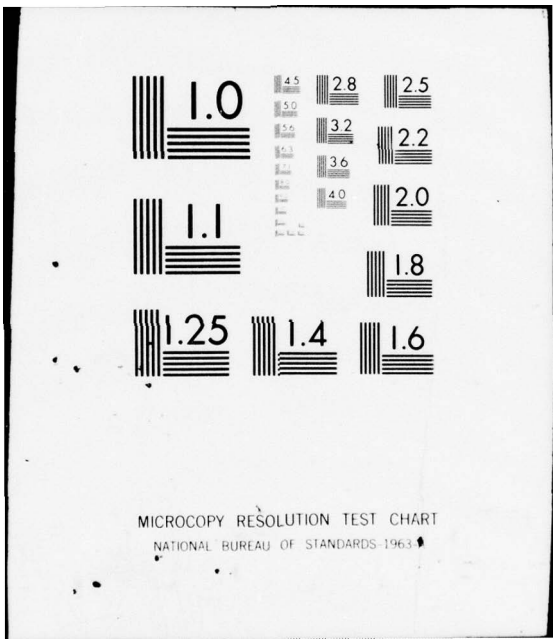
AD-A044 834 NEW YORK CITY-RAND INST N Y F/G 5/10
A PILOT APPLICATION OF DELPHI TECHNIQUES TO THE DRUG FIELD: SOM--ETC(U)
JUN 73 L T THOMPSON

UNCLASSIFIED R-1124 NL

1 OF 1
AD A044834

END
DATE FILMED
10-77
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963

A PILOT APPLICATION OF DELPHI TECHNIQUES TO THE DRUG FIELD: SOME EXPERIMENTAL FINDINGS

LORAN T. THOMPSON

R-1124

JUNE 1973

THE
NEW YORK CITY
RAND
INSTITUTE

Rand

The New York City-Rand Institute is a nonprofit research institution founded, as its Articles of Incorporation state, "... primarily to conduct programs of scientific research and study, and provide reports and recommendations, relevant to the operations, planning or administration of the City of New York." The Institute was established in 1969 by the City of New York and The Rand Corporation as a center for the continuing application of scientific and analytic techniques to problems of urban life and local government. It is governed by a Board of Trustees appointed jointly by the City and Rand.

Published by The Rand Corporation

9-6-73

ERRATA

R-1124 A PILOT APPLICATION OF DELPHI TECHNIQUES TO THE DRUG FIELD: SOME EXPERIMENTAL FINDINGS by Loran T. Thompson, June 1973

Please insert the two attached pages (pages 47 and 48) in your copy/copies of the above report.

The addition will necessitate your changing the pagination of your copy as follows: change the old page 47 to 49 and page 49 to 51.

PUBLICATIONS DEPARTMENT

Rand

SANTA MONICA CA 90406

Part III of Questionnaire (Round 2)

3.1. How many heroin addicts are there in New York City?

On the first round, 1/4 of the Mid-Estimates were below 35,000
1/2 of the Mid-Estimates were below 60,000
3/4 of the Mid-Estimates were below 100,000

Mid-Estimate: _____ [furnish Mid-Estimate first]

Low Estimate: _____ High Estimate: _____

3.2. In what year will marihuana be legalized?

On the first round, 1/4 of the Mid-Estimates were below 1973
1/2 of the Mid-Estimates were below 1976
3/4 of the Mid-Estimates were below 1980

Mid-Estimate: _____ [furnish Mid-Estimate first]

Low Estimate: _____ High Estimate: _____

NOTE: If in your opinion there is at least a 1 in 4 chance that marihuana will never be legalized, check here [ ] and leave the High Estimate item blank.

3.3. What percent of 25-year old males in Central Harlem have used heroin at least once?

On the first round, 1/4 of the Mid-Estimates were below 30%
1/2 of the Mid-Estimates were below 50%
3/4 of the Mid-Estimates were below 65%

Mid-Estimate: _____% [furnish Mid-Estimate first]

Low Estimate: _____% High Estimate: _____%

3.4. What percent of 25-year old males in Central Harlem have used marihuana at least once?

On the first round, 1/4 of the Mid-Estimates were below 50%
1/2 of the Mid-Estimates were below 75%
3/4 of the Mid-Estimates were below 85%

Mid-Estimate: _____% [furnish Mid-Estimate first]

Low Estimate: _____% High Estimate: _____%

Part III of Questionnaire (Round 2)

3.5. What percent of individuals who have used heroin at least once found it an unpleasant experience the first time?

On the first round, 1/4 of the Mid-Estimates were below 20%
1/2 of the Mid-Estimates were below 40%
3/4 of the Mid-Estimates were below 65%

Mid-Estimate: _____% [furnish Mid-Estimate first]

Low Estimate: _____%

High Estimate: _____%

(1)

(6)

A PILOT APPLICATION OF DELPHI TECHNIQUES TO THE DRUG FIELD: SOME EXPERIMENTAL FINDINGS

(10)

LORAN T. THOMPSON

(14) R-1124

(11) JUNE 1973

(12) 59p.

THE
NEW YORK CITY
RAND
INSTITUTE

Rand

PREFACE

In the summer of 1970 the author conducted an experiment to estimate the feasibility and usefulness of employing Delphi techniques as a forum for examining drug-related issues. This Report discusses findings from the pilot study.

SUMMARY

While Delphi techniques have traditionally been used to elicit the opinions of experts on questions whose true answers are essentially numerical, many questions of interest to policymakers in the field of drug-abuse prevention and treatment do not lend themselves directly to quantification.

The purpose of the experiment described in this Report was to examine some variants of Delphi formats that might be more adaptable to issues of interest to decisionmakers in this area. The focus of the experiment was on methodological considerations, rather than on the actual content of the questions asked.

In two Delphi rounds of questioning, a group of volunteers (whose work was not professionally related to drug research) were asked to complete a questionnaire of items relating to drug use, prevention, and treatment. The questionnaire consisted of two parts dealing with descriptive items, and one part which required an evaluation of the feasibility and desirability of various "policy options." For most questions, repondents were asked to rate their level of confidence in the answer they supplied.

For questions whose multiple choices were non-quantitative, a number was associated with each choice for the purpose of measuring consensus levels. Response trends were summarized numerically according to a scheme which gave the greatest weight to responses with a high associated confidence self-rating.

Feedback furnished on the second round of questioning included graphical displays summarizing response trends and levels of consensus among the participants on the first round, together, in some cases, with a verbal summary of the reasons which respondents gave for their answers. Participants were invited to provide new assessments on the second round of questioning in light of numerical and verbal information. Although some convergence behavior was observed with practically all questions on the second round, the degree of convergence differed substantially among items.

Confidence ratings can be useful in this form of Delphi questionnaire, both as a mechanism for weighting answers, and in providing an overall assessment of the quality of the group response. However, certain difficulties are associated with the use of confidence scores; these are discussed in Section IV of this Report.

Drug-related questions are particularly prone to ambiguity and multi-dimensionality. In light of this fact, a useful feature of Delphi is the capacity to modify or decompose questions on the basis of first-round responses.

If Delphi is to be used in an actual study of drug-related problems, the questionnaire should be designed with a specific set of decision problems in mind. Various suggestions relating to the implementation of Delphi in future drug studies are included in Section V of this Report.

CONTENTS

LIST OF TABLES

I. INTRODUCTION

In the summer of 1970 the author conducted a pilot study to estimate the feasibility and usefulness of employing Delphi techniques as a forum for examining drug-related issues.* This work was initiated with the thought that it could be incorporated into a large-scale study of drug-abuse prevention programs in New York City.

The experiment discussed in this Report deviates in format and perspective from most Delphi studies that have been conducted in the past. Because interest in earlier studies has been focused on the quality of group responses generated in the feedback process, questions in those studies were selected for which a unique (usually numerical) "true" answer existed. In the drug field, most important questions of interest are not directly quantifiable--and for those few that are, answers are generally not known with any precision. It was clear in the early stages of the study, therefore, that we would have to experiment with new questionnaire formats if we hoped to develop a modified Delphi methodology which could be helpful to decisionmaking in the drug-abuse area. Since empirical validation of group responses would have been impossible for most questions of interest, we concentrated instead on issues of questionnaire design, presentation of feedback, and observed convergence behavior for various categories of questionnaire items.

With this perspective, we designed a three-part questionnaire on issues relating to drug abuse, administered in two rounds of questioning to 23 professionals and non-professionals interested in urban problems in New York City (but not directly involved in studies of drug abuse). Because we were specifically concerned with *methodological* considerations, items on the questionnaire are in no way representative of questions that would be appropriate or useful in a more formal study involving a panel of "experts" in the drug field. Moreover, *the actual numerical values which arose should not be interpreted as reliable or authoritative estimates.*

---

*For a more complete description of Delphi techniques, see Refs. 1 and 2.

II. THE EXPERIMENT

QUESTIONNAIRE DESIGN

In considering the sorts of questionnaire items that would be candidates for a formal Delphi study, we found it convenient to distinguish between items that are essentially *descriptive* (questions, for example, pertaining to the mechanism through which an individual becomes a drug user), and those which are primarily *evaluative* (such as items involving the relative merits of alternative drug prevention programs). Descriptive questions, in turn, decompose naturally into two sub-categories: those for which the information sought is inherently numerical ("how many heroin addicts are there in Bedford-Stuyvesant"), and those for which it is not. (The boundary between these categories is admittedly somewhat vague. A "descriptive" question on the recidivism rate in therapy programs would clearly be relevant to an overall *evaluation* of such programs, for example.)

The pilot questionnaire (cf. Appendices) is accordingly divided into three parts. The multiple choice items in Part I are descriptive and non-numerical. Although we were not always successful in this effort, we attempted to design questions for this part with "uni-dimensional" answers. By invoking this requirement, we can in principle define a mapping from the set of choices for a particular item into the real line; the images of such a mapping can be used in estimating consensus and convergence levels for the group (this is described in more detail below). For all questions in Part I we associated with each of the four choices an integer between one and four; these assigned values appear in sans serif type to the left of each choice in Appendix I.

Part II consists of evaluative items in the form of "policy objectives" for which respondents are asked to furnish two ratings. The Desirability rating, on a scale from -3 to +3, is a measure of the degree to which the respondent would be pleased or displeased if the objective in question were achieved. The Feasibility rating, on the other hand, reflects the relative ease with which the objective could be implemented,

and is closely related to the likelihood of its coming to pass in the near future; the feasibility scale runs from 0 (not feasible) to 3 (very feasible). Presumably the objectives which are most promising in the eyes of the evaluators are those with high scores on both desirability and feasibility.

Part III, finally, consists of descriptive items whose correct numerical answer is unknown. For each of the Part III items, respondents were asked to supply a judgmental median and an interquartile range.

A critical feature of Delphi procedures is the facility for discriminating among levels of knowledge or confidence. Various methods can be designed for achieving this. In a recent experiment involving short-range forecasts of events in national and international affairs, for example, T. A. Brown[3] attempted to select elite sub-groups on the basis of their responses to a multiple-choice current events test. We chose instead to accompany each question in Parts I and II with a "confidence" self-rating, an approach which has been used in past Delphi exercises.[1, 2] The confidence scale in our questionnaire runs from a low confidence of 0 ("this is a sheer guess") to a high rating of 3 ("I would be prepared to defend this opinion before a professional audience"). Confidence ratings were omitted in Part III, on the assumption that levels of knowledge would be reflected in the "tightness" of judgmental distributions, and hence in the distance between low and high judgmental quartiles. (This assumption may not be justified, a point which will be discussed later.)

In drawing up a set of questionnaire items, we were particularly interested in comparing the effect of feedback under contrasting circumstances. A question such as 1-1 (Addictive properties of heroin), we anticipated, would generate a fairly high level of consensus on the first round. Is there, then, an advantage in including "easy" questions like this in a Delphi questionnaire, and if so, do we obtain added knowledge from a feedback round? Our feeling was that such a question, while almost certainly not generating any new information, can prove helpful to respondents in answering other items on the questionnaire,

and in addition may help identify individuals whose answers systematically deviate from the group norm.

In contrast with these ostensibly straightforward questions, we included a number of items (e.g., Legalization of heroin for addicts) for which we felt there would be high disagreement on the first round. One of our objectives was to compare convergence behavior in such situations with that on less controversial issues.

We also selected questions which we felt would be of varying familiarity to respondents. While virtually everyone has given some thought to the prospect of legalizing marijuana, for instance, practically no one had considered immunizing school aged children against heroin.* Our *a priori* hypothesis was that opinions on unfamiliar issues would tend to be more pliable than those on very familiar ones.

OUTLINE OF THE DELPHI PROCEDURE

Three interviewers administered the first-round questionnaire to 23 participants; of these, 19 also completed the second-round questionnaire. Tabulated results in this report are based on this subset of responses, which decompose into the following categories:

	Professionals in Urban Research	Others
Male	6	3
Female	5	5

Approximately four-fifths of the respondents were between 23 and 35 years of age.

After marking his answer to a question in Parts I or II, the respondent was asked to explain why he held that view, and to offer information which he considered relevant to the question. This dialogue was carried on informally, and comments were transcribed by the interviewer for subsequent use in feedback. Except for the instructions,

---

* Since the time of the experiment, this issue has been given some attention in the press.

Part III required no assistance from the interviewer; however, he remained present while the respondent completed this part of the questionnaire.

A week after completing the first questionnaire, participants were given the Round 2 questionnaire. Except for some minor editing and the presence of feedback information, the second questionnaire was identical to the first. Respondents completed the second version on their own, and were not asked to supply reasons for their answers. They did not have records of their own first round responses.

FEEDBACK

An underlying premise of the Delphi approach is that, by and large, a respondent is better equipped to answer a question if he has some information on how other individuals have responded to the same question. One can conceive of various forms which this feedback information might take; roughly speaking, however, these can be classified as either "numerical" or "verbal." In the pilot study, we experimented with both numerical and verbal feedback. The latter was included on some questions in Part I, and for all Part II items.

Unless the number of participants in a Delphi exercise is extremely small, designing verbal feedback always poses a dilemma. If one includes all comments from all participants, the volume of feedback rapidly becomes prohibitive and its function self-defeating. On the other hand, the editing of first-round data must necessarily be somewhat arbitrary. When opinions are aggregated and condensed, certain participants will inevitably (and sometimes justifiably) feel that their opinion has not been adequately represented in the edited feedback version. The question of validity also poses a difficult problem. If a respondent opposes the legalization of marijuana with the authoritative statement (which he may believe to be true) that "a recent Johns Hopkins study has established a causal link between marijuana smoking and gangrene," should this reason be included in feedback?

These are nontrivial questions. In the experiment, we adopted the policy of (a) combining and summarizing opinions, generally in the form of one-sentence statements, and (b) editing out only those statements of scientific "fact" which appeared to draw on outside authority, but for which the respondent could not cite experimental evidence. Since this latter criterion is somewhat vague and may be influenced by the Delphi moderator's subjective attitudes about the issues involved, this sort of editing would probably not be desirable in an actual study. In such a setting, of course, participants would presumably be experts in their field and probably *would* be able to cite specific references in such cases.

The set of reasons supplied on most questions could be dichotomized, and they were accordingly grouped into two columns representing opposing views in the feedback questionnaire.

Some form of numerical feedback was supplied for all items on the questionnaire. Now if we were able to isolate the "well-informed" answers, it is reasonable that we would want to place the greatest (or exclusive) emphasis on these responses in generating feedback. Unfortunately, we have no crystal ball for facilitating this sorting-out procedure. What we do have at our disposal is a set of confidence self-ratings; the hope, of course, is that these ratings reflect to some degree the quality of the answers they accompany.

There are a number of ways in which we could discriminate among answers, using confidence ratings as a guide. Since the confidence scale runs from 0 to 3, one procedure would have been to use only answers accompanied by a confidence level of 3 to generate a feedback statistic. In many cases, however, this would have given exclusive influence to an elite corps of three or four individuals--a situation which we wished to avoid. At the other end of the spectrum, it did seem reasonable not to allow answers accompanied by a 0 confidence rating ("this is a sheer guess") to influence feedback data. The procedure we selected was a linear weighting scheme which, in effect, gave each response 0, 1, 2, or 3 "votes" in correspondence with the associated

confidence rating. On each question, these votes were accumulated for every possible choice and then normalized; the resulting proportions appeared in the feedback questionnaire as italicized percentages to the left of each choice in Part I, and as histograms in Part II.

Since individuals were asked to furnish judgmental quartiles* for the numerical items in Part III, a reasonable choice for feedback data would be quartiles of the group's medians, or alternatively medians of the quartiles. It happened that the two sets of statistics obtained from our data were quite close (cf. Appendix D), and we chose to supply quartiles of the mid-estimates. No verbal feedback was provided for Part III, since respondents were not asked to explain their answers on the first round.

There is no direct way of inferring the effect of verbal (as compared with numerical) feedback from our data. In various informal discussions after the experiment, however, many of the subjects of this study said they had been more impressed by the numerical statistics than by the summary arguments provided in Parts I and II.

---

*As used in this Report, the term "quartiles" will be used to denote the .25-, .50-, and .75-fractiles of a distribution. These are also referred to as the lower quartile, median (or mid-estimate), and upper quartile, respectively.

III. DIFFICULTIES

MODIFICATIONS BETWEEN ROUNDS

One of the benefits of sequential questioning rounds is that one can make use of first-round responses to make modifications in the questionnaire itself before returning it to respondents. One might want to do this either to (1) sharpen the format of existing items, or (2) include subsidiary or related items which may have been suggested by responses on the first round (including, in particular, "verbal" feedback).

Observations of first-round behavior prompted a number of minor changes in the questionnaire between rounds. These were motivated by the first consideration mentioned above rather than the second (since our experimental interest was in methodology rather than in the content of questions *per se*).

As an example of such a change, consider question 1-5. On Round 1 it was posed as follows:

> It has been suggested that drugs fulfill a need for anti-social behavior among youngsters in ghetto areas. If drug prevention efforts were successful, it is _____ that these youngsters would redirect their energies toward other forms of destructive activity (e.g., gang wars).
>
> [ ] Almost inevitable
> [ ] Probable
> [ ] Improbable
> [ ] Extremely unlikely

The first statement was intended to be purely motivational. In a number of cases, however, individuals either disagreed with this introductory premise, or failed to see its connection with the question being asked. It became clear that the first sentence was confusing and not essential to the second, and it was omitted in the second questionnaire.

The order of Part I items on the first-round questionnaire resulted in a different sort of modification. When participants reached the

"relative threats" question (subsequently labeled 1-6), they had not yet reached the questions which asked them to compare the addictive properties of speed and heroin. In fact, several participants asked the interviewer such questions explicitly (but were not given answers). It was therefore advantageous to precede 1-6 by 1-1 and 1-2, and this change was made on the second round.

A number of changes were made in Part II. The first version of the questionnaire included the following policy objective:

> Isolation of addicts from the rest of society (e.g., through the establishment of "drug communities").

This objective was receiving so many interpretations that it confused even the interviewers by the end of the first round, and it was therefore dropped from the questionnaire.

Item 2-1 was originally phrased as

> Massive step-up of narcotics enforcement efforts on the part of police and customs officials.

A number of individuals felt the need to distinguish between enforcement efforts directed at the pusher and those aimed at the addict. A few respondents also asked whether narcotics included "soft" drugs. The item was changed on the second questionnaire to read as:

> Massive step-up of efforts on the part of police and customs officials, aimed at reducing the *trafficking* of hard narcotics.

Although respondents were asked for both a Desirability and a Feasibility rating for policy objectives, they were asked for only a single confidence score. In reality, confidence on the two ratings were sometimes in conflict; participants were therefore informed by the interviewer that the confidence rating applied only to Desirability. To reduce ambiguity, two confidence ratings were requested on Round 2, although Feasibility confidence scores were not used in the data analysis.

A final change involved a question in Part III. Question 3-2 requested judgmental quartiles on the year in which marijuana will be legalized. In the opinion of one respondent, there was a chance greater than one in four that this event would never take place. An option for this opinion was provided in the second questionnaire. (An alternative way to handle this question would have been to ask a respondent for his judgmental probability that marijuana would *ever* be legalized--and then request quartiles of the legalization date, *given* that it would be legalized sometime in the future.)

OTHER SOURCES OF AMBIGUITY

An important lesson to be learned from our experiment is that constructing a meaningful, well-posed, unambiguous question in the drug area is by no means a trivial task. This is attributable in part to an inherent "non-scientificness" of the field, which compounds the usual difficulties encountered in questionnaire design.

We ran into a number of difficulties in the questioning process which merit brief mention here, although they did not result in actual alterations of the second-round questionnaire. One of these problems concerned the instructions for Part III, which included a concise explanation of judgmental quartiles in terms of conditional probability (cf. Appendix D). To our frustration, the only individuals who needed no additional clarification of the directions were those who had worked with judgmental fractiles in the course of their research, and fully understood the concepts already! Because subjects in a number of experiments at The Rand Corporation and elsewhere reportedly have had no difficulty in grasping the concept of judgmental fractiles, the problem here presumably lies in the explanation itself. Perhaps the explanation should be left out altogether: subjects might simply be asked to complete statements such as:

> There is a 25 percent chance that the true value is below _____.

The more serious ambiguities in questions were discussed in the preceding section. Almost every item posed some difficulty for someone. In questions 1-1 and 1-2 on the addictive properties of heroin and speed, the ambiguity for some was in the actual choices offered. Two or three individuals decided not to check the "no one" choice because "out of a hundred million people, there's bound to be some character who's immune to almost anything." We had not intended this interpretation, and if a similar item is included in a future questionnaire, the "no one" option should be replaced by, say, "an extremely small percentage of the population."

In question (1-3), Marijuana and sexual experience, one respondent felt the issue was dichotomous. While grass can heighten physical sensation, he maintained, the emotional component is lessened, leading to a self-oriented experience. What, he asked, is meant by the all-encompassing term "sexual experience"?

Interestingly, there were few interpretive difficulties in Part III. In particular, no one raised the issue of dichotomy in question 3-5 ("What percent of individuals who have used heroin at least once found it an unpleasant experience the first time?").

IV. SUMMARY AND DISCUSSION OF DIRECT FINDINGS

DEVELOPING A MEASURE OF CONSENSUS

An important objective of the Delphi experiment was to determine the degree to which feedback affects convergence of opinion. If we were observing behavior on a single question alone, we might be able to make an intuitive judgment of whether or not the group exhibited "significant" convergence by observing trends in responses between rounds. When we are dealing with more than one question, though, we would like some numerical measure of convergence with which to contrast behavior on different items.

Although the choices for questions in Part I are, in raw form, non-numerical, we associate with each choice an integer between one and four (five in the case of question 1-3), as discussed in Section II. One way to summarize data from either round for a question from Part I or Part II, then, is by means of the mass function f(i), where f(i) is the (possibly weighted) number of votes cast for choice i on that question. One method for simulating a continuous distribution is to assume that for each i, the f(i) votes are distributed in an equi-spaced manner along the interval $[i - \frac{1}{2}, i + \frac{1}{2}]$; and this is the procedure we have adopted. Thus, for example, if f(2) = 11 for a particular item, we represent these votes as an equi-spaced distribution with unit height on the interval [1.5, 2.5]; i.e., we assume that a single vote was cast at each of the eleven points 1.5, 1.6,..., 2.4, 2.5.

There are a number of ways in which we could measure the tightness of this adjusted distribution. One measure is the width of some inter-fractile range, i.e., the size of the smallest interval required to capture the central k percent of the votes, for some specified k. We have selected this family of measures, with k fixed at 50. We will call this measure the "interquartile width," and it is defined to be the difference between the .75- and .25-fractiles of the adjusted distribution.

Let us examine a few characteristics of the interquartile width (IQW). Suppose (to take an extreme case) that, for the Desirability of a certain policy objective, all n votes were cast for a rating of -2. Then the adjusted mass function consists of n equi-spaced points of unit height on the interval [-2.5, -1.5], and the IQW will be roughly .5. Suppose, on the other hand, that respondents are particularly polarized on the Desirability of a particular policy objective: assume, in fact, that all votes were cast for either a +3 or a -3 rating. Two cases can arise here. If min $\{\frac{1}{n}f(-3),\frac{1}{n}f(+3)\} \geq .25$, then the IQW will be between 5 and 6, reflecting the high level of controversiality. If the above condition on f(-3) and f(+3) is not satisfied, then the "dissenting" votes are swamped by the majority opinion, and the IQW will be of order .5; this is explained by the fact that the central half of the votes are contained in the interval [-3.5, -2.5] despite the presence of minority votes at +3. While the high sensitivity of the IQW to the value .25 in this particular case may be open to criticism, it should be kept in mind that this exaggerated polarization is an anomaly which does not usually have a counterpart in real situations. To consider a final example, suppose the adjusted distribution is uniform over the interval [-3.5, + 3.5]. This situation will result in an IQW of order 3.

It should be pointed out, finally, that we assume a roughly "linear" relationship between choices in the sense that, for example, the degree of disagreement reflected by a +3 versus a +1 Desirability rating is comparable to the degree of controversy manifested by a 0 and -2 rating. While we do not claim that such linearity is precise, we think it is an adequate representation to first order.

CONVERGENCE BEHAVIOR

We see immediately from the Appendices that all questions but one in the first two parts exhibit *some* convergence tendency between rounds. This in itself is an interesting result; we had speculated before conducting the experiment that particularly controversial issues might generate *divergent* behavior between rounds.

What is also remarkable about the results is the variety of distinct patterns that were generated.

We will consider first items which displayed a particularly high degree of convergence. The most dramatic member of this class is the Desirability rating for the Compulsory Injection policy objective (2-5) (see Table 1). Although more than 50 percent of the first-round (weighted) votes were cast for a rating of -3, a majority of the remaining votes fell in the plus range, making this item the second most "controversial" of the policy objectives on the first round, with an interquartile width of 4.71. In contrast with the first-round distribution, the second-round was the second *tightest*, with an IQW of 0.97. Nine out of an n of 19 furnished a lower desirability rating for this objective on the second round than on the first, while only two individuals raised their desirability rating (their ratings remained non-positive). Although we cannot draw definitive conclusions from this single question, it is perhaps significant that of all the questionnaire items, the Compulsory Injection issue appeared to be the least familiar to the greatest number of people. A likely explanation for the high degree of activity on this item is the fact that most individuals had no preconceived thoughts on the issue, and were therefore particularly susceptible to the influence of feedback.

The most pronounced convergence among the multiple choice questions in Part I occurred on items 1-6 (Relative threats) and 1-7 (Economic background). On the second round there was fairly high agreement that (a) speed poses a somewhat greater threat to the user and to society than heroin, and (b) youngsters from middle to high income backgrounds tend to be somewhat less difficult to treat in therapy programs than those from ghetto areas. Although the second round IQW for item 1-7 is only 0.58, the narrowness of the response spread is offset by an extremely low confidence level (the median first-round confidence level was 1.00), so that the consensus on this question is not particularly exciting. Confidence on the Relative Threats question runs somewhat higher, with a first-round median of 1.73.

Table 1

RATINGS FOR COMPULSORY INJECTION OBJECTIVE ITEMS (2-5)

Subject No.	1st Round Desirability	2nd Round Desirability	1st Round Confidence
1	2	-1	1
2	-3	0	1
3	-3	-3	2
4	0	-2	2
5	2	2	2
6	2	-2	2
7	-3	-3	3
8	-3	-2	1
9	-1	-3	2
10	-3	-3	3
11	-3	-3	3
12	-3	-3	3
13	-3	-3	3
14	-3	-3	3
15	2	-2	3
16	0	-3	2
17	3	-3	1
18	3	-2	3
19	1	-3	2

At the other end of the spectrum are two very controversial items in Part II which were minimally affected by feedback. The Desirability of objective 2-9 (Legalization of heroin for addicts) was characterized by a dramatic first-round IQW of 5.34 (the maximum possible IQW for Desirability is 6.0). Respondents were apparently not affected by the numerical or verbal feedback on this question, the second-round IQW remaining at a high 5.0. The question on police enforcement (2-1) exhibited similar properties. Although ambiguity between enforcement efforts aimed at pushers as opposed to users might have caused a first-round IQW as large as the 4.42 which was obtained, this apparently was not the determining factor, since the IQW remained at a high 4.39 even after the question was reworded. Evidently the legalization of heroin for addicts, and the massive step-up of police enforcement, belong to the "inherently controversial" category of issues. One would suspect that additional Delphi rounds would accomplish little in bringing group opinion on these matters closer together. Additional Delphi iterations could be used to articulate and evaluate the reasons used to support either side of the argument in such a case.

In terms of convergence, the most "stubborn" question in Part I was question 1-2 on the addictive properties of speed. This can probably be attributed simply to a lack of knowledge on the subject: the median first-round confidence was only 0.92, the lowest of the seven items in that part. The distribution is also characterized by the uninteresting second-round median of 2.36 (2.5 would be "dead center").

There is, finally, a class of items in Parts I and II which exhibit a high level of consensus on both rounds, and hence for which Delphi is not particularly helpful. Interestingly, one of these was the desirability of legalizing marijuana for individuals over 18 (item 2-6), which received a median second-round rating of +2.56, and an IQW of not more than 1.56 on both rounds. The Desirability ratings for items 2-10 (Public school instruction), 2-3 (Encounter sessions in youth centers), and 2-8 (Establishment of more therapy programs) also exhibited high consensus on the first round, but nevertheless succeeded in becoming

even more tightly distributed on the feedback round. Question 1-3 (Marijuana and sexual experience) displayed similar behavior in Part I.

We have so far not discussed convergence behavior among Feasibility ratings in Part II. There was some convergence tendency on the Feasibility ratings for all items except 2-1 (Public school instruction), which was already characterized by a relatively tight IQW of 1.05. One of the more active items was 2-1 on police enforcement. The drop in IQW from 1.70 to 1.22 was accompanied by a shift in median from 1.75 to 2.43. Unfortunately, it is impossible to determine whether this is the effect of feedback alone, or whether it pertains to the changed wording mentioned earlier.

Unlike the Desirability distribution for the legalization of marijuana objective (2-6)--which was already tight on the first round--the corresponding Feasibility IQW shrank from a relatively high 1.77 down to 0.96. Since the upper quartile of the distribution remained constant between rounds, the convergence is attributable entirely to an upward change in the lower quartile: initial skepticism seemed to diminish between rounds.

By and large, participants probably had least knowledge on the five numerical questions in Part III; using the IQW of the collection of medians as a measure of consensus, some convergence tendency can be discerned on all questions (cf. Appendix III). The most pronounced convergence occurred on question 3-2 ("in what year will marijuana be legalized?") in which the IQW shrank from 7 to 3.

Questions 3-3 ("What percent of 25-year old males in Central Harlem have used heroin at least once?") and 3-5 ("What percent of individuals who have used heroin at least once found it an unpleasant experience the first time?") were the only items in Part III for which the group median of individual mid-estimates experienced a change between rounds.

CONFIDENCE RATINGS

The use of confidence ratings involves more subtleties than we had anticipated. Conceptually, confidence scores in a Delphi process serve

two functions: to assign relative weights to responses for feedback purposes, and to provide a kind of overall assessment of the quality of the group response after feedback. In the case of the experimental questionnaire, participants were asked to rate their confidence on each item with an integer between 0 ("this is a sheer guess") to 3 ("I would be prepared to defend this statement in public"). In retrospect, it was not strategic to define zero as the lower confidence limit; zero apparently holds an unpleasant stigma for some individuals (seven participants did not rate themselves with zero for any item) which might have been removed if the scale had started at, say, one. But a more subtle problem presented itself during the experiment. The interviewers observed that those individuals who seemed actually to be more familiar with narcotics problems would frequently assign low confidence ratings because "through extensive reading I realize that this is an extremely complex question." This poses a dilemma: should the response of such an individual be processed in the same way as that of the individual who rates himself with low confidence because he has never heard the term "encounter session" before? My feeling is that it should not, and I would suggest as a possible remedy the inclusion of a "familiarity" rating. Although such a rating would not necessarily be used in generating feedback, it could be of substantial value in interpreting the end-product distribution of responses, particularly on numerical items of the type found in Part III. Alternatively, if the respondents were "experts" in the area of drug abuse and treatment, they might be asked a question such as "In your professional judgment, to what extent has this issue been settled in the field?"

In the Delphi experiment, respondents were asked for confidence ratings on both rounds of questioning. It was surprising to us that there was a sharp increase in overall confidence levels between rounds (see Table 2). The cause for this phenomenon is not clear. It may be due to a reinforcement effect of feedback, or it may be caused simply by the fact that respondents have given more thought to the questions by the time they are given the second questionnaire, and hence are more

"familiar" with the issues. These second-round confidence ratings may actually be introducing additional noise into the evaluating system -- and perhaps they should be eliminated altogether in this round.

Table 2

COMBINED CONFIDENCE TALLIES FOR PARTS I AND II

	Confidence Rating			
	0	1	2	3
Round 1	31	85	124	83
Round 2	8	64	162	89

Some crude analysis of the data suggests a general tendency for "extreme" answers (choices 1 or 4 on Part I, Desirability ratings of $\pm$ 3 in Part II, etc.) to be accompanied by high confidence ratings. For example, the calculation which treated each response by each respondent as a data point yields a sample correlation coefficient of 0.49, which is significantly different from zero at the .01 level. In a sense this is just what one would expect: people who make strong statements are usually confident in their point of view. Looking at the matter from another perspective, though, one might have anticipated that Desirability ratings in, say, the [-1, +1] range would be accompanied by high confidence -- since there is really nothing to be confident about (one might think of this as "fervent apathy"). This hypothesis was not supported by the data, however.

Because of the correlation between "extremeness" and confidence, the general effect of weighting was to push the distributions away from center. Weighting the Desirability distribution never decreased the percentage of $\pm$ 3's, and in fact never decreased normalized membership in the {-3, -2, +2, +3} category. Beyond this polarization effect, weighting did not materially affect the percentage of votes falling on either side of middle.

FEASIBILITY AND DESIRABILITY

The data suggest that Feasibility and Desirability scores are not independent. A correlation of these two ratings across all questions and all respondents, for example, yields a sample coefficient of 0.42 (significantly different from zero at the .01 level). A plausible explanation for this may be that individuals tend to project their personal preferences onto the community at large: "If *I* feel that a particular objective is undesirable, then so will the community, and it will not gain the widespread popular acceptance necessary for implementation."

V. IMPLICATIONS FOR FUTURE RESEARCH

The most challenging aspect of future applications of Delphi techniques to the drug field will almost certainly be the design of a cohesive set of questionnaire items that are both well-posed and useful to the decisionmaker. On the one hand, it became apparent during the study that developing concise questions which will be given similar interpretations by all respondents will inevitably involve substantial pre-testing. The usual difficulties in questionnaire design are compounded in the context of drug abuse by disagreement over underlying assumptions, and by the absence of an agreed-upon vocabulary.

Equally important is the incorporation of Delphi techniques into a larger decisionmaking framework. Because our orientation in this study was methodological, we were not directly concerned with the implications of a set of group responses for policy decisions. But if Delphi is to be used as an aid to decisionmaking in the area of drug abuse, it will be crucial to have a clear idea of how the results of such an exercise will be used, before engineering the questionnaire and Delphi procedure itself.

One can think of a variety of contexts in which an administrator might make use of Delphi procedures. At one level, a new agency head might be interested in comparing the attitudes and biases in his own staff with trends in the research community at large. To reduce the potential distorting effects of direct, personal interaction to a minimum, he might choose to employ a Delphi approach simply as a means of surveying the attitudes of his staff. The spread of group opinion might also point to issues which appear to be "inherently controversial." If such results were coordinated with experts' assessments of those areas which are most crucial for research in the immediate future, this information might be used to reorient the direction of future studies in the drug area. It might be useful, in this context, to include in the Delphi questionnaire a list of researchable issues; respondents could be asked to rate the relative priorities for research in these areas from the point of view of the policymaker as well as to estimate

the current state of the art for each one. Delphi would be a particularly appropriate arena for this, since a list of research areas could actually be generated by experts in the first round of questioning, to be later returned to the group for evaluation.

Delphi is also promising as a forum in which policy questions can be compared and evaluated. Part II of the pilot questionnaire illustrates one way in which this evaluation could be carried out, but there are a number of possible variations on that approach. From the point of view of decisionmaking, it may be more useful to force individuals into a comparison of alternative approaches, rather than to allow them to award the highest rating to a large number of proposals. Such comparisons might be made in the context of certain global objectives (e.g., maximizing the effectiveness of public school programs).

In another setting, Delphi might be used more directly as an input to the policymaking process. In situations in which a decisionmaker must make choices under uncertainty, for instance, he might use a set of group responses to update his prior estimates on the probabilities of alternative states or hypotheses. In other words, he would treat the results of a Delphi procedure as experimental data. This application would require the decisionmaker to assess the probability of obtaining a given response pattern, given each of the possible states of hypotheses. While this may be a fairly straightforward procedure for numerical items (such as those in Part III of the experimental questionnaire), it is not immediately clear just how one would extend this approach to a collection of opinions on policy objectives.

The question of how and whether a Delphi exercise could be helpful in aiding a decisionmaker in the drug area to choose among alternative courses of action in some specific context would be an interesting issue for further research. If, indeed, the responses of a panel of experts can be useful in this context, the question remains of how these experts would behave in an anonymous feedback situation. Presumably such a group would differ from the pilot group of this experiment in two important respects: their level of familiarity with drug-related problems, and their degree of commitment to fundamental assumptions regarding drug use.

Appendix A

QUESTIONNAIRE INSTRUCTIONS

GENERAL INSTRUCTIONS FOR ROUND 1 QUESTIONNAIRE

This questionnaire is part of a larger study designed to test the effectiveness of Delphi techniques as a means for collecting opinions on drug-related issues.

After this set of questions has been asked of a number of other respondents, we will return to you with a second questionnaire; in addition to containing a number of the questions you are being asked today, this will contain some information on the responses of other participants, as well as some additional questions. We expect that on both questionnaires there will be some items which you will not feel particularly well qualified to answer. However, we hope that you will try to furnish some answer to even the most difficult questions; you will be given the opportunity to rate your level of confidence on each one.

We request that you *not discuss the questionnaire with anyone* until the completion of the experiment.

Many thanks for helping us out.

GENERAL INSTRUCTIONS FOR ROUND 2 QUESTIONNAIRE

This is the second, and last, questionnaire on drug attitudes and familiarity. Most of the questions in this group will be identical to those which you answered previously.* This time, however, you will be told something about how other respondents answered each question. What we are asking you to do is provide revised answers to each question, having in hand this information on the responses of other participants.

On each item you will be given statistical information reflecting response trends on the first round of questioning. (In calculating these statistics, more weight was given to answers with a high associated confidence rating than to those with a low rating.) In addition to this numerical information, you will be given for certain items a list of "reasons" which were important to other individuals in answering the question.

We ask once again that you *not discuss this questionnaire with anyone* until the end of the experiment.

---

*Two or three questions have been slightly reworded.

Code Sheet (For Use on Parts I and II)

DESIRABILITY:

+3--Very desirable: I would very much like to see this implemented, and in my opinion it should be a top-priority objective

+2--Desirable: I would like to see this implemented, but it is not as important a policy objective as items rated +3

+1--Somewhat desirable; certainly a low-priority matter

0--I am indifferent on this matter

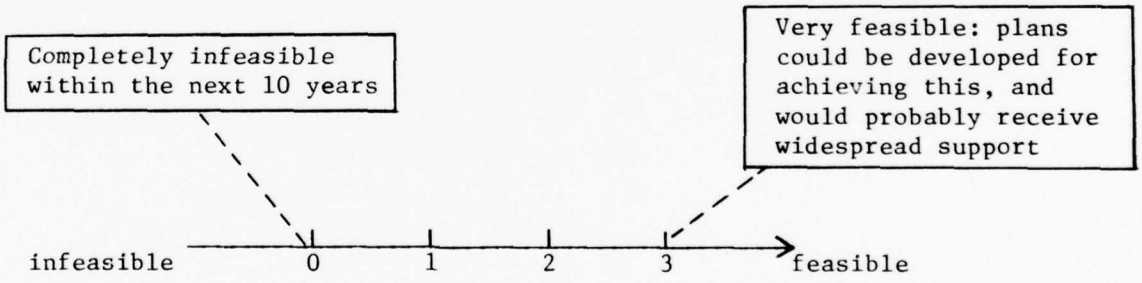
-1--Somewhat undesirable: I would not be particularly happy if this were implemented, but I do not consider it to be a particularly important issue

-2--Undesirable: I would definitely not want to see this implemented

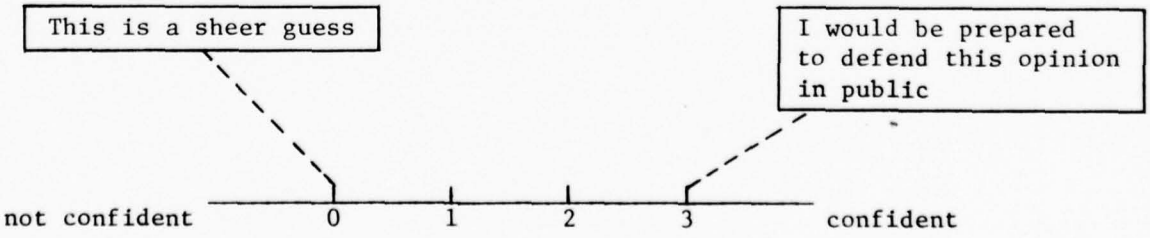
-3--Very undesirable: this is one of the worst things that could happen

---

FEASIBILITY:



CONFIDENCE:



Appendix B

DATA FOR PART I OF QUESTIONNAIRE

Data on the following pages pertain to the seven questions constituting Part I of the questionnaire; they are in the order in which they appeared in the Round 2 version.

The italicized percentages to the left of each choice represent the weighted response trends on Rounds 1 and 2, respectively. Percentages for Round 1 may differ slightly from those which appeared as feedback for the Round 2 questionnaire, due to the fact that the former were calculated on the basis of the 19 individuals who completed both rounds of questions, while feedback statistics were calculated using all 23 sets of first-round results.

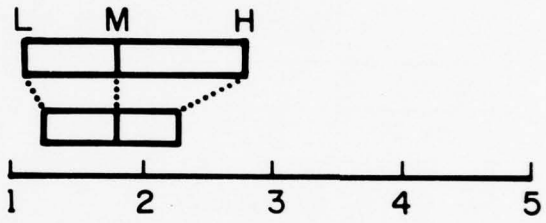
In the experiment, verbal feedback was furnished only for questions 1.6 and 1.7.

Data for Part I

1.1 [Addictive properties of heroin]
It is possible for _____ to use heroin several times a week for several months without becoming physically dependent on it.

Round 1 Responses*	Round 2 Responses	
42%**	33%	1 No one
28%	53%	2 Only a few individuals
21%	14%	3 Many individuals
10%	0%	4 Almost anyone

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	1.10	1.81	2.79	1.69
Round 2	1.25	1.82	2.29	1.04

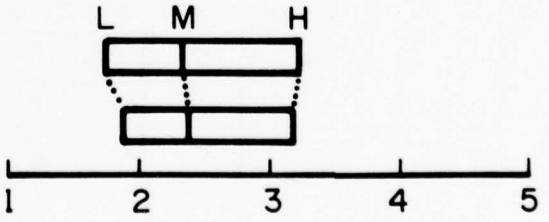


Median Confidence Rating:
Round 1 = 1.43
Round 2 = 1.92

1.2 [Addictive properties of speed]
It is possible for _____ to take high doses of "speed" (methedrine) several times a week for several months without becoming physically dependent on it.

Round 1 Responses*	Round 2 Responses	
14%**	6%	1 No one
43%	51%	2 Only a few individuals
24%	26%	3 Many individuals
19%	17%	4 Almost anyone

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	1.75	2.33	3.25	1.50
Round 2	1.88	2.36	3.19	1.31



Median Confidence Rating:
Round 1 = .92
Round 2 = 1.88

---

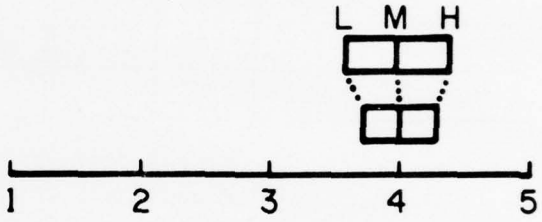
*Adjusted to reflect confidence self-ratings. See Section II of this Report.

**Percentages may not add due to rounding.

1.3 [Marijuana and sexual experience]
As a rule, sexual experience is _____ by marijuana.

Round 1 Responses*	Round 2 Responses	
0%**	0%	1 Greatly diminished
15%	8%	2 Somewhat diminished
6%	0%	3 Unaffected
61%	85%	4 Somewhat enhanced
18%	8%	5 Greatly enhanced

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	3.56	3.98	4.39	.83
Round 2	3.71	4.00	4.29	.58

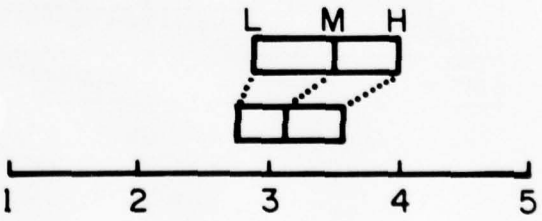


Median Confidence Rating:
Round 1 = 1.75
Round 2 = 2.14

1.4 [Anti-marijuana arguments]
It is _____ that most arguments that have been made to youngsters about the dangers of marijuana have done more to make them suspicious about warnings against the use of any drug than to dissuade them from using marijuana.

Round 1 Responses*	Round 2 Responses	
3%**	0%	1 Extremely unlikely
8%	10%	2 Improbable
40%	63%	3 Probable
50%	28%	4 Almost certain

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	2.88	3.50	4.00	1.12
Round 2	2.74	3.14	3.59	.85



Median Confidence Rating:
Round 1 = 2.11
Round 2 = 2.15

---

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

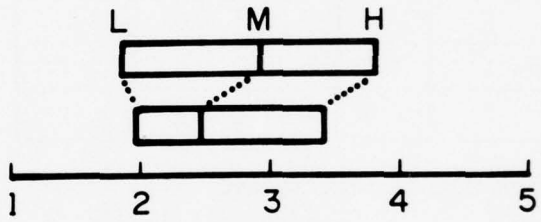
**Percentages may not add due to rounding.

1.5 [Redirection of energies]
If drug prevention efforts were successful, it is _____ that youngsters in ghetto areas would redirect their energies toward some form of destructive activity (e.g., gang wars).

Round 1 Responses*	Round 2 Responses	
17%**	0%	1 Extremely unlikely
23%	52%	2 Improbable
23%	26%	3 Probable
37%	23%	4 Almost inevitable

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	1.86	2.93	3.82	1.96
Round 2	1.98	2.47	3.41	1.43

Median Confidence Rating:
Round 1 = 1.43
Round 2 = 1.69



*Adjusted to reflect confidence self-ratings. See Section II of this Report.

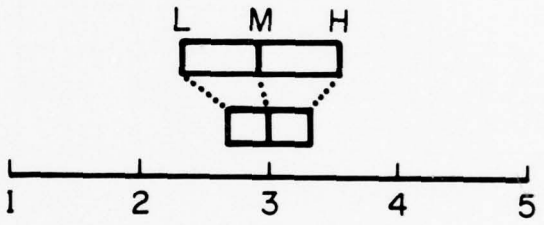
**Percentages may not add due to rounding.

1.6 [Relative threats]
The use of "speed" (methedrine) in the absence of a medical prescription poses a _____ threat to the user and to society than the use of heroin.

Round 1 Responses*	Round 2 Responses	
0%**	0%	Much smaller
32%	12%	Somewhat smaller
42%	76%	Somewhat greater
26%	12%	Much greater

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	2.28	2.92	3.53	1.25
Round 2	2.67	3.00	3.33	.66

Median Confidence Rating:
Round 1 = 1.73
Round 2 = 1.85



VERBAL FEEDBACK

--The heroin user will rob to support his habit.

--Unlike heroin, speed can cause permanent deterioration of the nervous system.

--While the heroin user is lethargic, the speed user is hyperactive and unpredictable.

--There is a chance that speed will be mixed with other chemicals (deliberately or inadvertently), with very detrimental effects.

--There are legal ways to handle the heroin situation, but there is nothing you can do for permanently damaged individuals.

---

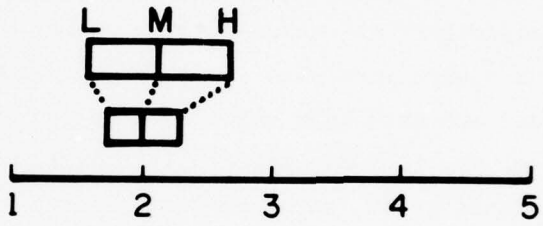
*Adjusted to reflect confidence self-ratings. See Section II of this Report.

**Percentages may not add due to rounding.

1.7 [Economic background]
Heroin addicts from middle to high income backgrounds tend to be ____ to treat in therapy programs than addicts from ghetto areas.

Round 1 Responses *	Round 2 Responses	
21%**	7%	Much less difficult
47%	85%	Somewhat less difficult
32%	7%	Somewhat more difficult
0%	0%	Much more difficult

	QUARTILES OF RESPONSES*			
	LOW	MED.	HIGH	IQW
Round 1	1.58	2.11	2.71	1.13
Round 2	1.71	2.00	2.29	.58



Median Confidence Rating:
Round 1 = 1.00
Round 2 = 1.43

VERBAL FEEDBACK

--The notion of therapy is less alien to youngsters from higher economic backgrounds.

--Higher income individuals have more opportunity and hence more motivation than lower income people.

--Lower income individuals have a relatively negative attitude toward health in general.

--Ghetto residents are easier to get through to than higher income people because they are more familiar with addiction and its effects.

--Heroin use on the part of higher income youngsters is apt to be symptomatic of more serious psychic problems.

--Higher income youngsters know exactly what they're doing when they go in for drug use; they are therefore less motivated to kick the habit.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

**Percentages may not add due to rounding.

Appendix C

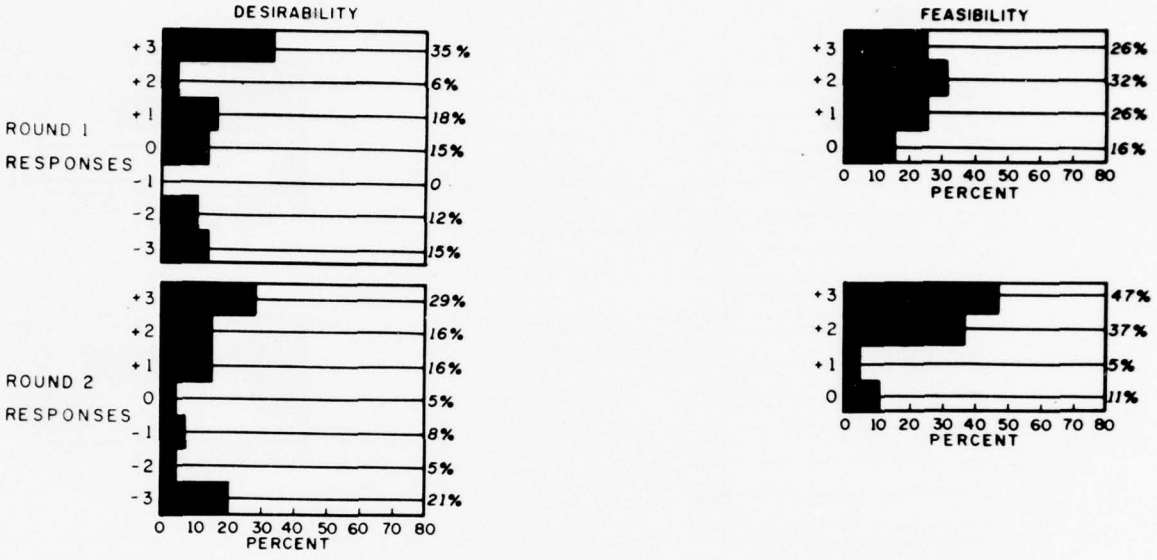
DATA FOR PART II OF QUESTIONNAIRE

Data on the following pages pertain to the ten policy objective items constituting Part II of the questionnaire; they are in the order in which they appeared in the Round 2 version.

Response trends on the Desirability rating, as illustrated by the pair of histograms, were calculated using the linear confidence weighting scheme described in Section II, while response trends for Feasibility are unweighted. Similar histograms (for Round 1 data only) were furnished as feedback in the Round 2 questionnaire. Since feedback data were calculated on the basis of 23 questionnaires, however, they may differ slightly from those which appear in this Appendix; the latter were calculated on the basis of the 19 individuals who completed both rounds of questions.

Verbal feedback, as it appeared on the Round 2 questionnaire, is included for all questions.

2-1. [Police enforcement]
POLICY OBJECTIVE: Massive step-up of efforts on the part of police and customs officials, aimed at reducing the trafficking of hard narcotics.



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	-1.63	+1.00	+2.79	4.42
Round 2	-1.75	+1.17	+2.64	4.39

FEASIBILITY

LOW	MED.	HIGH	IQW
0.85	1.75	2.55	1.70
1.75	2.43	2.97	1.22

Median Confidence Rating for Desirability:
Round 1 = 1.75
Round 2 = 2.07

VERBAL FEEDBACK

--Increased enforcement would decrease the flow of drugs.

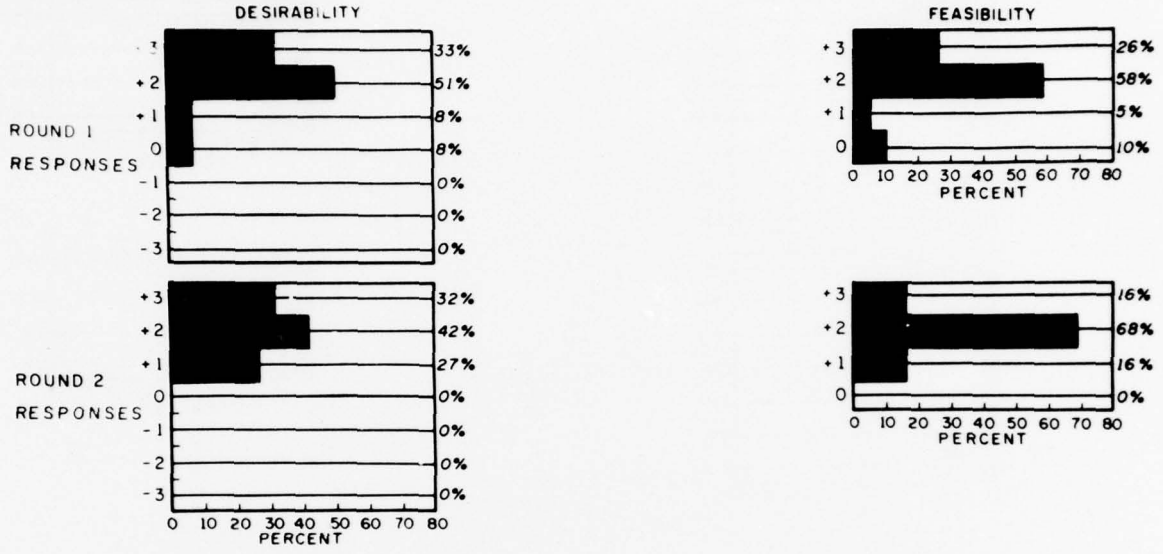
--Enforcement reassures large segments of the society who are afraid of drugs.

--We will never control narcotics with the law: drugs are too profitable and too easily made to be controlled.

--Increased enforcement will merely serve to push the price up.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-2. [Youth centers]
POLICY OBJECTIVE: Establishment of more community youth centers in ghetto neighborhoods.



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+1.69	+2.18	+2.75	1.06
Round 2	+1.43	+2.06	+2.71	1.28

FEASIBILITY

LOW	MED.	HIGH	IQW
1.66	2.09	2.55	0.89
1.64	2.00	2.37	0.73

Median Confidence Rating for Desirability:
Round 1 = 2.15
Round 2 = 2.18

VERBAL FEEDBACK

--Social outlets for youth are important.

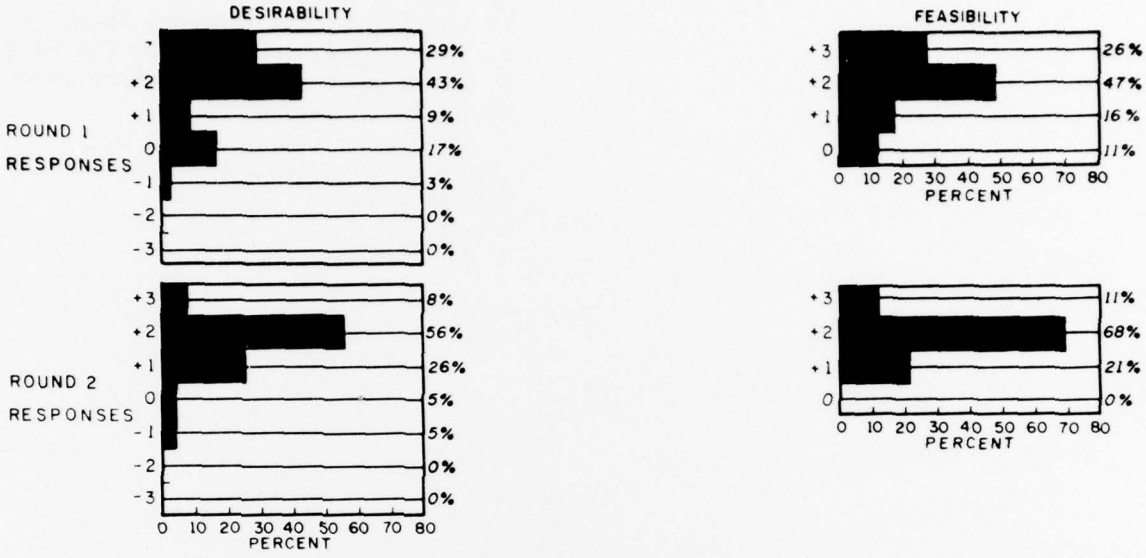
--Youth centers provide more help than the families of drug users.

--Youth centers don't solve the youngster's problem, which is getting a job and getting out of the ghetto.

--It is difficult to raise money necessary to run youth centers well.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-3. [Encounter sessions in youth centers]
POLICY OBJECTIVE: Regularly scheduled group encounter sessions in community youth centers.



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+1.08	+2.00	+2.62	1.54
Round 2	+1.07	+1.75	+2.19	1.12

FEASIBILITY

LOW	MED.	HIGH	IQW
1.42	2.00	2.55	1.13
1.56	1.92	2.29	0.73

Median Confidence Rating for Desirability:
Round 1 = 2.06
Round 2 = 2.04

VERBAL FEEDBACK

--Encounter sessions can be useful in helping kids deal with their problems.

--Only "Good Guys" would come to encounter sessions.

--Encounter sessions are characterized by a certain glibness which may not entail any real meaning or depth.

--In encounter sessions individuals simply attack each other to relieve a sense of failure.

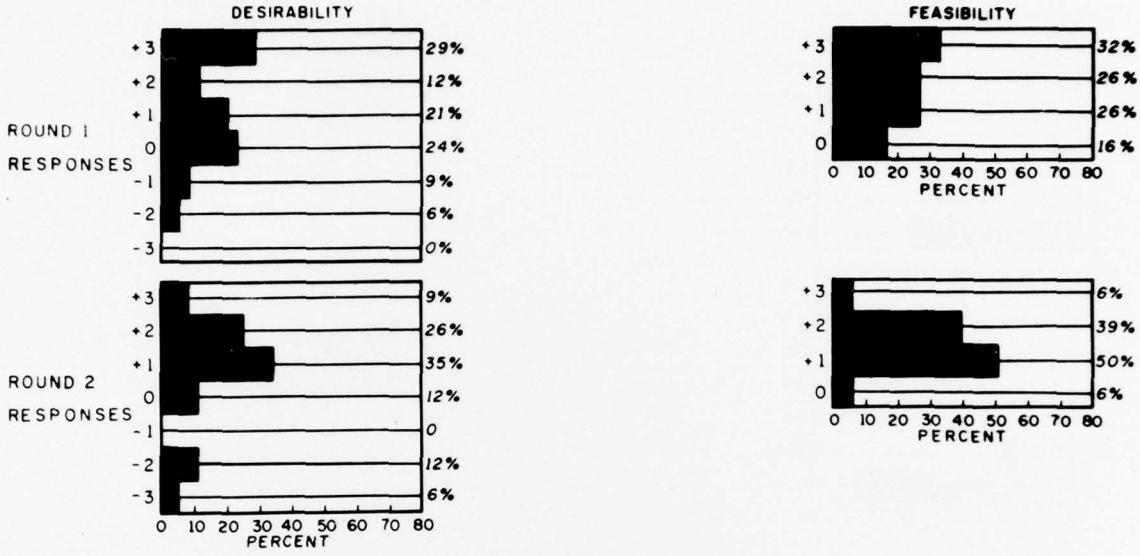
*Adjusted to reflect confidence self-ratings. See Section II of this Report.

VERBAL FEEDBACK (CONT.)

--The use of encounter sessions as a prevention effort assumes there is a psychological component to the drug problem; most youngsters just take drugs for kicks.

--Youth centers should be a place of happiness.

2-4. [Encounter sessions in public schools]
POLICY OBJECTIVE: Encounter sessions as an extracurricular activity in public schools.



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	-0.06	+1.07	+2.65	2.71
Round 2	+0.13	+1.08	+1.89	1.76

FEASIBILITY

LOW	MED.	HIGH	IQW
0.85	1.80	2.71	1.86
0.89	1.38	2.00	1.11

Median Confidence Rating for Desirability:
Round 1 = 1.83
Round 2 = 1.93

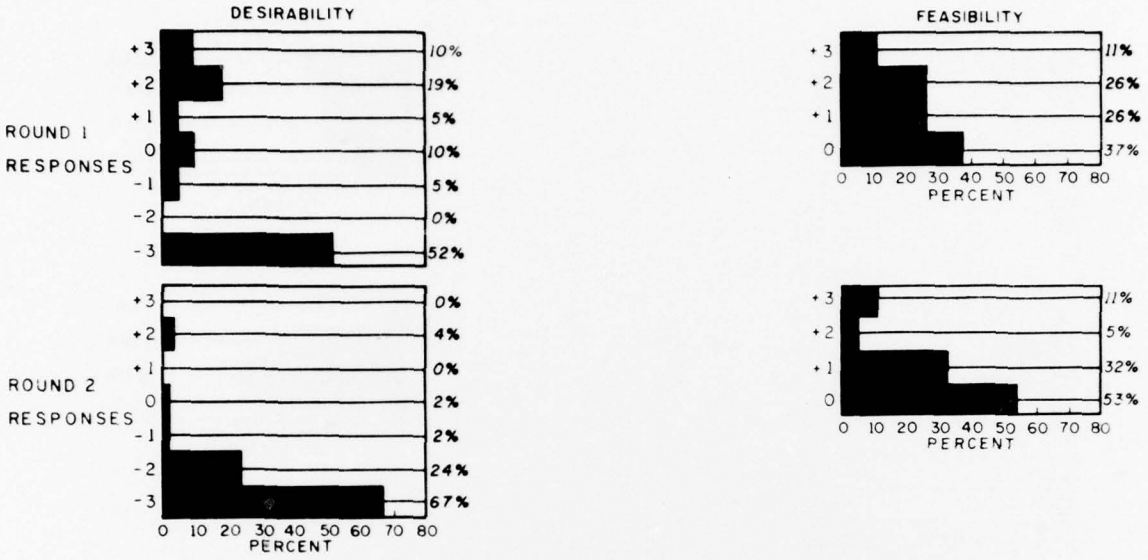
VERBAL FEEDBACK
[ALSO REFER TO STATEMENTS ON PRECEDING QUESTION]

--If you hold encounter sessions as an extracurricular activity in schools, you're just preaching to the converted.

--If encounter sessions are held as an extracurricular activity in schools, those students in the sessions who are in fact addicts will become either idolized or despised.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-5. [Compulsory annual injection]
POLICY OBJECTIVE: Compulsory annual injection of all school-age children with a chemical which would block the effect of heroin for a one-year period. (Assume such a chemical exists and assume it has no injurious side-effects.)



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	-3.02	-2.55	+1.69	4.71
Round 2	-3.13	-2.75	-2.16	0.97

FEASIBILITY

LOW	MED.	HIGH	IQW
0.18	1.00	1.95	1.77
-0.03	0.45	1.21	1.24

Median Confidence Rating for Desirability:
Round 1 = 2.29
Round 2 = 2.42

VERBAL FEEDBACK

--If you can keep kids off heroin when they're young, then because of a "maturing out process" they won't use it when they're older.

--Chemical prevention of heroin use is an infringement on the student's personal freedom.

--The drug problem has profound social roots; is there a chemical solution to a social problem?

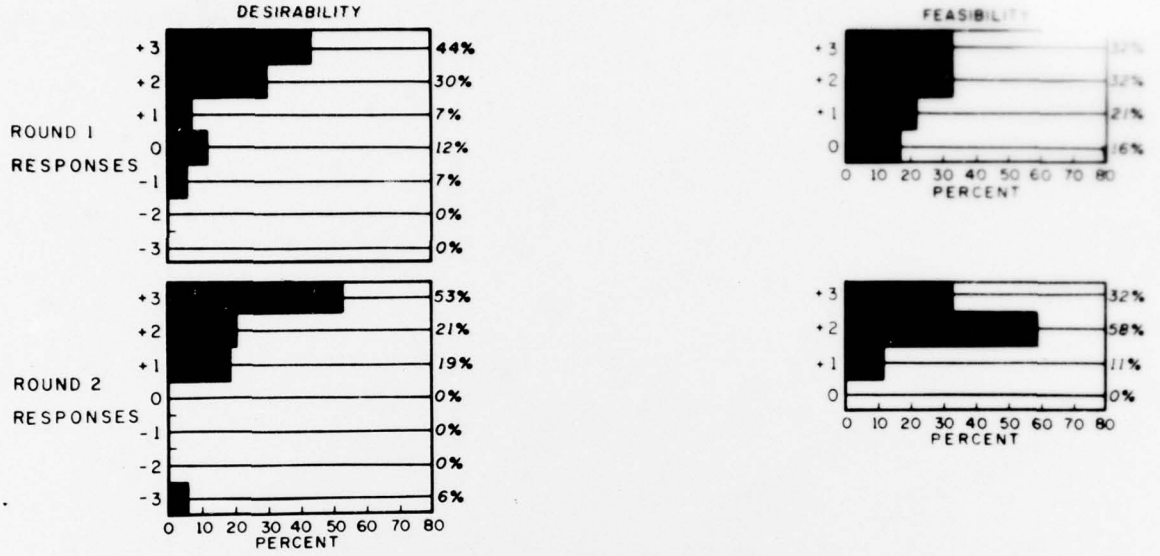
--Individuals should have the choice to destroy their minds and bodies if they want to.

--Unnecessary chemicals should be kept out of children's bodies.

--Chemical prevention of heroin use sets a dangerous precedent.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-6. [Legalization of marijuana]
POLICY OBJECTIVE: Legalization of marijuana for individuals



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+1.42	+2.31	+2.93	1.51
Round 2	+1.47	+2.56	+3.03	1.56

FEASIBILITY

LOW	MED.	HIGH	IQW
0.94	1.92	2.71	1.77
1.75	2.18	2.71	0.96

Median Confidence Rating for Desirability:
Round 1 = 2.55
Round 2 = 2.45

VERBAL FEEDBACK

--The anti-marijuana law represents a social bias against the young and increases their alienation.

--All evidence indicates that marijuana is less dangerous than alcohol.

--As many people should be kept out of the criminal system as possible.

--The status of "getting away with it" may spread to other crimes through a hero-worship of the law breaker.

--There is a danger of submersion in a drug culture "full of kids nodding" if marijuana is legalized.

--Marijuana is an entryway into the use of hard drugs.

--Marijuana is a crutch.

--There is no control over people driving under the influence of marijuana.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

--It is easier to go from marijuana to hard drugs because of the illegality linkage between the two.

--The law should not impose on an individual's liberty if it can be avoided.

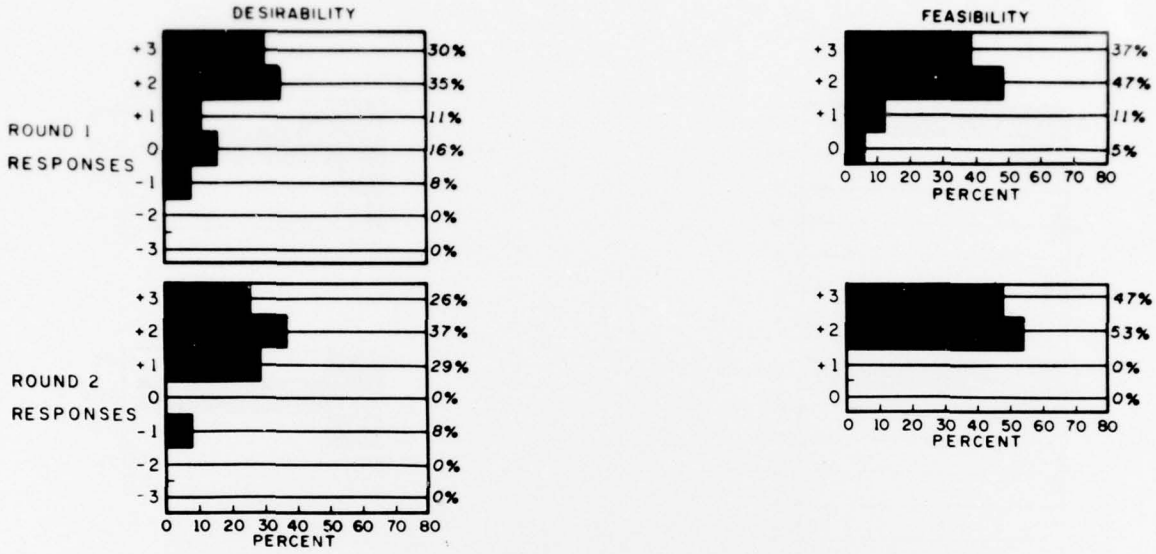
--Many cultures where marijuana is legal have become rather devitalized.

--Kids need something safe to rebel against.

--Not enough is known about the side-effects of marijuana.

2-7. [More methadone programs]

POLICY OBJECTIVE: Increase the number of methadone programs in New York City.



*QUARTILES OF RESPONSE**

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+0.56	+1.92	+2.66	2.10
Round 2	+1.09	+1.86	+2.55	1.46

FEASIBILITY

LOW	MED.	HIGH	IQW
1.69	2.22	2.82	1.13
1.98	2.45	2.97	0.99

Median Confidence Rating for Desirability:
Round 1 = 1.96
Round 2 = 2.00

VERBAL FEEDBACK

--There are long waiting lists for methadone, and it should be made available for anyone who can benefit from it.

--It is better that someone be hooked on something harmless than on something harmful like heroin.

--Methadone programs would reduce the crime associated with heroin addiction.

--Methadone may not solve the problem, but you can treat psychiatric problems later.

--People can return to a somewhat normal life via a methadone program.

--Methadone leaves people with a psychic problem: the need for a needle is still there.

--There will start to be a black market in methadone.

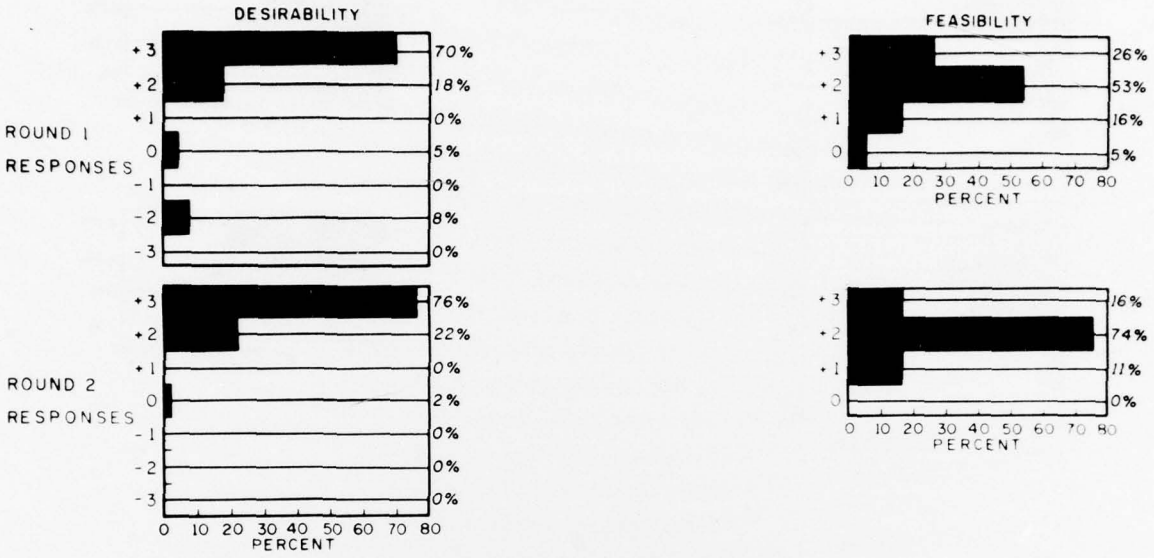
--The side-effects of methadone haven't been sufficiently explored.

--Methadone doesn't create a sense of responsibility in individuals; psychologically they haven't made the transition back into the real world.

--With methadone you end up with an individual who's blocked on heroin but addicted to something else.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-8. [More therapy programs]
POLICY OBJECTIVE: Increase the number of therapy programs (such as Phoenix House).



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+2.21	+2.79	+3.14	0.93
Round 2	+2.51	+2.84	+3.17	0.66

FEASIBILITY

LOW	MED.	HIGH	IQW
1.58	2.05	2.55	0.97
1.70	2.04	2.38	0.68

Median Confidence Rating for Desirability:
Round 1 = 2.05
Round 2 = 2.55

VERBAL FEEDBACK

--There are long waiting lists for therapy programs.

--Therapy programs suit the need for many addicts who want to get off.

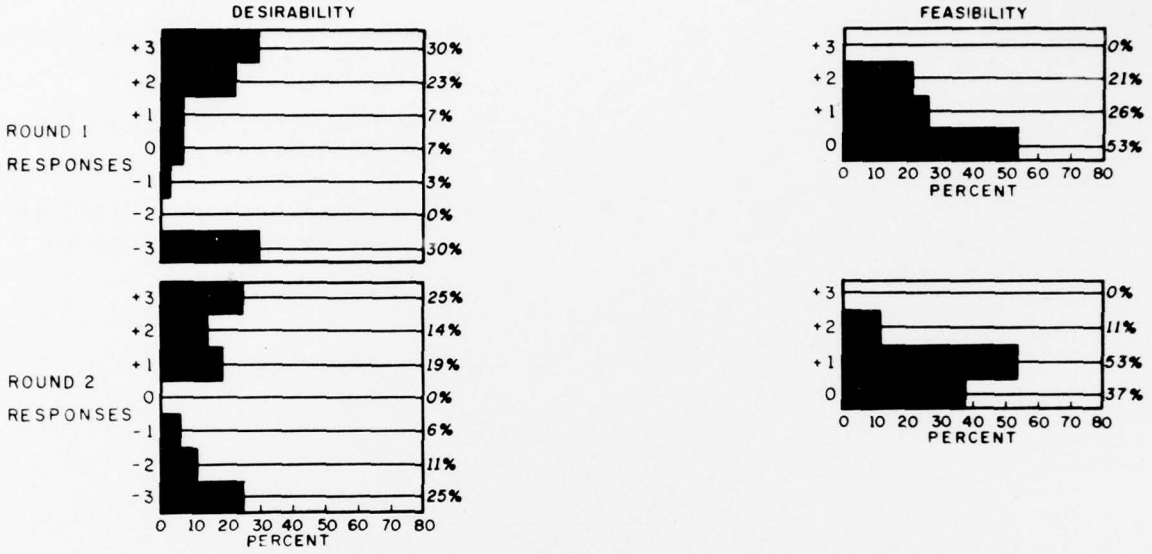
--Therapy program success rates are high enough to justify expansion in this area.

--So many people drop out of therapy programs so early that they are not justifiable from a cost-effectiveness standpoint.

--There aren't enough well-qualified people around to run therapy programs.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

2-9. [Legalization of heroin]
POLICY OBJECTIVE: Legalization of heroin for addicts (under medical supervision).



QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	-2.67	+1.64	+2.67	5.34
Round 2	-2.50	+0.93	+2.50	5.00

FEASIBILITY

LOW	MED.	HIGH	IQW
-0.03	0.45	1.35	1.38
+0.18	0.75	1.23	1.05

Median Confidence Rating for Desirability:
Round 1 = 1.58
Round 2 = 1.71

VERBAL FEEDBACK

--Legalization of heroin would break up the criminal aspect of addiction.

--If heroin were legalized, money which is now used in enforcement efforts could be reallocated in more constructive directions.

--Legalization of heroin is a more direct attack on the addiction problem than other alternatives.

--People should decide their own destinies.

--Legalization of heroin isn't fair to addicts; it would be better for them if they could get along without such an addictive drug.

--Legalization of heroin might perpetuate addiction rather than cure it.

--There is a risk that addicts will give themselves an overdose, whether or not heroin is made legal.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

VERBAL FEEDBACK (CONT.)

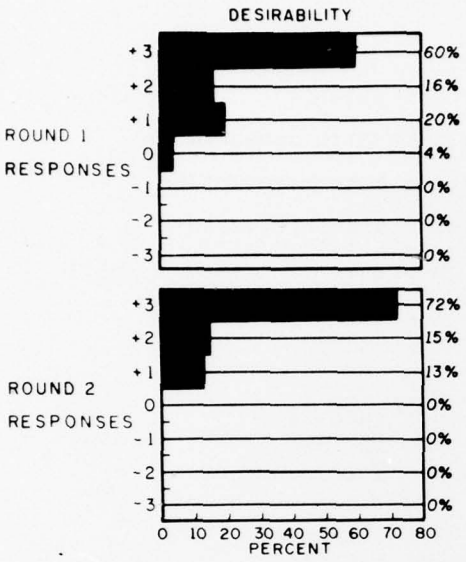
--Legalization of heroin would reduce the incentive for treatment programs.

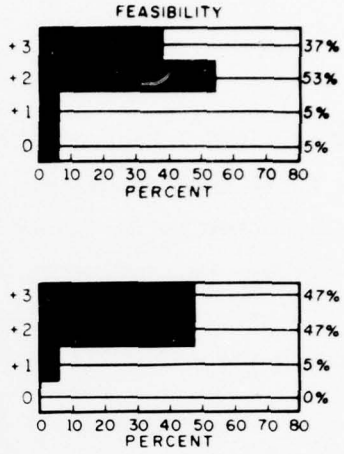
--Methadone offers an alternative to the legalization of heroin.

--Addicts can become non-addicts through treatment.

--Distinction between addicts and non-addicts may be difficult.

2-10. [Public school instruction]
POLICY OBJECTIVE: Public school instruction on the physical effects of drugs.





QUARTILES OF RESPONSES*

DESIRABILITY

	LOW	MED.	HIGH	IQW
Round 1	+1.54	+2.67	+3.08	1.54
Round 2	+2.29	+2.82	+3.15	0.86

FEASIBILITY

LOW	MED.	HIGH	IQW
1.77	2.25	2.82	1.05
1.92	2.44	2.97	1.05

Median Confidence Rating for Desirability:
Round 1 = 2.55
Round 2 = 2.55

VERBAL FEEDBACK

--Instruction on the physical effects of drugs provides the opportunity to make an enlightened choice.

--School instruction is the one way in which peer group pressure can be counteracted.

--Some people are impressed with scientific information.

--You have to instruct individuals about drugs somehow, and public school is the best public system.

--Instruction is a greater deterrent to drug use than telling youngsters after they're hooked.

--There are not many people who know enough about drugs and relate well enough to students to do a good job of instruction.

--Public school teachers get resentful when they have to take on another course.

--Young people may be so turned off to authority that school instruction wouldn't get to them.

--Instruction will inevitably become involved with rhetoric.

--Instruction on drugs is analogous to instruction on V.D.; many people aren't affected by it.

*Adjusted to reflect confidence self-ratings. See Section II of this Report.

Appendix D

DATA FOR PART III

INSTRUCTIONS*: Each of the numerical questions in this part appeared on the first questionnaire. This time you will be given some information on how other participants responded to these items. This information may or may not influence your answers to a given question.

After reading the italicized information for a given item, please furnish the following three assessments:

Mid-Estimate. This should be a number such that, in your opinion, the true value is just as likely to be below it as above it.

Low Estimate. *Suppose an individual knew the true answer, and told you that it was in fact lower than your Mid-Estimate.* Think of a number such that you would feel (given this added information) that the true answer was just as likely to be below it as above it. This is your Low Estimate.

High Estimate. Now suppose an individual knew the true answer, and told you that it was in fact higher than your Mid-Estimate. Think of a number such that you would feel (given this added information) that the true answer was just as likely to be below it as above it. This is your High Estimate.

---

*These instructions were supplied with the Round 2 questionnaire.

Table D.1

DATA FOR PART III

Question No.	Round	Quartiles of Mid-Estimates[a] Low	Med.	High	IQW of Mid-Estimates
3-1	I	35,000[c]	75,000	100,000	65,000
	II	45,000	60,000	95,000[c]	50,000
3-2	I	1973	1976	1980	7
	II	1975	1976	1978	3
3-3	I	32.5[c]	50	60	27.5
	II	35[c]	40	60	25
3-4	I	57.5[c]	75	85	27.5
	II	65	75	80	15
3-5	I	17.5[c]	50	67.5[c]	50
	II	37.5[c]	45	55[c]	17.5

[a]These numerical values are presented for the purpose of analyzing the Delphi method, and should not be interpreted as reliable estimates.

[b]These data were calculated using the responses of the 21 individuals who furnished complete sets of Part III answers on both rounds. The statistics may therefore differ slightly from feedback data, which were calculated on the basis of 23 completed questionnaires.

[c]Interpolated value.

REFERENCES

1. Dalkey, Normal C., "The Delphi Method: An Experimental Study of Group Opinion," The Rand Corporation, RM-5888-PR, June 1969.

2. Brown, B., S. W. Cochran, and N. C. Dalkey, "The Delphi Method II: Structure of Experiments," The Rand Corporation, RM-5957-PR, June 1969.

3. Private discussion with T. A. Brown.