in developing such models: the data base for constructing models, the different audiences, and the degree of measurability.

The data for constructing logic models will come in different forms. For example, they may be given as:

(1) expectations - team policing will reduce local crime by 10 percent.

(2) issues - should we use team policing or conventional patrol?

(3) questions - what effect does team policing have on officer productivity?

(4) assumptions - team policing uses group planning and that will increase officer morale which in turn will increase productivity.

(5) guidelines - team policing will have 16 components: (1) officers assigned to teams from 20-30 in size. (2) ...

In developing a logic model we translate such data into a sequence of events whose occurence or non-occurence can be measured from the actual interventions.

As suggested earlier, the interests and perceptions of different evaluation audiences may lead to different logic models - different in content and in level of detail. It may be necessry to group logic models by type of audience and at some point set priorities on which models are to be used.

Finally, there is the question of measurability. How specific must an evaluation audience be in specifying the events in their logic model? The answer is--as specific as possible. The audience can only be as specific as its experience and job allow. We would like to see an audience able to specify the measures and comparisons it would want used to answer a question. Frequently, the audience will not be able to go that far. In such cases the grantee (evaluation designer) will have to either make a judgment decision on what are sufficient measures and comparisons or develop alternatives and work further with the audience to reach a decision. When a logic model is developed to a point where the events are measurable and the comparisons required to test assumptions are specified, the model is referred to as a "testable logic model." The logic model to some degree controls how much of the actual intervention and its environment must be described and modeled. The measurement model represents sets of measurements from the intervention and its environment that are needed to answer testable questions.

B. MODELING CONVENTION

In the modeling work in Chapter IV, we tried to use a certain convention. Two types of modeling formats were used; one representing a sequence of logic, the other sets of measurements related by a flow through various states.

1. <u>LOGIC MODEL</u>

A logic flow model is written as -

A ⟶ B

"<u>if</u> A occurs <u>then</u> B occurs"

The boxes are events that can be observed through measurement. The arrow represents an assumption that A causes B or that A is a necessary precondition to B occurring.

Logic models are used since questions, expectations, and issues can be conveniently described as the occurrence or non-occurrence of events. Depending on the situation the events may be stated very generally -

Team Policing is Implemented ⟶ Crime is Reduced

or in measurable terms -

Team Policing as defined by LEAA manual is implemented in two areas ⟶ Part I UCR crimes go down in team areas 20% more than in control areas

2. <u>MEASUREMENT MODEL</u>

A measurement model is written as:

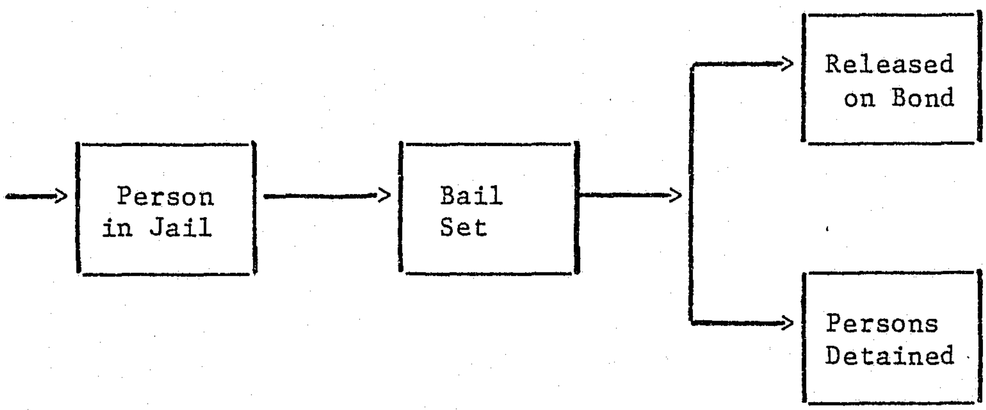
⟶ Number in S_1 ⟶ Number in S_2

The arrow represents some flow - people, funds, cases, etc.--through various states. The boxes are "states" the flow finds itself while in the program or in the environment--for example "employed," "unemployed" are examples of "states" that a person in the labor force can be in. Thus, both the environment and the intervention are described in terms of "states."

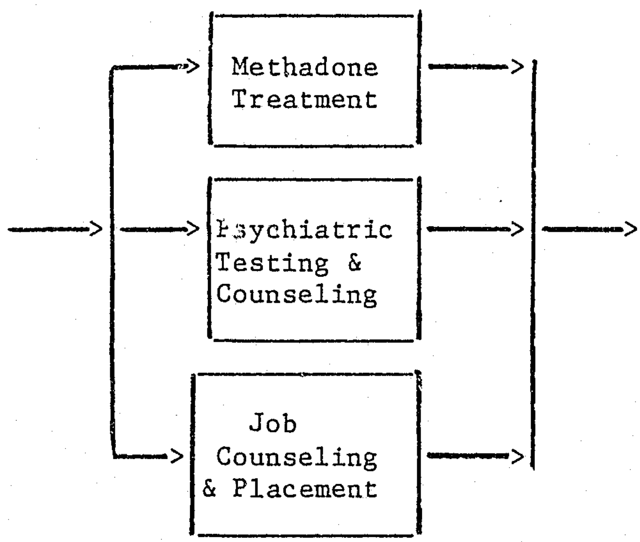
The models constructed should have the following properties.

> The model must capture all possible states of the flow. A simple example is shown to illustrate this point. "People" were selected as the appropriate flow for the manpower training program. The following states were considered for such a model:
>
> - Not yet born.
> - Not in the Labor Market.
> - Employed.
> - Unemployed.
> - Deceased.

The model must show what transitions among states are possible. Thus, both the possible states and transitions among states are specified for a flow. If more than one type of flow (e.g., funds and people) is considered for incorporation in a given model, then each of the flows must have its own set of states and transitions. For this paper when states are mutually exclusive--e.g., a person can only be in one at a given time--they will be drawn in series or as branches -



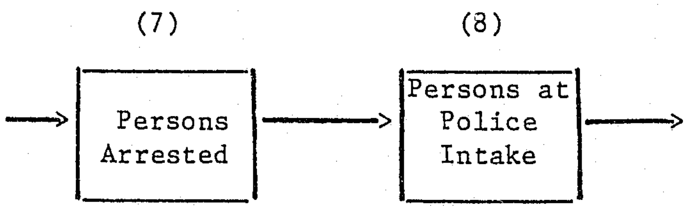
When a person can be in several states at one time, the states will be shown in parallel. For example while in a drug treatment program a client may receive several types of services simultaneously:



The measurement model will be used to identify a measurement point. In general, besides a description of the "state" itself, one can measure,

- rates in and out of a "state,"
- the level of the state,
- average time the media stays in a state.

<u>In order to keep the exercise in Chapter IV simple, we only refer to measurement of the level of each state and only use one of the media</u>. Therefore each box will be called a measurement point and will be noted by a number referencing an accompanying set of definitions of exemplar measurements.



(7) number of people arrested for criminal offense
(8) number of people going through intake procedures

The measurement models can be used to present summary descriptions of the intervention to those whose questions are to be answered. Each environmental state and in-program state identified can be supported by descriptions, plans, manuals... describing the "state" as planned or as it exists.

3. EXAMPLES OF THE MODELS

Examples of portions of the logic models and measurement models are given in Exhibits 2 and 3 for a hypothetical methadone treatment program. These figures show one way to present the models. Other approaches could be used; however the important point to note is the level of information presented in the models. Some flows have been omitted for clarity.

Exhibit 2 (and its supporting table) is a measurement model that, if this were not an example, would represent measurement points from a real project in a real city. This figure is linked by the flow of people through the program--i.e., process states--and through the environment of the progam--i.e., input, outcome and impact states. In the table accompanying Exhibit 2 each state is defined, the measures of interest identified and the current availability of a measurement and data collection system noted. We have shown for this example, four states (1a, 1b, 10 and 11) where there is no measurement/data collection system in place. Many more arrows and measurement points could have been added to more completely describe the project.

Exhibit 3 is a testable logic model developed from the literature. The first row describes how the program is expected to work and what it is expected to result in. The second row gives the corresponding questions raised about the program in the literature.

The first two and last three questions ask for descriptions of how certain events of interest occur. Thus they are asking for monitoring type information on the environment and on the program. The third question asks for a test of one of the linking assumptions--does the treatment cause the desired effect. Here we require a research design, a more intensive type of evaluation than monitoring.

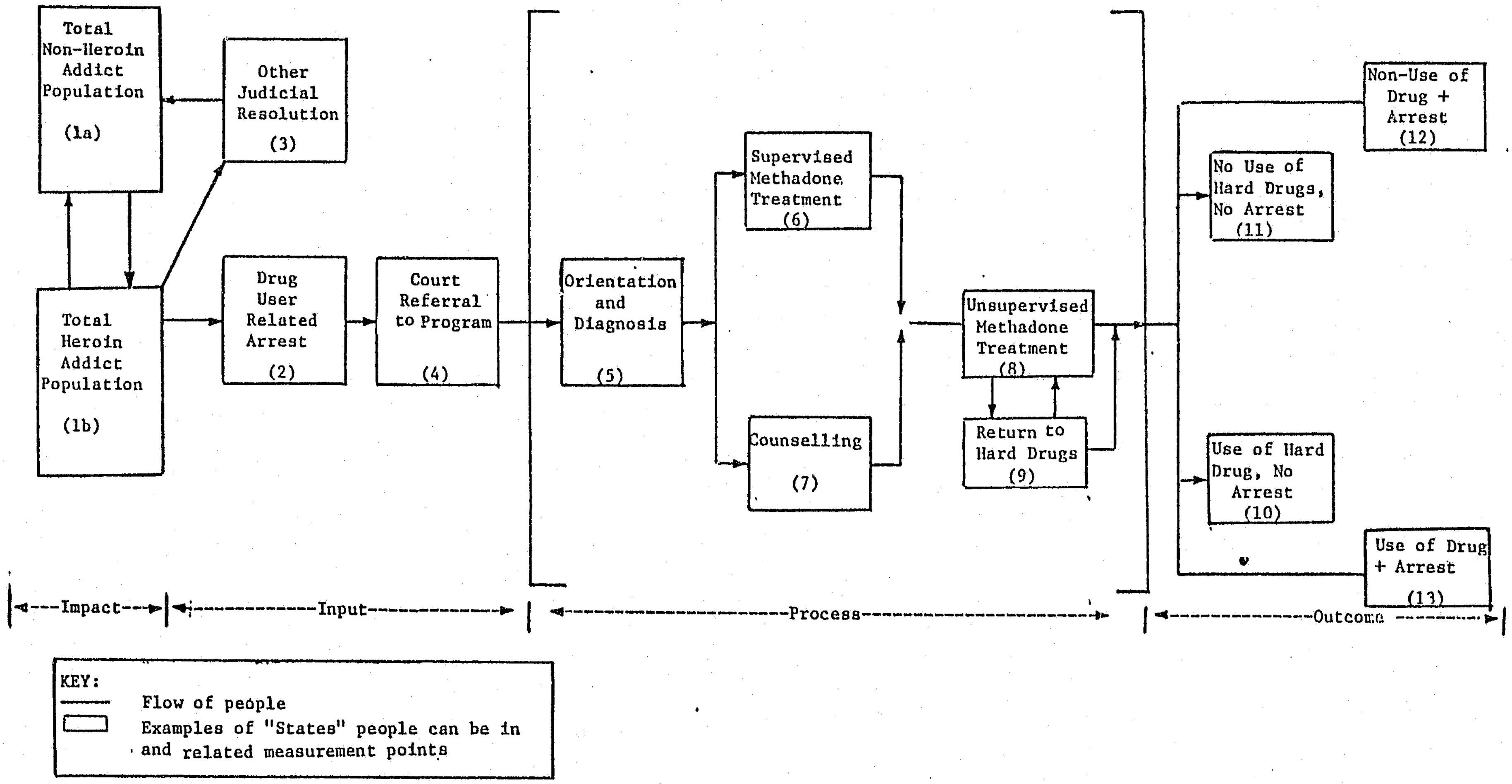


EXHIBIT 2: Measurement Model For Methadone Treatment Program (Type A) In City X

EXHIBIT 2 (Continued)

Measurement Point	Definition of The State	Measures of Interest	Measurement and Data Collection System Available
1a,b	Total populations addicted and not addicted to heroin who were never in program	number of people	No
2	Addicts arrested for any offense	number	Yes
3	Arrested addicts who are not in program while under jurisdiction of court	number, rates to and from other states	Yes
4	Arrested addicts referred to to program by court	number	Yes
5	Arrested addicts in orientation and diagnosis component	number, flows to other states	Yes
6	Clients in supervised mathadone treatment	number, time in component	Yes
7	Clients in counseling component	number, time in component	Yes
8	Clients in unsupervised methadone treatment component	number, time in component, rates	Yes
9	Clients in program who return to using "hard" drugs	number, time on hard drugs	Yes
10,11	Ex-clients using or not using hard drugs after the program who are never arrested	number, time in each state	No
12	Ex-clients arrested but not using hard drugs	number	Yes
13	Ex-clients arrested and using hard drugs	number	Yes

Events and Assumptions	Set up a Type A Methadone Treatment Program (Capacity = 100 Supervised Clients)	Client Population will Accept and Respond to Methadone Treatment		Client Population will not Return Illegal Hard Drugs After Leaving Program	Client Popula- will not Commit crimes to Support an Illegal Habit	Community Addict Population Will Go Down Community Drug Use Related Crime Rate Will Go Down
Evaluation Questions	1. Was the program set up as by Type A guidelines?	2. Did the client population accept and respond to treatment?	3. Did Methadone program Type A significantly reduce the addiction rate of people treated?	4. Did the client population return to drugs after the program?	5. Did the client population commit crimes after the program?	6. Did the community addiction rate & drug use related crime rate go down?
Measurements and Comprisons Desired to Answer Evaluation Question	2,4,5,6,7,8 (Compare site descriptions of state and levels guideline specification for Type A projects.)	3,5,6,7,8,9 (Compare flows among states 3,5,6,7 and between states 8 &9.)	8,9,10,11,1a,1b (Compare addiction rate of client population after program with addict rate of control group.)	8,9,10,11 (Measure flows between states 8, and 9,10, 11.)	10,11,12,13 (Measure states at month intervals for 1 year after program.)	1a, 1b, 2,12,13 (Measure states 1a, 1b, 2,12,13 over a three year period.)

KEY:
- ▭ Event
- ——> Linking Assumption
- (#) Measurement Point on Figure 2

EXHIBIT 3: Logic Model For Methadone Treatment Program (Type A)

The last row presents the measurements and comparisons the literature seems to call for and the evaluation designer might want to answer these questions. The figure 2 measurement points are used to identify the measures desired. Note that the third, fifth and sixth questions cannot be answered with the measurement and data collection systems currently in place. The necessary measurement systems would have to be developed and/or installed, at some cost before these questions could be answered.

III. HOW THIS APPROACH APPLIES TO A PHASE I STUDY

The material above presents a general background to the approach applied during the feasibility study. This chapter focuses in more detail on how the approach developed might apply to a Phase I study based on our experience. It is given here to further explicate how the work in Chapter IV was performed and for further use in the continued NEP development of efforts.

A. HOW TO USE THE MODELS IN A PHASE I

The models have five uses in the NEP Phase I: representing interventions and issues in a summary manner, organizing material and files, synthesizing information, developing typologies and designing evaluations.

1. REPRESENTING INTERVENTIONS AND ISSUES

Examination of a project or the current issues in a topic area can generate a great deal of information at different levels of detail. The flow models present a useful device for displaying the significant components of an intervention or issue and the relationship among components. It organizes and codes information at a level of detail that one can work with and still have the essential dimensions of the task covered for the purposes at hand.

In an NEP Phase I study, "flow diagrams of existing project interventions" were called for (product 2). These flow diagrams would generally be functional diagrams equivalent to the project functions rather than measurement models.[1] Similarly flow diagrams could be used to present the product (1): Issue Paper, and product (3): Framework.

1. Working Papers 783-05 and 783-09, op. cit.

2. ORGANIZING MATERIALS AND FILES

The data collection for an NEP is quite extensive: available knowledge on the topic area is collected, expectations and questions on the intervention from various sources are collected, descriptions and data on projects are collected. The flow models can be used to organize and code such information.

In Section B above the models were coded to refer to definitions of measures. Such coding can be extended to use the models as a map to important files, for example:

7.4.6d

—> | Person in Deferred Prosecution Program | —>

where

7 - refers to a definition for the measurement point.

4 - refers to a file which contains a description of this program "state."

6d- refers to a file which contains the data systems used by this project that cover this measurement point.

3. SYNTHESIS OF AVAILABLE INFORMATION

The measurement flow models can be "exercised" by using available data and knowledge to estimate such factors as:

- o rates in and out of a state,
- o levels for various states, or average time in a state,
- o and, dependencies among states.

By identifying key states and flows the models enable one to systematically collect, organize and map available information on to the models.

4. DEVELOPING TYPOLOGIES FOR THE TOPIC AREA

Since the flow models themselves represent information (on actual or expected project structure) they can be used in comparative analyses. For example, one can use all the project models to examine patterns and identify if there are distinct sets of projects within a topic area. For example, there may be five types of Youth Service Bureaus distinguished by different intervention activities (represented by different sets of "states").

5. DESIGN OF EVALUATION

The information in the logic models and measurement models can be used to guide evaluation design work. Ideally a complete evaluation design would consist of:

- a statement of the questions to be answered,
- the measures and comparisons (evidence) to be used to answer questions,
- a measurement and data collection plan,
- measurement instruments and/or data collection instruments,
- analysis plan (processing data and making necessary comparisons),
- plan for exercising quality control over measurement, data collection and analyses,
- formats for presenting evidence (results),
- estimated budget for measurement, data collection, analyses, and dissemination,
- estimated quality of the data and confidence one can have in the evidence.

With the logic models we have a representation of the questions people want answered in the form of expected events and the measures used to describe these events. With the measurement models we have a representation of the measurements that can be taken off operating projects and their environment; a representation which covers those measures applicable to the questions of interest while preserving important states and relationships that exist in the projects and environment.

By comparing the logical model and measurement model we can identify the type of measures and comparisons that can be made available, at some cost, to answer specific questions. Moreover to the extent that the models have been used to "organize material and files," "synthesize available information" and "develop typologies" we will have additional information to answer such evaluation design questions as:

- Where are these gaps in our current state of knowledge that an evaluation can fill?
- What types and quality of measurement systems are already in place?
- How many new measurement systems will have to be developed?
- What magnitude of rates, levels, effects, changes, are we trying to detect?

B. GUIDELINE FOR BUILDING AND USING MODELS IN AN NEP PHASE I

Based on our efforts to create logic and measurement models from the Phase I reports, we developed a sequence of steps for our own use that grantees of NILECJ might follow in building and using these models, a more detailed application of Chapter II.

The process consists of steps taken to construct the models and steps taken to use the models. It is assumed that all the NEP data collection has taken place and the data are available.[1] Exhibit 4 illustrates the process. On the left the data base created by an NEP is shown. It consists of:

1. When doing an NEP Phase I the grantee would carry out data collection and the model building steps simultaneously. Obviously, some of the model building steps would help guide data collection. The order in which that should or could be done is left up to each grantee.

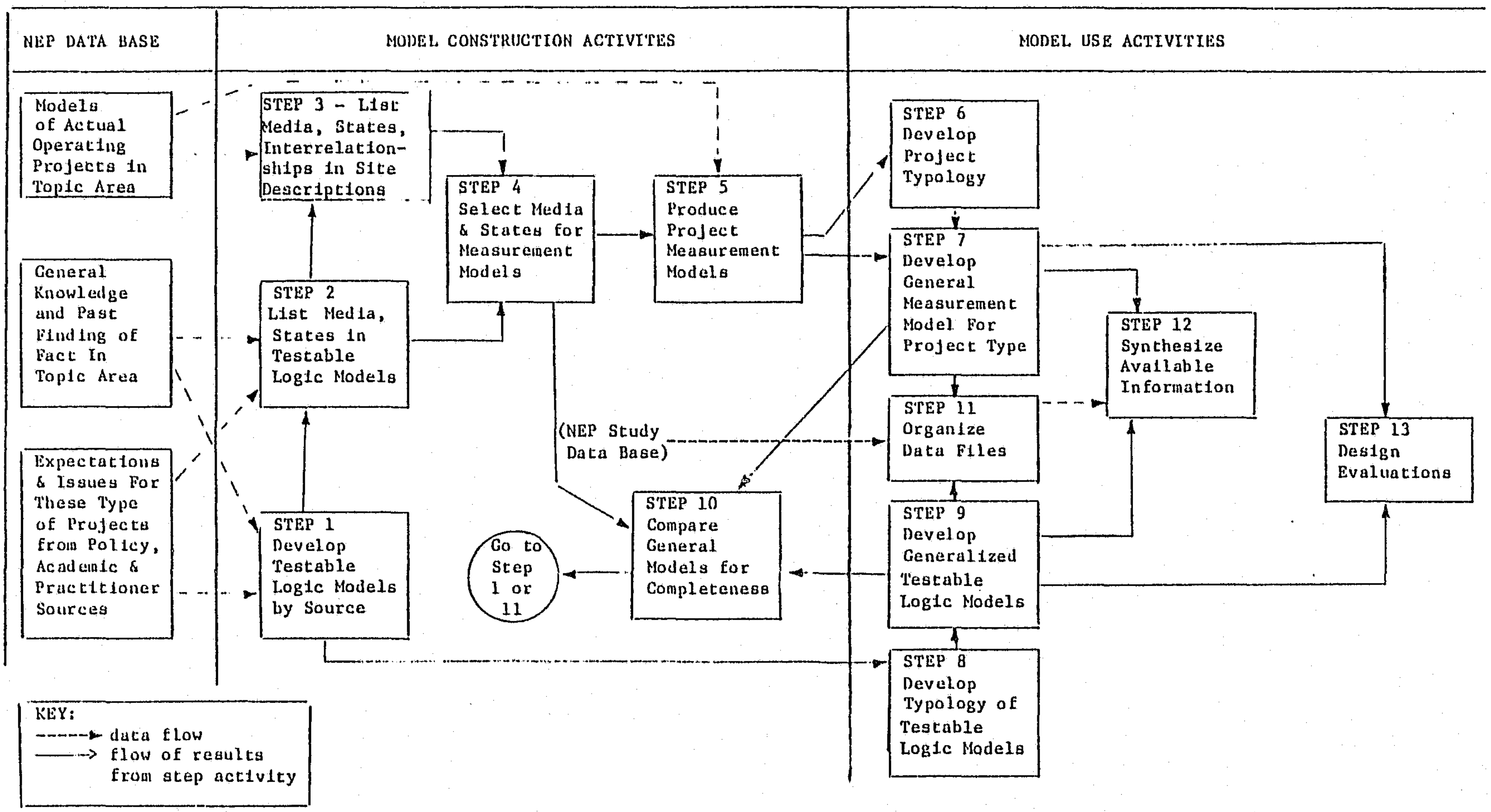


EXHIBIT 4: Building And Using Models in NEP

- models of field projects surveyed and visited (NEP Product 2),
- general knowledge in the topic area (NEP Product 1),
- past findings of fact (NEP Product 1),
- expectations for these type(s) of projects from policy, academic, and practitioner sources (NEP Product 1).

In the center are 6 steps defined as the basic model building activity.

Step 1 - Develop Testable Logic Model For Each Source

Step 2 - List Flows And States Referred To In Testable Logic Models

Step 3 - List Flows And States Identified In Site Descriptions

Step 4 - Select Key Flows And States For Measurement Models

Step 5 - Produce Project Measurement Models

Step 10- Compare Generalized Testable Logic And Measurement Models For Completeness

On the right are seven steps using the models to produce NEP type products. These uses are versions of the five listed above in Section C: developing representations, organizing files, synthesizing available information, developing typologies and evaluation design.

Step 6 - Develop Typology Of Project Types

Step 7 - Develop General Measurement Model For Each Project Type

Step 8 - Develop Typology Of Testable Logic Models

Step 9 - Develop General Testable Logic Model For Each Type

Step 11- Organize Data File To Support Generalized Models

Step 12- Synthesize Available Information

Step 13- Design Evaluations.

The results of these steps would be included in NEP Products 3, 4, 5, and 6 - i.e., framework, assessment of what is known, and evaluation designs.

Each of the 13 steps is briefly described here.

Step 1: Develop Testable Logic Models By Source

This step consists of extractions from the NEP data base descriptions of issues, questions and expectations and translating them into event statements in a testable logic flow.

Step 2: List Types Of Flows And States Identified In Testable Logic Models

Using the data bases and results of Step 1, start Step 2 by listing all the flows identified. Then for each type of flow list the various "states" and measures identified. Organize the states under the heading -

impact environments
input/outcome environments
process

Thus one would end up with a chart such as Table 1 filled in.

Step 3: List Types Of Flows And States In The Project Site Models and Descriptions

Step 3 uses the results of Step 2 as a guide and extracts from the project descriptions a list of flows and states identified during site visits. For each project then, a table similar to Table 1 is produced.

Step 4: Select Flows And States For Measurement Models

Using the results from Step 2 and 3 (i.e., the complete set of Flow/State Lists), the investigator decides which flows to use in constructing measurement models. The flows selected should be on the basis of the following criteria:

- A small number of flows (typically one or two) is sufficient to describe the program.
- All measurements of interest can be obtained from the flow.
- The definition of the flows is simple.
- The flows are easily traced in the system to be modeled.

Once the flows are selected, the "states" to be used to describe the environmental and intervention sectors should be selected so that questions developed in the testable logic models can be answered.

TABLE 1

LIST OF STATES AND MEDIA

For Flow #1	States Identified	Measures Identified
Impact Environments 1 2 3 . . .		
Input/Outcome Environments 1 2 3 . . .		
Process (Intervention) 1 2 3 . . .		

For Flow #2	States Identified	Measures Identified
Impact Environments 1 2 3 . . .		

Step 5: Produce Project Measurement Models

Using the decisions made in Step 4 and the project data base created by the NEP, construct measurement models for each project. The flow diagrams created should specify a logical starting point for the flow used. This point might be where the flow interfaces with the project or where the flow interfaces with the criminal justice system. The point of departure or model initialization will depend on the "environment" of the program/project. Based on the starting point selected, list sequentially states of the flow as they are affected by the various activities which take place in the environment of the project or under its control, as well as outside of the program. The listing of states should be generalized and specified in terms of measurement points and measures (e.g., number of people arrested for a particular criminal offense or number of clients referred to a residential inmate aftercare program).

The outcomes and expectations of the program intervention activities should be shown as the final states of the flow in the diagrams constructed.

The boundaries separating the intervention from the input/outcome environment and the impact environment should be clearly identified.

Step 6: Develop Project Typology

Using the project measurement models developed from Step 5, examine the array of projects and note the similarities and differences in the flows and states. The variations, noted in the various dimensions may result in different project flow diagram configurations in terms of measurement points and measures. It may be useful at this point to distinguish projects by creating a typology.

Step 7: Develop General Measurement Model For Each Project Type

Upon completion of Step 6, create a generalized measurement model for each project type which covers all individual projects in that group.

To accomplish this task, construct a table listing each and every measurement point and measure specified for each project of a given type. Review the array of information with that of Step 6 and, wherever appropriate, generalize the measurements. The product of this effort should be used to create the generalized flow diagrams.

Step 8: Develop Typology Of Testable Logic Models

Take the Testable Logic Models from Step 1 and review them to decide if they can be combined or if a natural pattern emerges suggesting another typology. For example one may find different sources have drastically different views of what project objectives are. It would be misleading to combine these into one model.

Step 9: Develop A Generalized Testable Logic Model For Each Type

For each of the "types" identified in Step 8, combine the individual testable logic models into an all inclusive testable logic model.

Step 10: Examine And Compare General Models For Completeness

The generalized flow diagrams, together with the product of Step 4, should be examined now to ensure that the important states and flow have been considered appropriately in the "general models." That is to say, at this stage of the activity the main emphasis has focused on projects which were actually site visited. Now there is a need to return to the full complement of projects falling within the topic area (i.e., results of Step 4) to see whether any important measurements were not represented in the final models created. If such measurements are omitted, the model should be modified, as applicable and desirable. It may be that some of the other projects fall outside of the scope of the models.

Step 11: Organize NEP Data Files

Once the general models are available they can be used to extract important information from the NEP data base, to code it and to file it. For example, one could create the following files:

- measurement systems used by each project organized by measurement points in the measurement model,
- data and information on project process, input/outcome, and impact organized by states and flows in the measurement models.

Step 12: Synthesizing Available Information

If enough data are available, the measurement models can be exercised to ascertain whether or not the hypothesized effects listed in the logic models are likely. If sufficient data are available to "test" the measurement model created, the measurements contained therein should represent the types of information (i.e., measures) that could be captured in operational project settings.

Step 13: Design Evaluations

With the 2 sets of general models, Step 7 and Step 9, and the Step 12 Synthesis, one now has a sufficient data base to begin designing evaluations which will improve the current state of knowledge (as defined by the results of Step 7 and 12) and provide information of interest to policy makers, practitioners and researchers (as defined by Step 9).

IV. REPORTING ON ATTEMPTS TO CONSTRUCT BOTH LOGIC AND MEASUREMENT MODELS FROM SELECTED NEP TOPIC AREAS

The subsequent discussion centers on the actual feasibility study undertaking and illustrates some of the points made in Chapters II and III above.

Based on review of several NEP Phase I studies, the concensus has been that they did not generally produce the most useful types of measurement models although much of the information for these models was collected and in many cases formally reported. Since there was a valid basis for the original requirement, an exploratory effort was initiated to investigate the feasibility of using the NEP Phase I work products (final products) as the source material for creating such models.

No distinct methodology was specified in the work description for the construction of measurement models. In this feasibility effort, successive attempts were made to establish more detailed description that could be followed for the remaining and future NEP's or any other comparable program. Each attempt made provided additional insight into the problems and increased our understanding of what was needed to create such models.

It is particularly important to arrive at specific testable logic models in some way or else the analyst does not know how much detail to develop or when to stop. The NEP Product 1 work was substituted for expectations of decision-makers during the design of the NEP. For this to work out in subsequent steps of a Phase I, quite detailed efforts in developing testable logic are often necessary.

In analyzing the four topic reports, the first step we took was to identify and detail the concerns and/or interest that an audience consisting of policy makers, practitioners and researchers voiced about the program. Such information pertaining to each audience level was extracted from the NEP source materials. The information was extracted for the purpose of creating testable logic models from rhetorical descriptions (or expectations) so that the logic could be used in structuring measurement models.

Our general theory is that development of logic models requires management and evaluator interaction in order to gain acceptance of the questions (measures and comparisons) that are to be answered. This feasibility effort was limited to the source materials. Without iteration and limited to the written reports, the investigation does not produce models for all questions that might be raised.

Some testable logic models could be drawn from each report. To gather such information, we developed a framework to assist in organizing the rhetorical information extracted from the source material. This framework (shown previously in Step 2, Chapter III.B.) was developed in order to indicate the types of relevant questions that measurement models might be constructed to answer. On the other hand, it also reveals in a logical way the information that sources believe is known about the program. During the exploratory study, some of the following questions were posed in the course of extracting information for rhetorical descriptions of possible media.

- In what environment does the "program" operate?
- Where does the authority for the 'program' operate?
- What government agency/agencies provide resources (i.e., funding support) for the program?
- What manpower resources are used?
- What formal and/or informal relationships exist between the program and other government agencies?
- What program intervention activities are used; what resources are employed for each intervention; are both program and outside (i.e., community) resources employed for the various intervention activities?
- How are clients acquired for the program?
- What types of clients are seen by the program?
- What data reporting systems are used and for what purpose?
- What measurements are taken and for what purpose?
- What outcomes are anticipated from the implementation of the program?
- What issues are addressed in the program?
- What policies have been established and how are they implemented?

We found that by assembling the extracted information from the source material in this format that the array resulting provided a viable way not only to capture the information on the universe of projects falling within the topic area, but also to refine what types of measurement models would need to be created.

The next step in the process examined the reports pertaining to the projects which were site visited by grantees. These visits to actual projects were conducted in an attempt to anchor the information collected on the "universe" of projects surveyed in the topic area. Using these site reports as the point of departure for modeling the program, information was extracted and assembled in the format described previously above (i.e., list of states and media). This proved to be a convenient way to organize information about the particular projects studied. Again, when actual individual projects were examined, we looked for and extracted specific information concerning each project that detailed how the intervention activities were accomplished. Some of the following questions were posed to assist in capturing the apparent appropriate measurable characteristics of each project:

- What operating environment is the project in?
- What is the project expected to accomplish?
- What authority is vested in the project?
- What resources (i.e., funding, manpower, other) does the project use to operate?
- How are the resources acquired?
- What formal and/or informal relationships exist between the project and other governmental and community agencies; how do these relationships effect its operations?
- How and from where are clients acquired; What types of clients are seen?
- What intervention activities are performed and in what sequence are they performed?
- Are particular staffing patterns employed for the intervention activities and how are they operated?
- What decisions are made by the project staff concerning client achievement of goals and how are such decisions made and/or are such decisions made by others outside the project or in some other way?
- What data reporting systems are employed, how and when are they used?
- What measurements are taken, when, by whom, how often?

We found that some of the actual individual projects reported on contained diagrams illustrating either how resources were used to achieve project expectations or how the projects related to the local environment of the criminal justice system. Such diagrams proved helpful in providing insights to the develoment of the logic models and to the identification of the media flowing in the project and program. Exhibits 5 and 6 are typical examples of the diagrams already contained in the Intensive Special Probation (ISP) and Juvenile Diversion studies.

Based upon the information extracted from the particular NEP Phase I topic areas employed for this feasibility study, a decision was made concerning the media of interest. In each case, <u>people or cases</u> were selected to represent the unifying concept. Other candidates such as dollars, personnel or staff of the program and especially information and/or data were considered and might be necessary for a full development. People or cases seemed to be the dominant connecting media and sufficient for this illustration. The unifying concept (i.e., people or cases) selected had the characteristics that (1) it is neither created nor destroyed during the process studied, although it may be transformed or transferred from one state to another, and (2) the flow of the unifying concept revealed something about the questions being asked pertaining to the "programs" expectations (testable logic). For instance, if an evaluator wanted to know if the intervention activities of the project or program "worked" it is likely that the selection of people or cases as the "unifying concept" would prove fruitful.

Upon making the determination of the unifying concept for each project examined within a given program, flow diagrams were created that traced the flow of people or cases in the individual projects studied along with the larger environment related to these projects. The flow diagrams developed attempted to capture the flows in terms of the <u>actual</u> intervention activities <u>reported on</u> in the project on-site visits. Exhibit 7 displays a typical ISP project

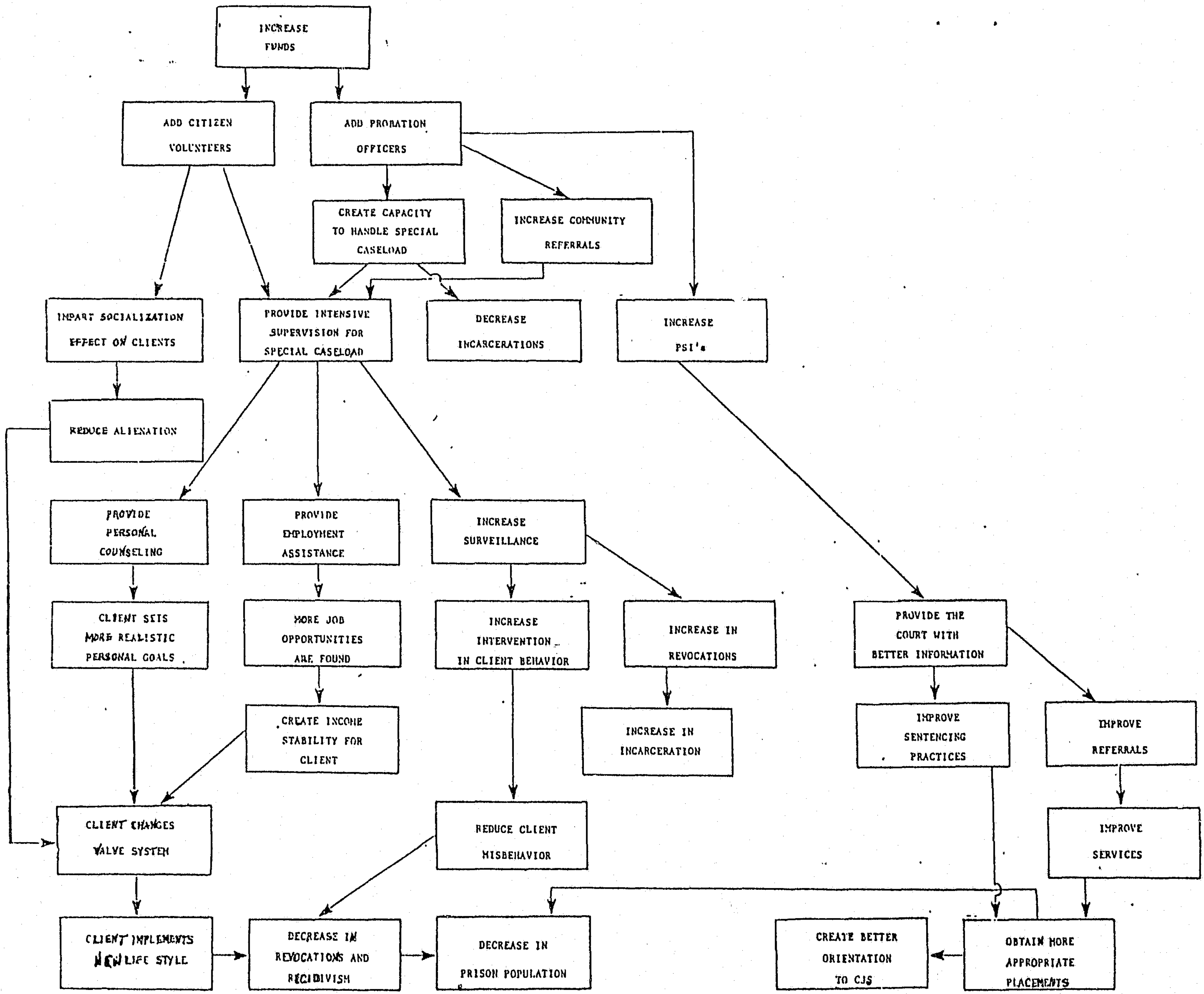


EXHIBIT 5: Intensive Special Probation Project Logic Flow Diagram

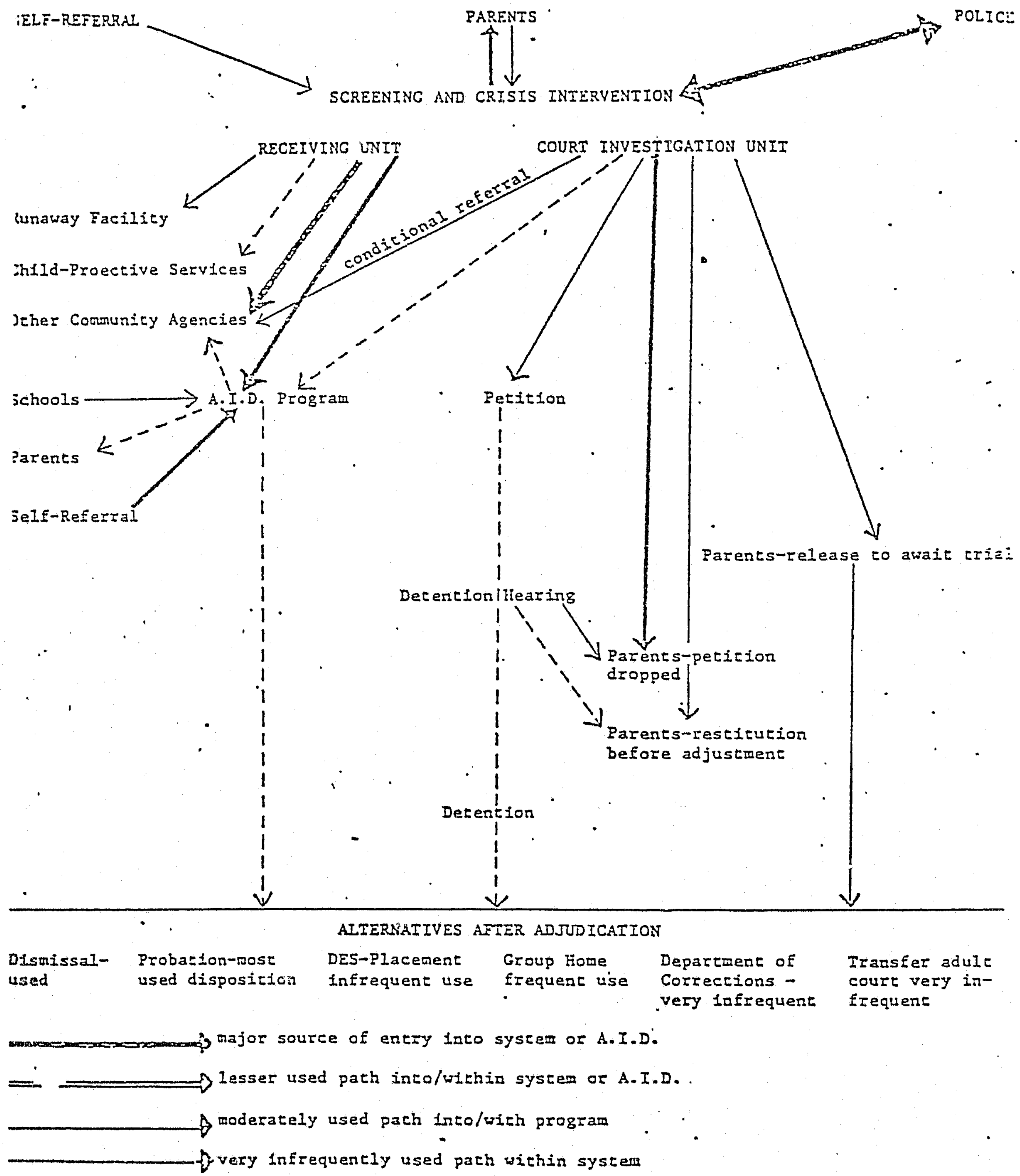


EXHIBIT 6: Juvenile Diversion Program Flow Diagram
(from Phase I report)

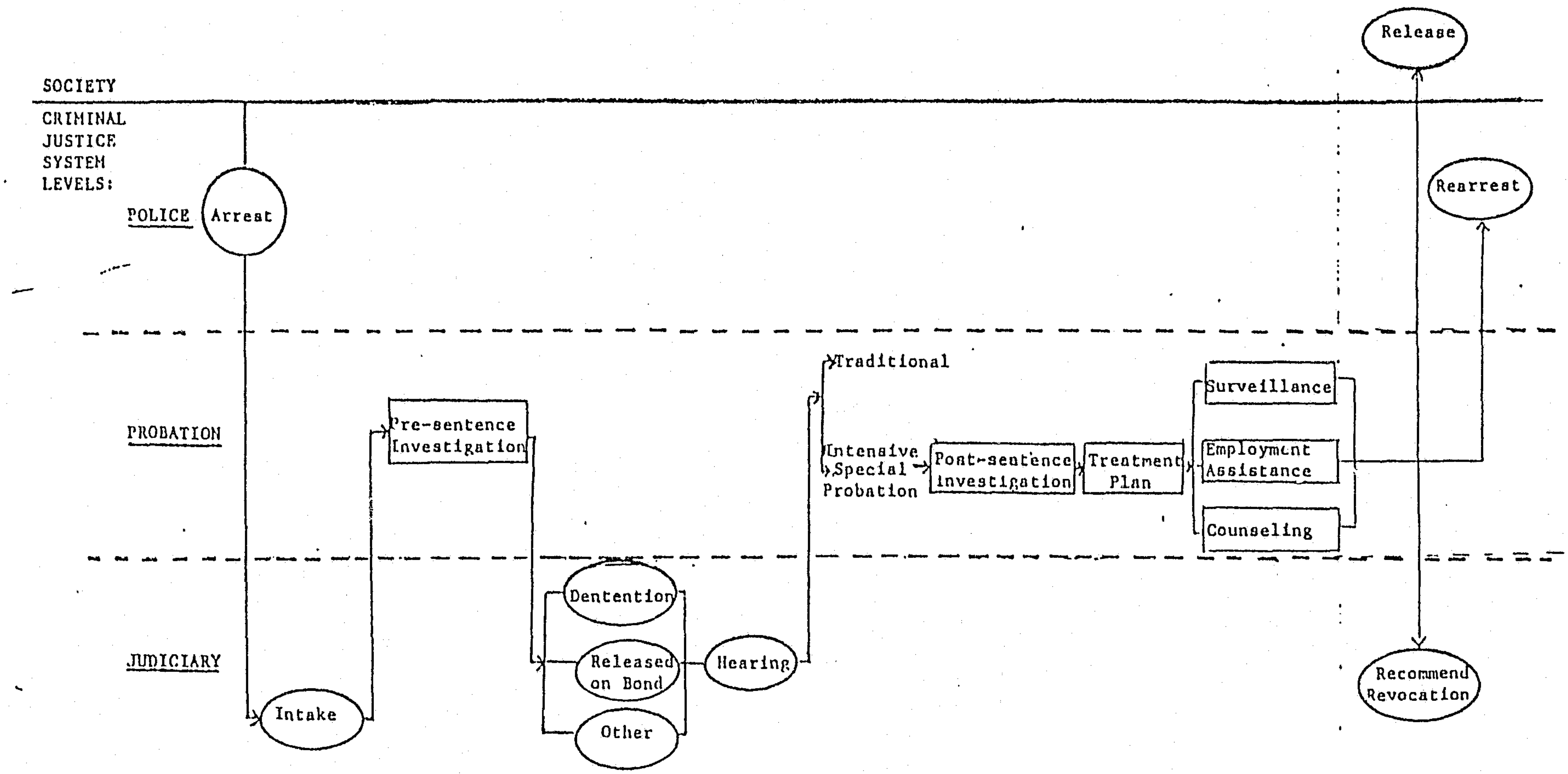
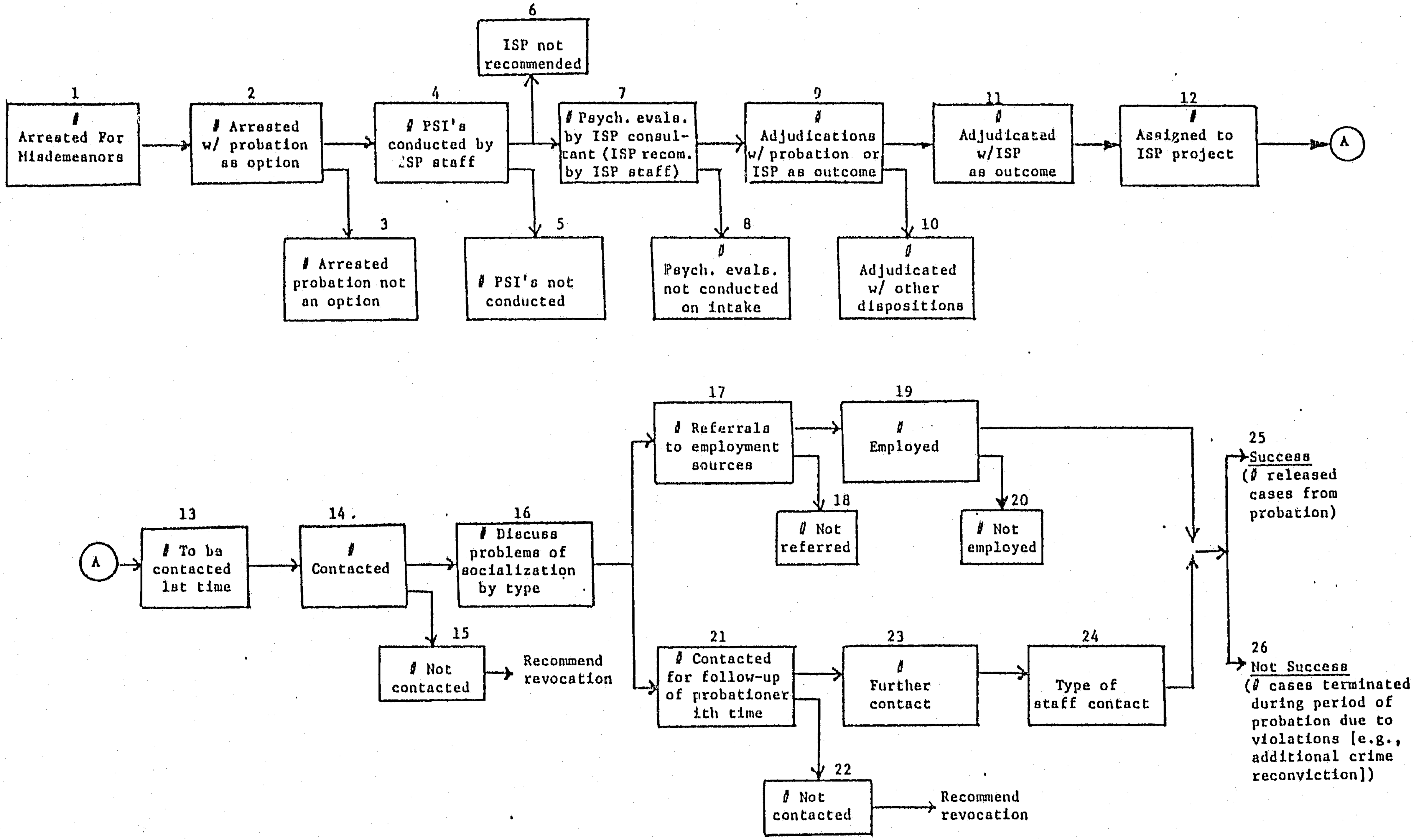


EXHIBIT 7: Intensive Special Probation Simplified
Case Flow Diagram for a Project Intervention
(constructed from information in the report)

in its operating environment. A general flow of people or cases is shown indicating key functions performed and showing which components of the criminal justice system are involved at each stage. Some flows (e.g., information) are not shown because they were not indicated in the source documents in detail. This and succeeding diagrams have been constructed for this exercise from information in the reports.

After these diagrams were constructed we reviewed information contained in the project data base and the list of states and media in order to create a measurement model of each project. The information pertaining to process, outcome and impact measures along the media at the project level were used along with project flow diagrams to construct the measurement model. The unifying concept (i.e., people or cases) was used in the modeling and each measurement was expressed in terms of a measurement point, the expected outcome of specific functions or events (e.g., the number of people arrested).

For each project, we attempted to draw (from the resource documents) activities of the project and its environment (and their relationship to each other), effort or personnel allocations, important known or intervening variables, in terms of points of measurement and measures. The various pathways show the process and outcomes indicated. Exhibit 8 represents a product of such an effort. This model depicts a particular ISP project which provides services only to adults arrested for misdemeanors. The primary activities of this project involve (1) the screening of such individuals, (2) the submission of recommendations to the judiciary concerning the potential referral of individuals to the project; (3) the provision of services to assigned cases in locating suitable employment; and (4) the continuance of contact with the assigned cases on a frequent basis over the period of probation to assist them in solving problems. Exhibit 9 illustrates another ISP project. The intervention activities of this project interact only after cases are assigned



#: Number
PSI: Presentence Investigation

EXHIBIT 8: Simplified Measurement Model Of An Intensive Special Probation (ISP) Project - Client Flow Diagram

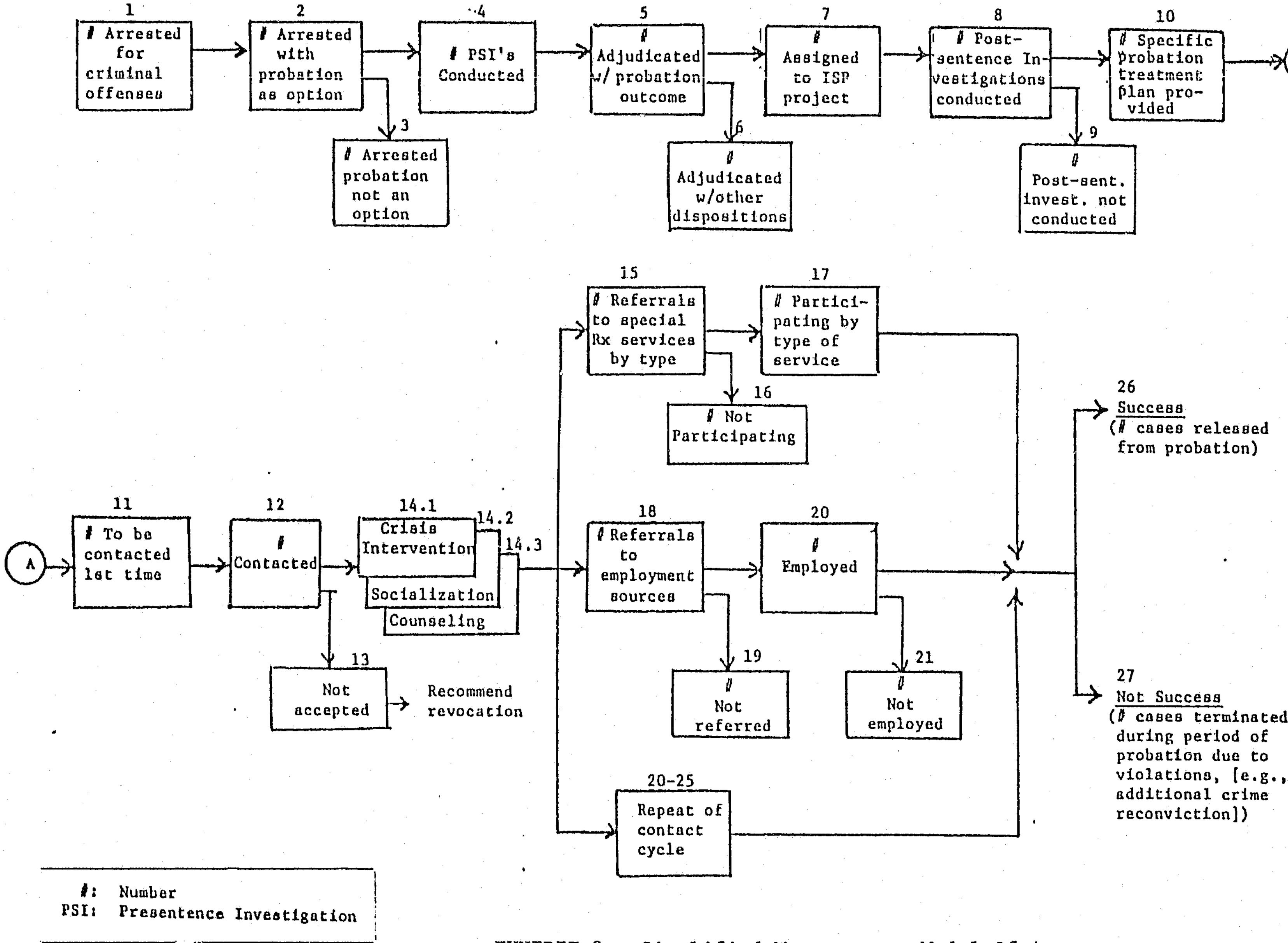


EXHIBIT 9: Simplified Measurement Model Of An Intensive Special Probation Project - Client Flow Diagram

to them. In addition, the project uses its resources in ways that expand the types of interaction activities made with their caseload and uses the resources of the community for special treatment services on a referral basis. This particular project aids its caseload also in locating employment and overcoming problems of socialization on a continuing basis during the period of probation.

After the individual measurement models were developed, we examined the array of projects to determine the need for creating a project typology which accounted for variations in media and states. Such a typology was not required in the topic areas studied. Only those characteristics that could be drawn from the Phase I reports were included.

To arrive at the generalized measurement model for the entire set of actual ISP projects that were site visited, a listing was made of all the measurement points and measures for these projects. (Table 2 lists all of the measurement points and measures of Exhibits 8 and 9 for instance.) This overall listing was examined. A synthesis was made of the measures necessary for a general ISP measurement model (shown in Exhibit 10). Table 3 illustrates the measurement points and measures from the synthesized diagram and how they apply to Exhibits 8 and 9. Since no other media was extracted from the source documents, this general model represents the universe of ISP projects described in the report. Particular projects described in the report are all degenerate cases of the general model.

Exhibits 11, 12, and 13 illustrate the general measurement models created for Juvenile Diversion, Residential Inmate Aftercare and Pre-Trial Release, respectively from material in the applicable NEP Phase I reports. Each of these models were constructed on the basis of the approach described above and again encompass the particular projects described.

TABLE 2

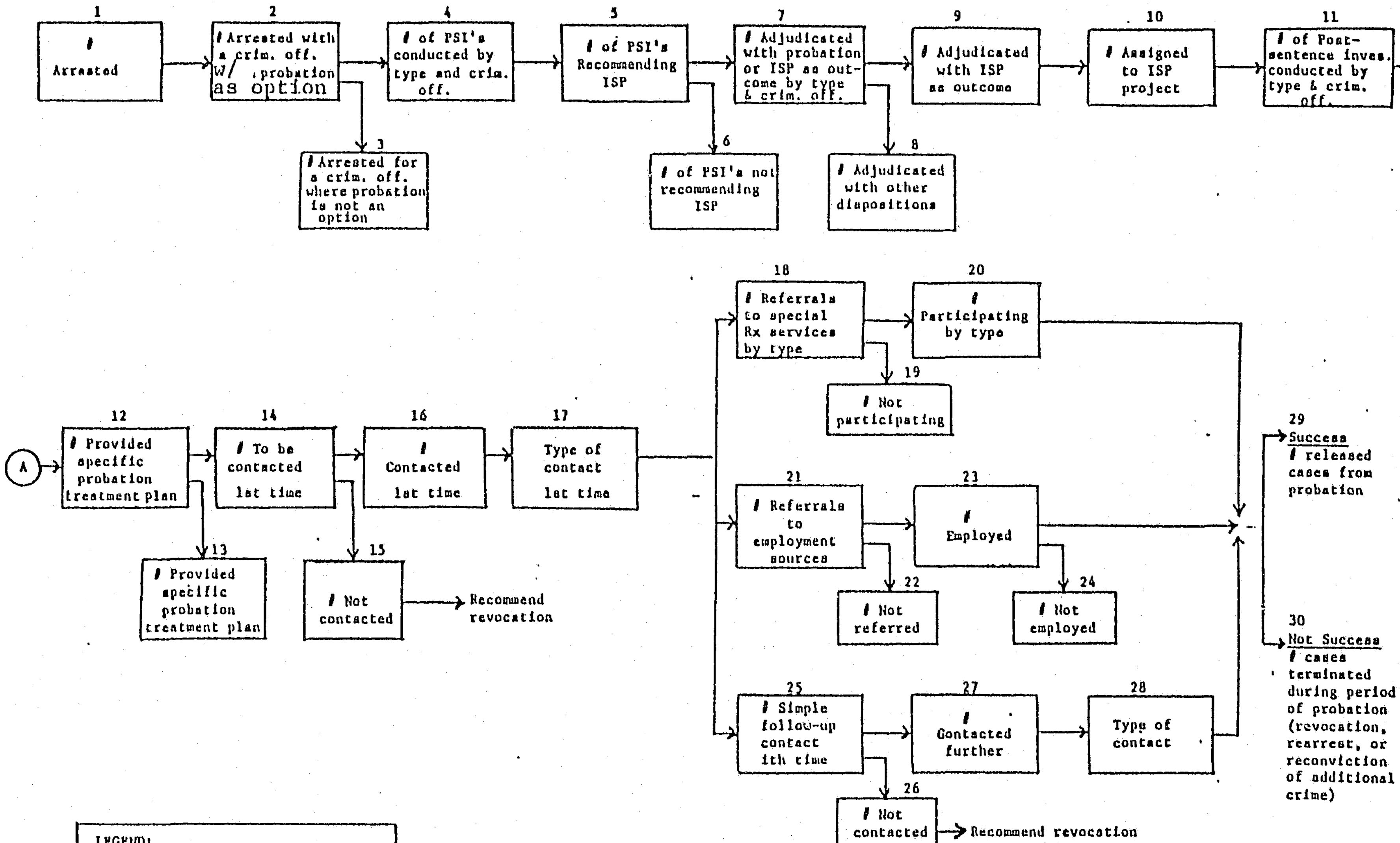
TABULATION OF MEASUREMENT POINTS AND MEASURES

(Exhibit 8)

Measurement Point	Measure
1	# Arrested for misdemeanors
2	# Arrested with probation as an option
3	# Arrested with no probation option
4	# Probation Screening Interviews (PSIs) conducted by Intensive Special Probation (ISP) staff
5	# PSIs not conducted
6	ISP not recommended
7	# Psychological evaluations conducted on intake by ISP consultant (ISP recommended by ISP staff)
8	# Psychological evaluations not conducted on intake
9	# Adjudicated with probation or ISP as outcome
10	# Adjudicated with other dispositions
11	Adjudicated w/ISP as outcome
12	# Assigned to ISP project
13	# To be contacted, 1st time
14	# Contacted
15	# Not contacted
16	# Discuss problems of socialization by type
17	# Referrals to employment sources
18	# Not referred
19	# Employed
20	# Not employed
21	# Contacted for follow-up of probation, ith time
22	# Not contacted
23	# Further contact
24	# Type of staff contact
25	# Success (# released from probation)
26	# Not success (# cases terminated during probation due to violations)

(Exhibit 9)

Measurement Point	Measure
1	# Arrested for criminal offences
2	# Arrested with probation as an option
3	# Arrested with no probation option
4	# PSIs conducted
5	# Adjudicated by type of probation outcome
6	# Adjudicated with other dispositions
7	# Assigned to ISP project
8	# Post-sentence investigations conducted
9	# Post-sentence investigations not conducted
10	# Specific probation treatment plan provided
11	# To be contacted, 1st time
12	# Contacted
13	# Not contacted
14.1	# Crisis intervention
14.2	# Discuss problems of socialization by type
14.3	# Counseling by type
15	# Referrals to special Rx services by type
16	# Not participating
17	# Participating by type of service
18	# Referrals to employment sources
19	# Not referred
20	# Employed
21	# Not employed
22-25	Repeat of contact cycle
26	# Success (# released from probations)
27	# Not success (# cases terminated during probation) due to violations



LEGEND:
- Number
crim. off. - criminal offenses
PSI's - Pre-Sentence Investigations
ISP - Intensive Special Probation
Rx - treatment

EXHIBIT 10: Intensive Special Probation Generalized Client Flow Diagram

TABLE 3

SYNTHESIS OF MEASUREMENT POINTS AND MEASURES OF GENERAL INTEREST
AS DERIVED FROM EXHIBITS 5 AND 6

Synthesis Measurement Point	Measure	Exhibit 8 Measurement Point	Exhibit 9 Measurement Point
1	# Arrested	1	1
2	# Arrested for criminal offense with probation as an option	2	2
3	# Arrested for criminal offense with no probation option	3	3
4	# PSIs conducted by type and criminal offense	4	4
5	# PSIs recommending ISP	7	-
6	# PSIs not recommending ISP	6	-
7	# Adjudicated with probation or ISP as outcome by type and criminal offense	9	5
8	# Adjudicated with other disposition	10	6
9	# Adjudicated with ISP as outcome	11	-
10	# Assigned to ISP project	12	7
11	# Post-sentence investigations conducted by type and criminal offense	-	8
12	# Provided specific probation treatment plan	-	10
13	# Not provided specific probation treatment plan	-	-
14	# To be contacted, 1st time	13	11
15	# Not contacted	15	13
16	# Contacted, 1st time	14	12
17	Type of contact, 1st time	16	14.1,14.2,14.3
18	# Referral to special Rx service, by type	-	15
19	# Not participating	-	16
20	# Particpating by type	-	17
21	# Referral to employment sources	17	18
22	# Not referred	18	19
23	# Employed	19	20
24	# Not employed	20	21
25	# Simple follow-up contact, ith time	21	-
26	# Not contacted	22	-
27	# Further contact	23	-
28	# By type of contact	24	-
29	Success - # released cases from probation	25	26
30	Not success - (# cases terminated during probation to violations)	26	27

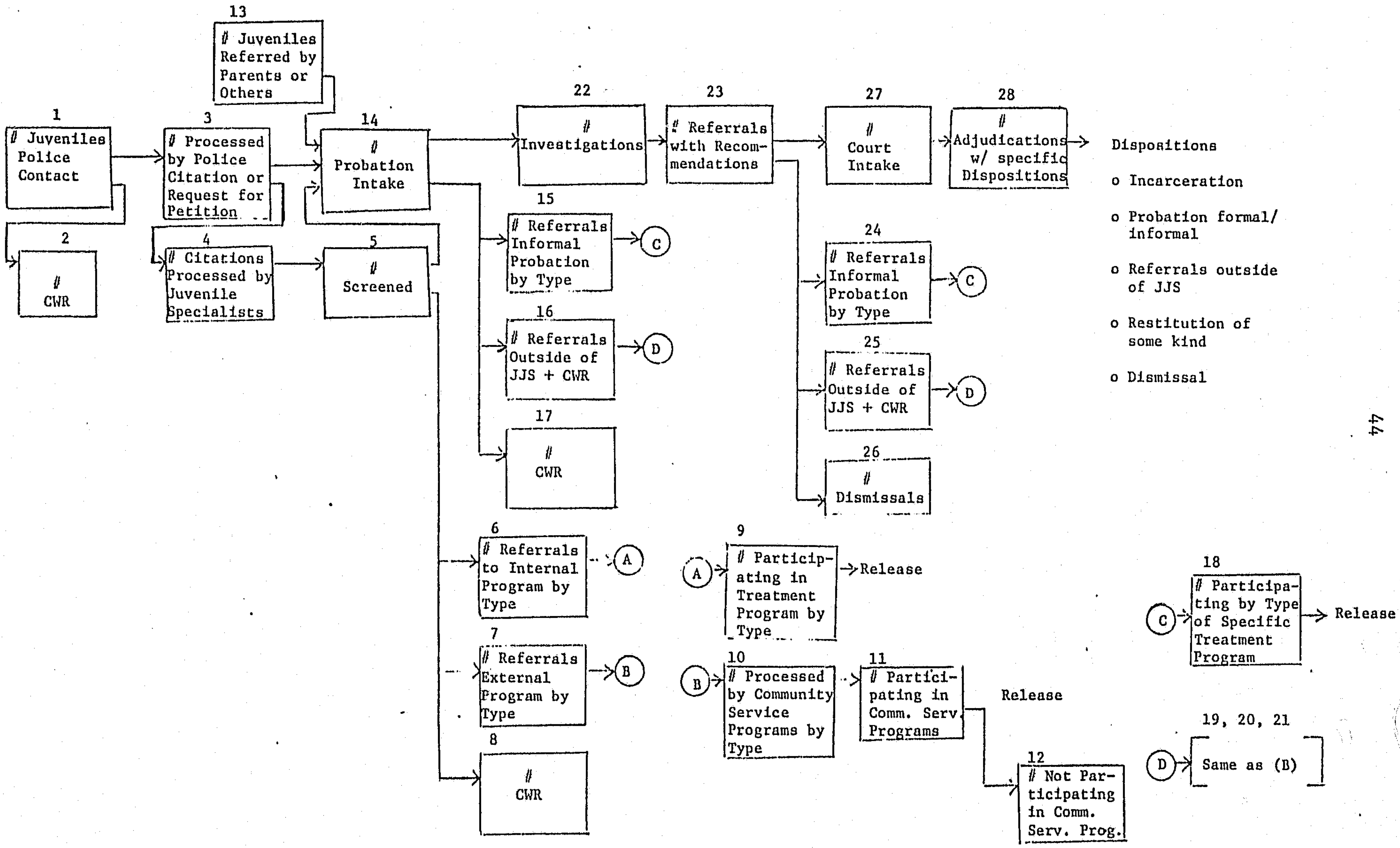


EXHIBIT 11: General Measurement Flow Model of Juvenile Diversion Within and Outside of JJS, including Measurement Points and States

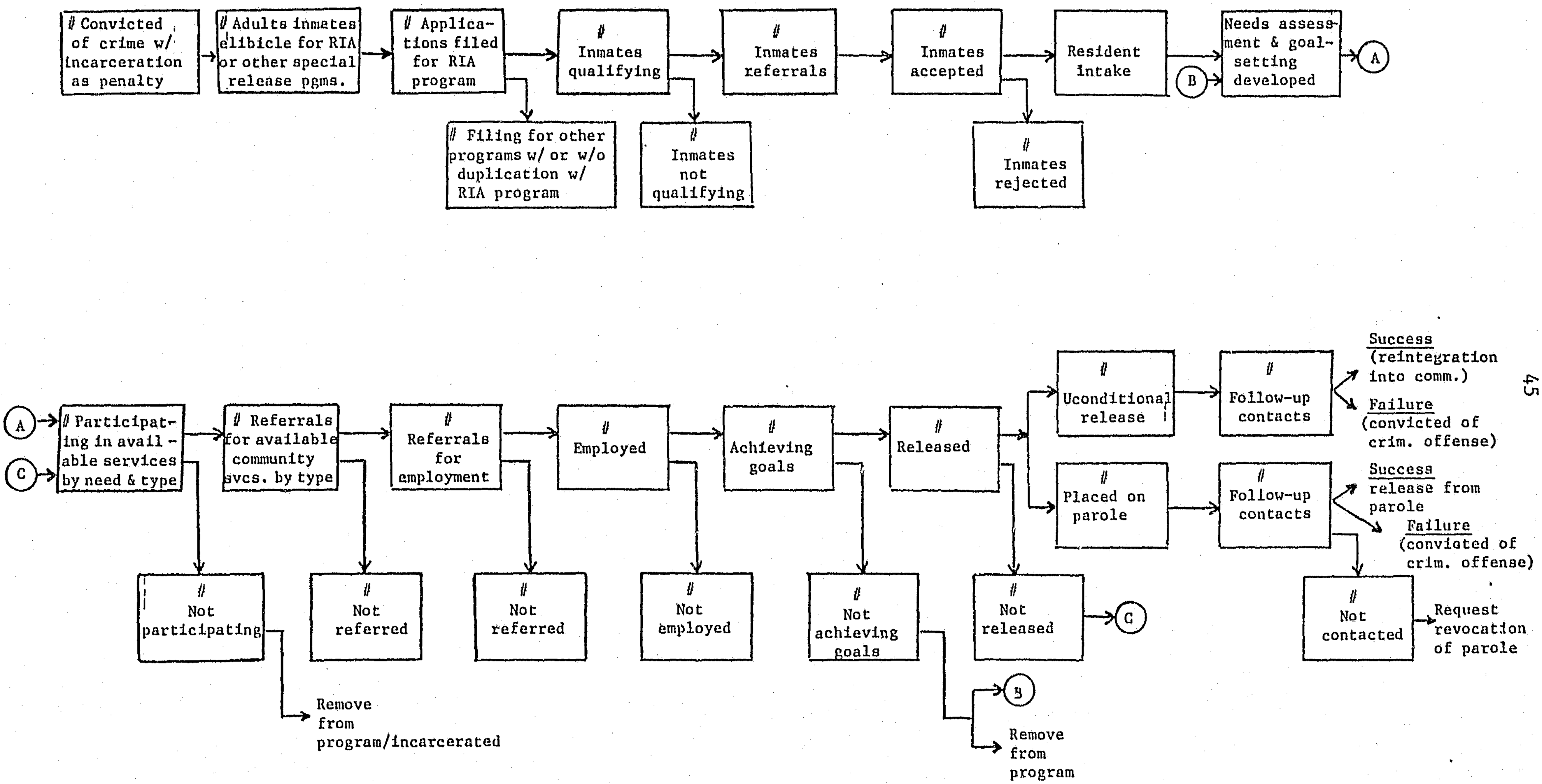


EXHIBIT 12: Residential Inmate Aftercare Client Flow Model With Some Measurement Points And Measures

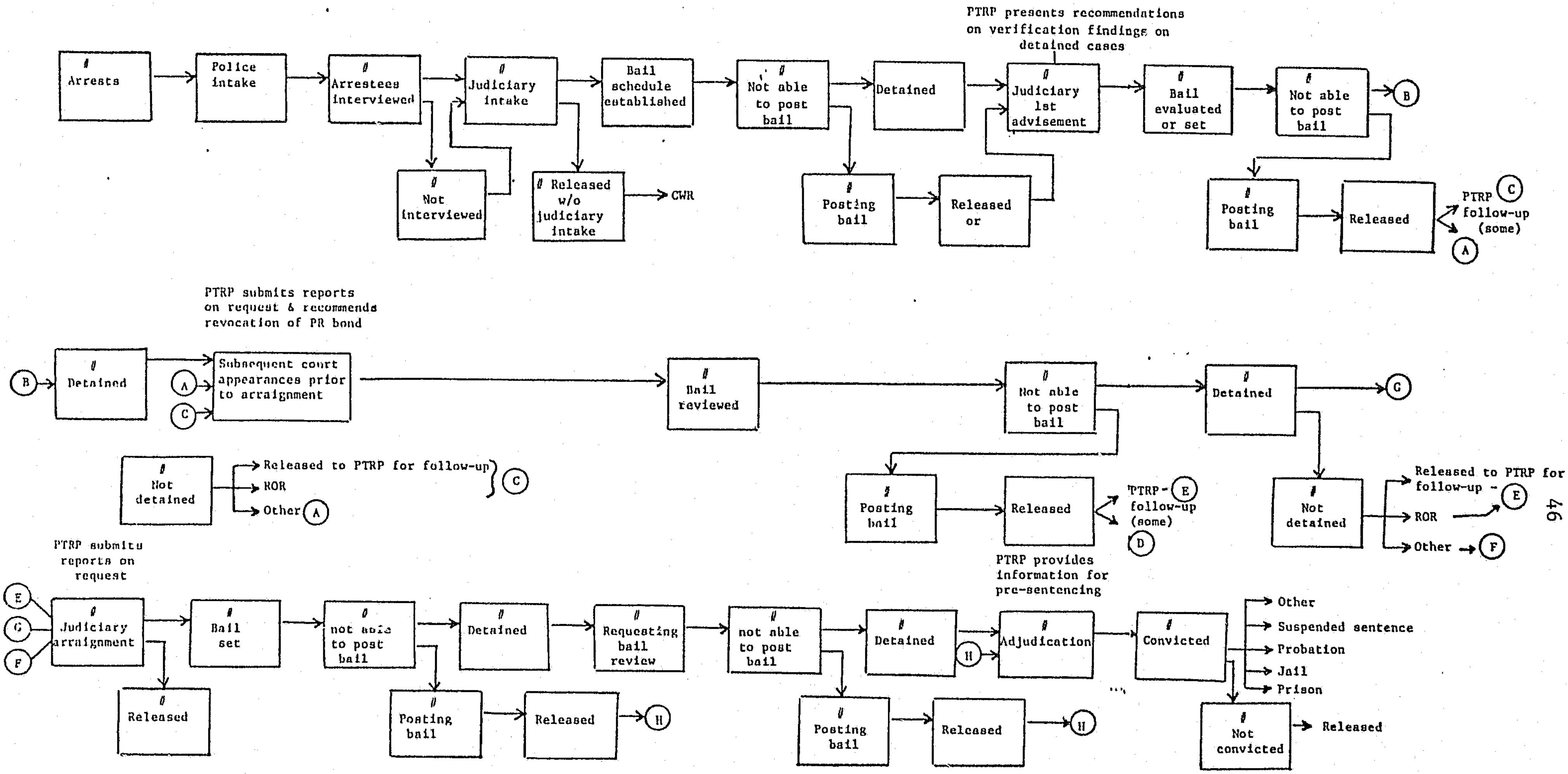


EXHIBIT 13: Measurement Model of Pre-Trial Release Program Showing Measurement Points and Measures

After creating the general measurement models for each NEP, we recognized several additional work requirements (e.g., organizing files for each model) involving (1) the data and information on project process, input/outcome and impact by media states and (2) measurement systems used by project in relation to measurement points. Likewise, we observed that the NEP Phase I products did not contain sufficient data to "test" the various measurement models developed. We noted that the general measurement models and the other materials generated through this process provided a sufficient data base to begin designing evaluations which could improve the current state of knowledge and provide information of interest to policy makers, practitioners and researchers for some questions, but not for others.

Having proceeded this far in development (i.e., general logic and measurement models), we explored potential evaluation design frameworks which would result from the linkage of logic and measurement models and use data bases or files from each NEP. In the work effort of developing the logic models from the NEP topic area studies, we observed that the audience, consisting of policy makers, practitioners and researchers, had a variety of specified questions, issues or expectations. These particular points of concern and/or interest were transformed to the extent possible, into input/outcome, process and impact measurements which they would want to see taken for question resolution. (NOTE: Since this entire modeling activity relies on the use of secondary sources, we extracted information from the source material and exercised judgment in specifying the candidate logic models constructed. Some are illustrated in the examples shown subsequently. Under "normal" conditions such logic models would best be developed by the evaluators through direct interactions with the various user audiences.)

After reviewing the array of logic models associated with a particular NEP, in conjunction with the general measurement model for the same NEP, we made the observation that <u>only</u> some of the logic models identified could possibly be resolved in terms of the measurement model's characteristics. Comparisons must be made between measurements contained on both the general measurement and particular testable logic models. In some cases the measurements identified were either identical or with slight modification of the data collection procedures might be made congruent. In other cases, the measurements identified in the logic models created just did not exist (i.e., they were not contained on the measurement model at all). In other cases, only some of the measurements required were contained on the measurement model. If measures are to be captured that answer some questions, extensive changes would potentially be required to existing data collection, reporting systems and in some cases to project activities. In the latter cases, nothing related to the questions being asked actually happens in any of the projects examined or in their environment.

For those testable logic models associated with an NEP that could be directly linked to its companion general measurement model, the objective would finally be reached of laying out a systematic process for designing evaluations for that question. Designs could be formulated to evaluate the "expectations" expressed by the testable logic models using a "testbed" of actual operating projects belonging to the measurement mode s typology.

A few examples are presented below that attempt to illustrate some of the notions discussed above. With this small an effort (this feasibility study), it is impossible to reproduce the amount of detail considered and examined, so samples are shown.

Example 1: NEP Topic Area - INTENSIVE SPECIAL PROBATION

Using only the NEP products, an array of logic models were developed. Such models were examined and compared with the measurements contained in the general measurement model of the ISP. A logic model was selected from this array to show how the product resulting from linkage of the two models leads to establishing a framework for evaluation designs. The logic model used is presented below.

ISP Logic Model for One Question

If probation officers are supported by citizen volunteers--in accordance with the guidelines defined for the number of citizen volunteers (1) under the authority of each probation officer and (2) assigned to each probationer such that provisions are made for the conduct of both counseling and employment assistance services as well as supervision/surveillance of special caseloads--then there will be a 10 percent reduction in the number of revocations and in the rate of recidivism involving the caseload supervised during the X years of their probation period supervision, where revocations and recidivism is defined in the LEAA Manual, as compared to other comparable ISP approaches using probation officers <u>only</u>.

Probation officers aided by citizen volunteers counsel, assist in locating employment, and supervise caseloads (special) according to established guidelines defining individual roles, functions & relationships to probation cases.	————>	Supervised caseloads during X yrs. of probation will result in 10% decrease in the number of revocations of probation and rate of recidivism as compared with other comparable ISP approaches using probation officers only.

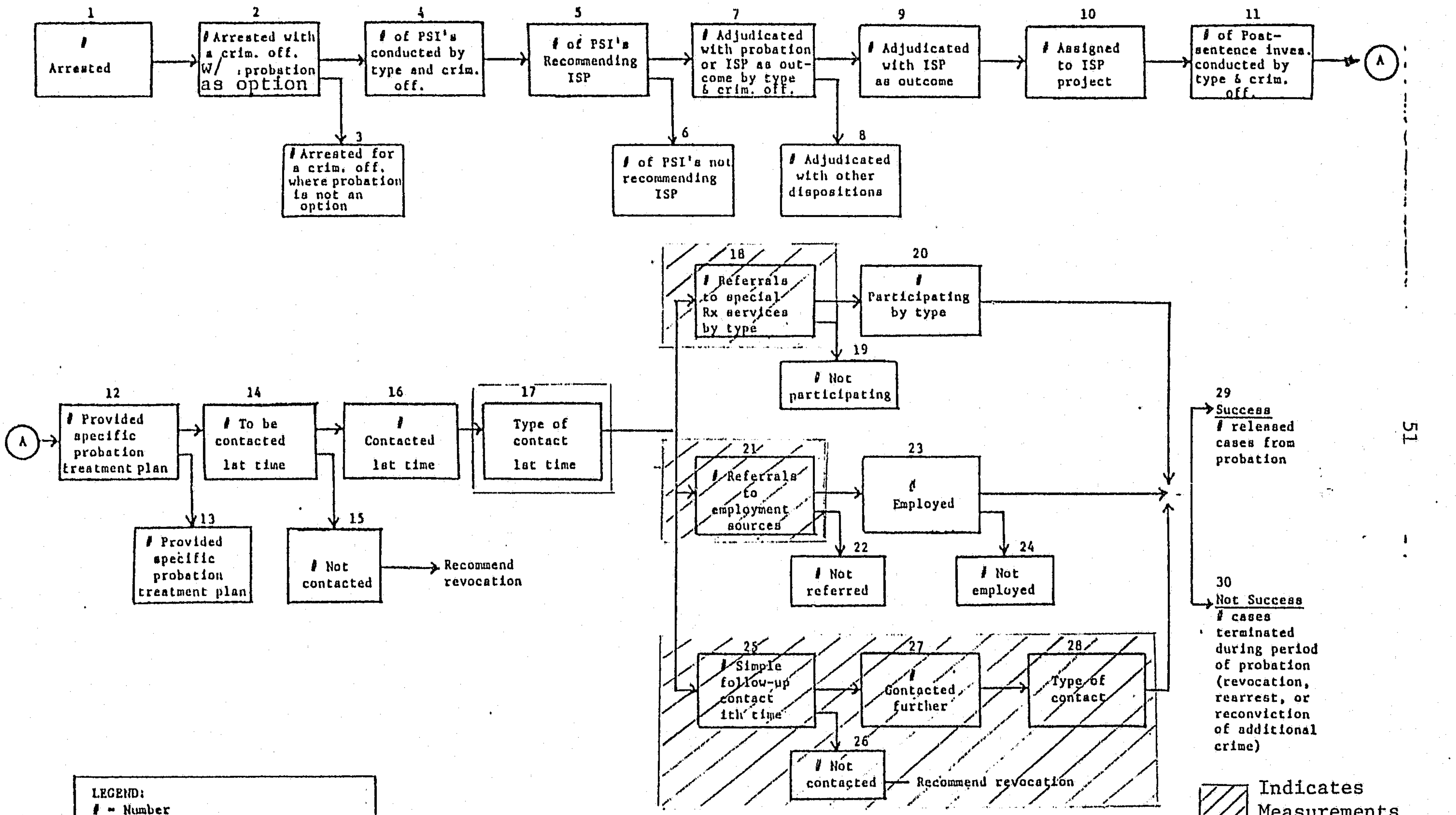
When this model is linked to the general measurement model (shown in Exhibit 14), we observe that all events identified in the logic model can be obtained (i.e., counseling, employment assistance services, supervision/ surveillance and outcomes, including revocation and success/not success). Since the flow model of Exhibit 14 represents a general model, some of the actual measurements such as measurement points 15, 17, 18, 21 and 25 which are listed in Table 4 need to be refined to indicate who are the ISP project staff performing the case interaction and how they do it.

Further examination of this table which lists points of measurement and measures identified a sizeable number of measurements that may effect the evaluation (e.g., the characteristics of cases assigned to the project). In such instances, the evaluator must draw upon the data base and files associated with the points of measurement of a particular project and determine those measurements considered necessary to most appropriately provide a response to the question, issue or expectation stated.

The evaluation design in addition to the plans needed to collect, analyze and process the measurement required would include also the identification of a list or group of projects under the typologies stated, specifically projects having similar characteristics involving the events and staffing called out in the logic model.

Example 2: NEP Topic Area - JUVENILE DIVERSION

A logic model was selected from the group of models developed from the source materials:



LEGEND:
- Number
crim. off. - criminal offenses
PSI's - Pre-Sentence Investigations
ISP - Intensive Special Probation
Rx - treatment

▨ Indicates Measurements Directly Affected by Logic Model Statement

EXHIBIT 14: Intensive Special Probation Generalized Client Flow Diagram Showing Sections Answering A Particular Expectation

TABLE 4

MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES CONTAINED ON ISP GENERAL MEASUREMENT MODEL

Measurement Point (No.)	Definition of the State	Measures of Interest
1	Total population in local agency jurisdiction arrested for criminal offenses	Number of people
2	Total arrestees having probation as an option	Number of people (adults)
3	Total arrestees having no probation option	Number of people (adults)
4	Total by type and criminal offense having PSI's	Number by category
5	Total arrestees having PSI's which recommend placement in ISP program	Number
6	Total arrestees having PSI's which do not recommend ISP program assignment	Number
7	Total arrestees adjudicated with either traditional probation or ISP as outcome by type and criminal offense	Number by category
8	Total arrestees adjudicated with other dispositions	Number
9	Total arrestees adjudicated with ISP as outcome by type and criminal offense	Number by category
10	Total arrestees assigned to ISP program	Number
11	Total ISP cases have post-sentence investigation by type & crim. offense	Number by category
12	Total ISP cases provided a specific treatment plan	Number
13	Total ISP cases not provided a specific treatment plan	Number
14	Total ISP cases to be contacted, 1st time	Number
15	Total ISP cases not contacted, 1st time	Number

TABLE 4 (Cont'd)

MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES
CONTAINED ON ISP GENERAL MEASUREMENT MODEL

Measurement Point (No.)	Definition of the State	Measures of Interest
16	Total ISP cases contacted, 1st time	Number
17	Total ISP cases contacted by type of contact, 1st time	Number and type of contact
18	Total ISP cases referred to special treatment service by type	Number and type of treatment service
19	Total ISP cases not participating in special treatment service by type	Number and type of treatment service
20	Total ISP cases participating in special treatment service by type	Number and type of treatment service
21	Total ISP cases referred to employment sources	Number
22	Total ISP cases not referred to employment sources	Number
23	Total ISP cases provided employment	Number
24	Total ISP cases provided employment	Number
25-28	Total ISP cases contacted for simple follow-up, ith time	Number

Logic Model

If diversion project for juveniles were managed and administered by the police sector of the criminal justice system where the only other programs under their authority are confined to cite-warn-release (CWR),-- "education" concerning the law and consequences inherent in its violation, intensive counseling or "services," and informal "probation"--then the regular system will formally adjudicate X percent fewer cases per year than other juvenile justice systems.

Linking the logic model, listed above, with the general measurement model, shown in Exhibit 15, along with the data base and files developed on this topic area, we observed that it was possible to structure an evaluation design. This evaluation, for all intents and purposes, would compare the effect or impact on court caseload (i.e., total number of juvenile adjudications over a one year time period) resulting from alternative ways of processing and diverting offenders in juvenile justice systems. The alternatives considered for this evaluation range from juvenile justice systems which process such offenders without referrals outside the system or formal/informal diversion programs all the way to juvenile justice systems which practice and apply minimization of penetration philosophies within each sector (i.e., police, probation and court), as well as refer juveniles to programs outside of such systems.

A project typology is required for development of this design and might consist of project grouping along the following lines:

1. juvenile justice systems which comply with the logic model statement;
2. juvenile justice systems which apply traditional diversion (i.e., discretionary judgments by their personnel to not process, informally or to refer to non-specialized community programs);

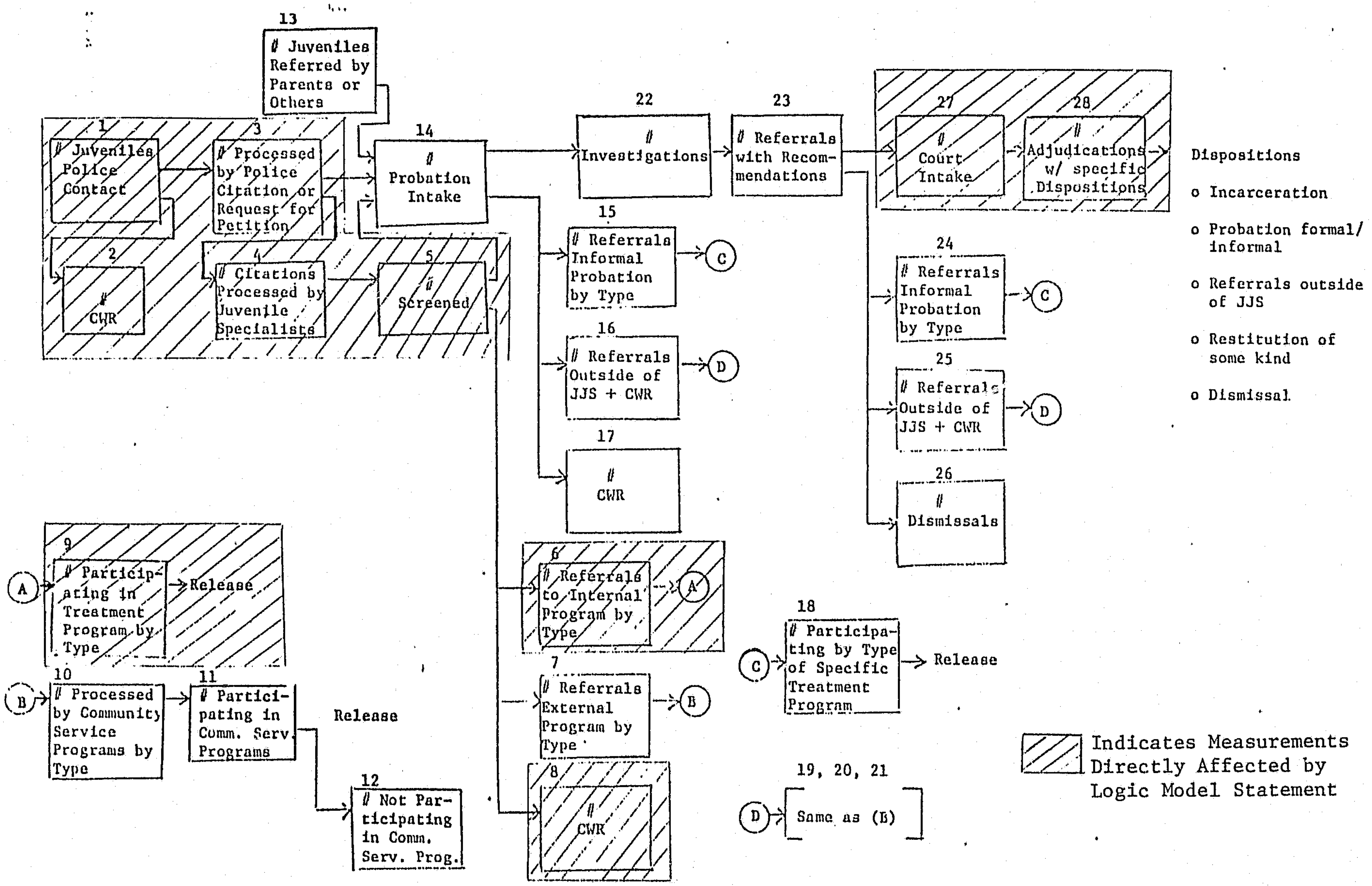


EXHIBIT 15: General Measurement Flow Model Of Juvenile Diversion Within And Outside Of JJS, Indicating Measurement Points And States For A Particular Testable Logic Model

3. juvenile justice systems which process offenders at each sector within such systems either informally or formally with and without referrals to non-specialized or specialized community programs; and

4. juvenile justice systems which are similar to the logic model statement and also refer cases to non-specialized and/or specialized community programs.

Based on the logic model statement, the key measurements contained on the general measuement model can be identified; these are listed in Table 5. Such measurements might require other relevant definitions (e.g., juvenile crime, state and local offenses, community and conditions, policy, and regulations) and measures. This would be in order to better define diversion in each part of the juvenile justice system that might become involved in the selection of projects.

Using the information, highlighted above, and the development of the remaining components involved in an evaluation design (e.g., measurement instruments and/or data collection instruments, analysis plan) an evaluation design to conduct such a comparison can be accomplished.

TABLE 5

MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES CONTAINED ON JUVENILE DIVERSION GENERAL MEASUREMENT FLOW MODEL

Measurement Point (No.)	Definition of the State	Measures of Interest
1	Total population of juveniles located in community that are contacted by police for "criminal offenses" by type, including characteristics of juveniles such as age, race, sex and prior history involving JJS	Number of juveniles, their ethnographic and prior history with JJS
2	Total juvenile population diverted by police	Number by key characteristics
3	Total juvenile population served with citation or processed further (request for petition)	Number by category
4	Total juveniles processed by juvenile specialists	Number
5	Total juveniles screened	Number
6	Total juveniles referred to internal police programs by type	Number
7	Total juveniles referred out to external diversion programs by type	Number by type of program
8	Total juveniles diverted by juvenile specialists	Number
9	Total juveniles participating in treatment programs by type	Number by type of program
10	Total juveniles processed in community services programs by type	Number by type of program
11	Total juveniles participating in community services programs by type	Number by type of program
12	Total juveniles not participating in community services programs by type	Number by type of program
13	Total juveniles referred to JJS by parents or others	Number
14	Total juveniles processed by probation officers	Number

TABLE 5 (Cont'd)

MEASUREMENT POINTS, DEFINITIONS OF STATES AND MEASURES CONTAINED ON JUVENILE DIVERSION GENERAL MEASUREMENT FLOW MODEL

Measurement Point (No.)	Definition of the State	Measures of Interest
15	Total juveniles referred to informal probation programs by type	Number by type of program
16	Total juveniles referred to programs outside of JJS with CWR	Number by type of program
17	Total juveniles diverted	Number
18	Total juveniles participating in informal probation programs by type	Number by type of program
19	Total juveniles processed in community services programs by type	Number by type of program
20	Total juveniles participating in community services programs by type	Number by type of program
21	Total juveniles not participating in community services programs by type	Number by type of program
22	Total juveniles involved in probation investigations	Number
23	Total juveniles provided referrals with recommendations through investigation officers	Number
24	Total juveniles referred to informal probation programs by type	Number by type of program
25	Total juveniles referred to programs outside of JJS with CWR	Number by type of program
26	Total juveniles dismissed from JJS based on investigation recommendation	Number
27	Total juveniles referred to judiciary for further (formal) processing	Number
28	Total juveniles adjudicated by JJS and provided specific dispositions	Number by category

V. SOME FURTHER OBSERVATIONS FROM THE FEASIBILITY EXERCISE

Examination of the full sets of Phase I products for each of the four studies--Juvenile Diversion, Intensive Special Probation, Halfway Houses (Inmate Aftercare), and Pretrial Release--in an attempt to develop measurement models produced some additional observations:

- More effort needs to be made in future Phase I studies to reduce theoretical issues to testable questions and focus the fieldwork on those questions.
- Measurement Models could become quite complex unless testable questions applicable to the actual operations of sets of similar projects are addressed.
- Finding existing local data to meet measurement needs was apparently difficult.

Each of these is addressed briefly below.

A. THEORETICAL ISSUES NEED TO BE REDUCED TO TESTABLE QUESTIONS TO FOCUS THE WORK

The grantees apparently had trouble deciding on sets of testable questions to be specifically addressed in the topic areas and in reducing them to testable logics. Expectations expressed by planners, theorists, academics, and in past work in each area tended to vary in the level of semantics, specific definition, and details of measurements that would be acceptable. Many of the extant policy and theory discussions are in terms (e.g., crime reduction, lower rates of recidivism, less labelling, better integration into the community) that sound quite solid as potential measures until an attempt to specifically operationalize them is made by the grantee (or by this feasibility effort) in terms of specific measurements to be taken at specific project sites. The desired measurements are then found to be too ambiguous and in need of much more detail definition, even to achieve general testable and measurement models.

The point that may need to be made more clearly to grantees is that reduction of the large numbers of rhetorical expectations to testable logics is really part of determining what must be learned from field projects and what must be modeled in the measurement models. Using lists of dozens of unsorted questions as a guide means that nearly everything examined must be modeled and the task becomes impossible. Using only a single question would, of course, usually oversimplify important questions that need answering. The grantees treated the "issues paper" work in a literate and thoughtful way, but did not always develop it far enough into testable models that might bring more order into the other, later tasks.

The problem that we found in using these studies is not usually one of too restrictive a set of questions, but rather of a plethora of possible questions that might be selected for examination and study. Faced with this, the grantee often oriented the modeling work around something other than the key issues developed earlier in their own study.

The questions chosen determine in part what must be modeled, what will be learned about operating project sites and what measurements would be necessary to provide answers through evaluations. For instance, the reader can easily see how additional elements could have been developed and included in the measurement models created as examples in this paper in response to other important testable questions (even though this feasibility effort was limited to material in present reports).[1]

---

1. In fact in several cases additional material was developed, but omitted here to make this document of reasonable size.

To some extent this problem may be alleviated by the new work statement. Special attention should also be paid to the problem by present and future NEP Phase I monitors in NILECJ.

B. GENERAL MEASUREMENT MODELS COULD BECOME QUITE COMPLEX UNLESS TESTABLE QUESTIONS APPLICABLE TO THE ACTUAL OPERATIONS OF SETS OF SIMILAR PROJECTS ARE ADDRESSED

The work of extracting a measurement model from reality starts with broad rhetorical expectation statements at the national level (to get testable logics) but then passes to constructions of equivalent functional flow diagrams and consideration of sets of measurements at and for a specific project. A constant effort was made to develop generalized measurement models that represented particular families of projects. Such an approach was used both by the grantees and in this feasibility effort. The exploration of feasibility resulted in two observations:

- The generalized measurement model that can be developed is likely to be a good guide to measures and measurement points across a set of similar projects. Additional process work will generally be necessary at field sites to develop exactly what is to be measured at any specific site to answer a specific question.
- The sites concentrate primarily upon delivery of service and are likely to be interested only in process, or at most outcomes. Most questions being asked at the national level are in terms of impact, or occasionally outcomes.

Let us briefly consider what these two observations mean:

As demonstrated in the chapter above, it is possible to develop--from Phase I reports--models representing interconnected measures and measurement points that are related to questions asked about the field projects and the operation of the field projects. These models (which could be even more fully developed than those shown in this paper) seem to serve as an excellent means of bridging discussions between those persons having questions, those measuring for answers, and those doing work at a field project. In fact without such models, such discussions are almost impossible to carry out. But what might seem to be significant evidence (answers) to different questioners and what measurements are to be made at particular points in the criminal justice system to obtain answers will vary. The models are therefore a necessary, but not sufficient, step for designing and conducting evaluations. That is, such models appear necessary for iterative discussions that finally result in agreement on the statements of the questions and the general types of measurements to be made to obtain the answers.

Developing the actual measurements for any specific group of projects will require an additional step, however. An evaluation made at any particular project will have to capture and to link specific explanatory and exogeneous variables which may be unique to that project. Unique variables necessitate measurements for a particular project which otherwise might not be considered, and which must be either controlled or accounted for locally so that they do not have a misleading affect upon the answers produced nationally. The severity and nature of criminal cases dealt with at each site, special local capabilities, and variations in law from site to site are examples that easily spring to mind. This probably indicates that the small pilot effort begun by the Office of Special Programs to develop both general models and of a process for developing specific local models from general ones on a site by site basis is probably a correct approach, although one that may take some time to complete. What this whole discussion--and indeed, perhaps the whole feasibility effort--implies about general national guidelines containing specified measurements to be used in all locations requires further thought.

All efforts to develop knowledge at the national level could profit from an early and broad recognition of the implications of this observation. Most of the national and state level expectations for and questions about programs--and especially those of most theorists in the four areas examined--are in terms of broad impacts, behavioral changes, and effects that take place at some distance (often in both time and space) from the intervention funded. This often means that, even when testable logic models and statements can be drawn from rhetorical ones, the statements are still in terms quite different from the concerns of most actual projects.

The actual projects almost universally concentrate on operating some portion of the CJS and upon the delivery of a service. The focus of a project and of its measurements are thus likely to be in terms of the service process (at most the focus is on immediate outcomes of that service). If answers to many national level knowledge questions are to be developed--beyond a knowledge of those programs that generally fail to operate well or generally succeed--then the gap between rhetorical national concerns and concrete project operations must be closed somewhat or a family of fairly complex measurement models will need to be developed.

Phase I grantees have an exceedingly difficult task at present because their charter and charge is to attempt to bridge this gap. In some cases, more specific sets of testable questions may have to be developed by LEAA if the questions are to be answered from study efforts and evaluations. This exercise, however, seems to demonstrate the feasibility of developing general measurement models representing selected similar projects.

C. PRESENCE AND AVAILABILITY OF SPECIFIC DATA

In determining if data are available to answer questions, each grantee had to grapple with both of the two problems above--what questions to address or test and what exact data are necessary for proof--and with the availability of such data in the field. The Halfway House study found more usable past evaluations available than most other Phase I studies. The other three studies examined here found succeedingly less data available or even recorded. Juvenile Diversion appears to have found the least amounts of broadbased data available and access to what is recorded is sometimes complicated by confidentiality requirements. The spotty availability of field data has arisen in two-thirds of the first Phase I studies. The possibilities are that recorded data are simply not there, that they are not in an obvious and available form, that more have been collected by grantees than they describe (known to be true in a few cases), or that more data are available, but simply have been missed in the field. Data availability remains a problem. This problem might have been improved had the specific questions been narrowed more as mentioned earlier. With this number of grantees--some of whom appear to have quite good field staffs--all encountering problems in finding available data, LEAA faces a problem that may have to be brought into sharper focus and dealt with in detail nationally.

At the national level, it appears that much more thought must be brought to bear on the detailed development of guidelines for a few carefully targetted, specific testable questions that are to be answered or else the efforts to know performance (and especially impacts of large number of projects nationally) may have to be abandoned. At the same time, increased emphasis by national and state officials on local monitoring of existing projects might cause better data to be available to future study efforts.

<u>Summary</u>

Overall, this feasibility effort indicates that it does appear feasible to develop measurement models to represent specific types of activities for answering specific questions. NILECJ should continue to press for the development and use of such models in the Phase I grants. More training should be provided to both monitors and grantees in the reasons for the approach and the methods of carrying it out.

VI. CONCLUSIONS AND RECOMMENDATIONS

Conclusions:

- The modeling approach laid out (Chapters II and III) and tested (Chapter IV) in this exercise proved to be an excellent way to organize information and data.
- It was possible to synthesize sample models and testable questions from material contained in the reports (samples in Chapter III).
- The amount of detail observed and recorded in the reports heavily influenced the level of specificity that could be included in a particular measurement model. The selections of questions given in the reports only secondarily affected the development of the sample measurement models.
- Questions (or issues raised) in a particular topic area are not always developed (in the four reports examined) to the point that they are testable nor always linked tightly to the models developed in that report from field visit data.

Recommendations:

- Questions should be developed in testable form for each Phase I study. Models related to those questions should then be developed and used by the grantee during each Phase I study and in preparation of the report. Illustrations are given in this report and this approach is already spelled out in the new work description. Following such procedures should make the information developed more readily available to to users.[1]
- The process used in this exercise (or a similar one) should be an integral part of all subsequent NEP work.
- Single project evaluation designs should be based on efforts similar to those illustrated in this report.

---

1. This implies that the grantee/NILECJ would slowly settle onto applicable, answerable, valuable sets of questions. Facilitating this is treated further under recommendations. Sample measurements that are more quality of process oriented are, for instance, more easily produced during a grant than in an exercise such as this one.

- Improvement in products of the National Evaluation Program can be achieved through:

  - Tighter development of issues and questions raised in a topic area into testable questions to be addressed within each Phase 1 study.

  - Better guidance and more training in using the measurement modeling approach.[1]

  - The National Institute may want to consider further proceduralization of the approach along the lines outlined herein.

- Improvement in topic area definition and issue development can be achieved through:

  - A set of testable question (logic) models should be selected nationally, or developed by a single grantee performing that function.

  - As part of the first step in defining topic areas, more effort in the topic definition state (nationally) and in the grantee issues effort (on each Phase I) must be made to develop testable questions and fit them to equivalent models of projects. By this tactic the measurement models can be much better defined when they are finally handed over to NILECJ and the work of the Phase I grantee effort made more manageable at an earlier point.

  - Grantees must have a starting set of questions early in the study to indicate the measurements of probable interest, field collections necessary, and a better approach to what must be studied at project sites. Grantees should not be limited to this starting set.

  - Other more appropriate testable question models may have to be created in the course of a study. The grantee should be allowed to do this and explain why it was done. Iteration with the monitor or with an advisory board will generally be necessary.

  - Basic work in this area could be carried out by NILECJ monitors, grantees, or by a combination of effort.

1. The materials developed here, the latest version of the NEP Phase I work description, and The Urban Institute 783 Series of Research Items would be suggested as training material for monitors and grantees alike.

END