Sanctions are officially imposed punishments aimed at enforcement of legal obligations. They are said to constitute the core, if not the defining characteristic, of the legal order. Inadequate sanctions are blamed for the failures of legal control in such divergent areas as international law, domestic crime, and civil rights. Despite the presumed importance of sanctions in the legal process, however, serious attention has rarely been paid to the topic.¹

This paper examines sanctions from the perspective of the social sciences. It consists of two parts: a brief review of existing knowledge relevant to legal sanctions, and a preliminary report of a field experiment aimed at generating more information on the subject. The first part begins with an oversimplified account of conventional wisdom in the area of sanctions, uses some research findings to show the inaccuracy of many of these conventional beliefs, and suggests the complexity of the problem. The second part narrows the questions to be examined, describes the research technique employed in trying to answer some of these questions, and gives initial results from that research. Both parts are presented here in a frank effort to elicit suggestions which will make the final report more useful as a social scientific contribution to jurisprudence.²

¹ ARENS & LASWELL, IN DEFENSE OF PUBLIC ORDER (1961).
² Writers in jurisprudence have long stressed the importance of sanctions, but often without systematizing the questions in this area. Notable efforts to deal systematically with the topic range from BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION (1823) to CARLSTON, LAW AND ORGANIZATION IN WORLD SOCIETY (1962).

For a cogent appeal to social scientists to attend to the issues of jurisprudence, see Harry Jones, A View from the Bridge, 39 Social Problems 41-42 (1965).
Notwithstanding the frequency with which the concept of sanction is used, there exists at present little empirical knowledge as to how, when, and whether sanctions work. For the most part, we rely on conventional wisdom in this area. Many legislators and administrators speak and act as if they believed some very primitive assumptions about sanction:

1. That it tends to reduce the frequency with which punished acts will be performed in the future by the offender and others.
2. That its effectiveness is an increasing function of its severity.
3. That it can be employed without incurring substantial costs to the society.
4. That other modes of increasing compliance need not be considered as alternatives, or at least not as incompatible alternatives, to punishment.\(^3\)

A brief survey of information bearing on these assumptions casts doubt on each of them. It also accentuates the scattered and inadequate nature of the evidence available and points to the need for relevant research.

**Preventive effect of sanctions on the offender and others.** The preventive effects of punishment on the offender are open to question. High rates of recidivism, in regard to crimes of property in particular, suggest that prison terms do little more than delay the recurrence of crime for one-third of the prisoners.\(^4\) The period in prison serves in many instances to alienate the offender further from society and to train him systematically in techniques for escaping future punishment.\(^5\)

---

\(^3\) These assumptions do not exhaust all the ideas concerning sanction effectiveness currently held by practitioners. In addition, policy makers seem to be guided by the belief that sanction impact increases with such elements as advance knowledge and clarity of the norm, compatibility with other norms, and the certainty of enforcement. These and other conditions—all presumably heightening the legitimacy of the norms—have been formally included in the literature of jurisprudence, most recently in Fuller, *The Morality of Law* 33-94 (1964). Although Fuller seems primarily concerned with moral imperatives, his eight conditions may be read as hypotheses bearing on the probable effectiveness of legal sanctions.


imposes on him, moreover, a stigma which tends to impede his economic and social adjustment upon release.6

In civil cases, other mechanisms may lighten the effect of sanction. In medical malpractice cases, for instance, doctors routinely insure themselves against the threat of court action7 and deduct the cost as a business expense.8 Even those convicted of malpractice benefit from the sympathy of their colleagues, so that in one recent study doctors found liable for malpractice reported that their practices had increased after losing their cases, because of referrals from sympathetic colleagues.9 Businessmen, potential targets of litigation, are usually able to avoid court action despite technical breaches of contract because of the culture and control system of the business community.10 When they are found liable, they apparently react to such sanction—following Mr. Justice Holmes' advice11—less as moral condemnation than as a business cost to be paid in return for the privilege of breach.

Similarly, the deterrent effect is open to question. In interviews conducted informally some years ago, a few instances were found in which tax violations began after the prosecution of a widely publicized case. The primary reasons given were that the convicted offender had been incredibly stupid and that his evasions had been of major proportions. "If that is the kind of thing the government waits for," said one informant, "they'll never come after me."12 Deterrence is supposed to work through fear of punishment. But if enforcement reduces fear of punishment, there is reason to suppose that the deterrent influence may

---


7 A nationwide survey of 20,000 physicians, conducted in 1963 by the Law Department of the American Medical Association, indicated that 94.3% of the 14,000 respondents had liability insurance. AMA Law Department, 1963 Professional Liability Survey, 189 A.M.A.J. 859, 862 (1964).

8 Deduction for malpractice insurance, apparently assumed as a legitimate business expense for private practitioners, was extended to company physicians under Rev. Rul. 60-365, 1961-2 Cum. Bull. 49.

9 Schwartz and Skolnick, supra note 6, at 138-39.


vanish or even be replaced by an incentive. Nor is it certain that fear of punishment deters. Andenaes twice searched the literature for evidence that it did and came up with very little the first time, substantially more the second.\textsuperscript{13} One of his best finds was a report on crime rates in Copenhagen before, during, and after a seven month period in which the police were rendered ineffectual because disarmed and re-organized by the German army of occupation. During these seven months "without police," robbery and larceny in Copenhagen apparently increased tenfold, whereas embezzlement, fraud, and homicide rates remained virtually constant.\textsuperscript{14} This suggests that the perceived danger of punishment serves as an immediate deterrent to the first type of crime but not to the second. It is difficult to interpret these findings with reference to a theory of sanction. The stability of crime in the second category might mean that such crime is not primarily deterred by fear of sanction, or it might mean that fear of detection and sanction remains high even in the absence of police because such crime can be readily detected without police help. If the first explanation carries any weight, it raises questions as to the circumstances in which fear of sanction explains compliance. Instead of accepting the blanket generalization that deterrence works universally (or the converse, that it never works), it seems more realistic to assume that sanction deters some violations under some circumstances. If this is so, it falls to the social scientist to try to specify the conditions under which deterrence works and, if possible, the intervening variables which explain why it does or does not.

\textit{Severity of sanctions.} Historical materials, case studies, and experimental evidence raise serious doubts about the proposition that sanction effectiveness is a linearly increasing function of severity of prescribed punishment. For one thing, when the threat of severe punishment is extended to a variety of offenses which are not considered serious hazards to community welfare, the penalties are unlikely to be enforced. In the classic instance of this, British law in the eighteenth century specified nearly two hundred capital offenses, including the stealing of goods valued at five shillings from a shop, at forty shillings from a dwelling place, and so forth. Juries reacted to these severe penalties by regularly specifying, in the face of sworn testimony to the contrary, that the stolen goods were worth one shilling less than the amount which would make


the act a capital offense. Parliament abolished the death penalty on all but four offenses by 1861, in important part because of pressure from commercial interests, whose avowed purpose was to secure reliable enforcement of property laws by making their penalties more acceptable to juries.\textsuperscript{15} Even where severe punishment can be effectively administered, it may not deter. Orsted describes a period in eighteenth century Norway when homicide rates increased because of the belief of suicidal persons that execution for murder was the only technique of suicide which might excuse them from eternal damnation. This led to the abolition of the death penalty in such cases on the ground that it acted as an incentive to, rather than a deterrent from, murder.\textsuperscript{16}

The literature of contemporary psychology also casts doubt on the severity assumption. In general, psychologists studying animal learning agree that punishment only reduces the likelihood of the punished behavior occurring when it contributes to the learning of alternative responses.\textsuperscript{17} If the punished individual for some reason repeats the punished response after punishment (R $\rightarrow$ Pun $\rightarrow$ R), that response may subsequently become more likely as a result.\textsuperscript{18} Under other conditions, punishment may temporarily inhibit the response it follows, but the response increases in frequency following the cessation of punishment.\textsuperscript{19} The more severe the punishment under such circumstances, the greater the motivation for the punished response. Even where a socially approved alternative is available, severe threat may not be the most effective way of promoting its adoption. This has been suggested by Janis and Feshbach in an experiment which shows that behavioral change can be greater under mild than under extreme threat.\textsuperscript{20} Subsequent

\textsuperscript{15} CANTOR, CRIME AND SOCIETY 222-24 (1939).
\textsuperscript{16} I1 ORSTED, EUNOMIA, 1815-22, 147.
\textsuperscript{17} SKINNER, SCIENCE AND HUMAN BEHAVIOR 182-93 (1953). The leading experiment is Estes, An Experimental Study of Punishment, 57 PSYCHOLOGICAL MONOGRAPHS 263 (1944).
\textsuperscript{18} Experimental findings on the subject are slim but suggestive. The leading animal experiment is reported by Muenzinger and Pawloski, Motivation in Learning—Comparison of Electric Shock for Correct Turns in a Corrective and a Noncorrective Situation, 42 J. EXPERIMENTAL PSYCHOLOGY 118 (1951).
\textsuperscript{19} Azrin, Effects of Punishment Intensity During Variable-Internal Reinforcement, 3 J. EXPERIMENTAL ANALYSIS OF HUMAN BEHAVIOR 123 (1960). More recently, the same experimenter reported data which suggested that punishment accomplishes this effect by raising the level of aggression. Thus, the observation might hold particularly for responses which are hostile toward others or destructive toward self. Azrin, Motivational Aspects of Escape From Punishment, 8 J. EXPERIMENTAL ANALYSIS OF BEHAVIOR 31 (1965). We are indebted to Frank Zimring, a third year student at the University of Chicago Law School, for bringing this work to our attention, as well as for other valuable suggestions.
\textsuperscript{20} Janis & Feshbach, Effects of Fear-Arousing Communications, 48 J. ABNORMAL & SOCIAL PSYCHOLOGY 78 (1953).
experiments qualify this finding but do not eliminate the possibility hypothesized in the original research.

Social costs of punishment. Also open to question is the assumption that punishment can be employed without consideration of its costs to society. The sanctioned person, as well as those who identify with him, may become alienated from those who invoke or impose the deprivation. In civil matters, this can lead to an intensification of hostility between parties whose continued cooperation is necessary for their own interests and/or for the welfare of society. It was this danger that led Harry Shulman strongly to advocate non-adversary proceedings as a substitute for winner-take-all judgments in labor-management relations. The same effect figures importantly in the philosophy of the family courts, as advanced by such proponents as Judge Alexander. When the opponent is the entire society, as represented by the prosecution in criminal cases, antagonism generated by punishment can produce alienation from the society. This may lead to withdrawal from social participation and/or an increased tendency to strike back at the society in varying ways. Such acts may be directed specifically against the kind of compliance enforced by the punishment. In the case of taxation, for instance, it may take the form of individual evasion, a culture of non-compliance such as is represented by moonshining, or a socio-political movement such as Poujadism, focused on thwarting tax collection. In addition, alienation from the state can result in more general resistance ranging from political apathy to revolution.

This is not to say, of course, that the imposition of sanction inevitably produces such effects. But there is enough evidence that they can occur to merit attention by policy makers and theorists alike. We need to know the conditions under which sanction produces side effects which nullify its contribution to social order. To what segments and proportions of the population may sanction be applied without im-


24 For a thoughtful discussion of the phenomenon among youth, see Matza, Delinquency and Drift (1964).
pairing the equilibrium of the social system? What is the effect of perceived legitimacy in reducing the socially disturbing side effects of sanction? How does legitimacy interact with varying degrees and types of sanctions? What are the conditions—such as power distribution, sanction invocation, value consensus—which affect the sense of legitimacy in various sociocultural settings? This is only a sample of the questions which must be articulated and explored if we are to understand the range of consequences which flow from the administration of sanctions.

Alternative modes of securing compliance. It is interesting to note that sanctions are rarely weighed against alternative means of securing compliance. Our legal system contains very few instances in which people are explicitly rewarded for compliance, rather than punished for deviance. The exceptions are most likely to be found in the regulation of economic behavior through devices such as tax incentives for investment and subsidies for crop limitation. Other countries have used incentives such as bonuses for larger families. But for the most part, incentives have been left to non-legal and especially to non-governmental agencies, such as insurance companies which provide lower rates for accident-free drivers. It was this lack of emphasis on incentives in legal policy which led Swift's Gulliver to the following comment:

Although we usually call Reward and Punishment the two hinges upon which all government turns; yet I could never observe this maxim to be put in practice by any nation, except that of Lilliput. Who ever can there bring sufficient proof that he hath strictly observed the laws of his country for seventy-three moons, hath a claim to certain privileges, according to his quality and condition of life, with a proportionable sum of money out of a fund appropriated for that use: He likewise acquires the title of Snipall, or Legal, which is added to his name, but does not descend to his posterity. And these people thought it a prodigious defect of policy among us, when I told them that our laws were enforced only by penalties without any mention of reward.25

The disinclination to use explicit rewards for legal compliance can be variously explained. The subject is too complex to explore in detail here. We suggest, however, that for reasons of economy explicit rewards tend to be employed where only small segments of the population are supposed to be their recipients. Legal norms, as Simmel pointed out, characteristically provide minimal standards which a majority are expected to observe.26 For the violator, the appropriate

25 Swift, Gulliver's Travels 64 (Modern Library 1931).
On Legal Sanctions

consequence is a deprivation. It would be possible in principle to give rewards to violators when they cease their violations, but this would seem likely to provide a motive for violation. Extension of rewards to all who observe the law would be expensive, difficult to administer, and ineffectual if the recipients were too numerous. If incentives were limited to the exemplars of legality—as semiseriously proposed by Swift—it might carry the implication that obedience to law is optional rather than obligatory. Moreover, emphasis on strict and manifest compliance with law might jeopardize a major benefit of a minimal, negatively defined code of conduct: the freedom of the individual to innovate, an especially necessary element in a dynamic social order.

Some of the same considerations apply to the use of conscience as a motivational support of law. Although religious systems have included appeals to conscience, as in the Ten Commandments and Canon Law, this element is largely absent in secular legal systems such as our own. While some of our statutes contain preambles giving the purposes of the legislation, such statements are characteristically used more for interpreting the meaning of the law than for moral exhortation. The body of the statute itself is almost invariably confined to a description of the standard of conduct and to a matter-of-fact statement of the consequences which will befall the violator. Avoidance of appeals to conscience was noted by Holmes as an evolving characteristic of our legal system.\(^2\) It has been explained by Cahn\(^2\) and Fuller\(^2\) on the ground that conscience alone cannot adequately enforce conduct for the majority. This consideration does not explain, however, why conscience is not combined with the threat of sanction. Among the possible explanations are the difficulty of the making of a direct appeal to conscience by the state, and the implicit assumption (as yet inadequately tested) that a combination of conscience and threat reduces the effectiveness that either would have separately.

Heavy, if not exclusive, reliance on sanction rests only in part on the absence of alternatives. Some evidence has been noted above which points to the effectiveness of sanction as a means of enforcing compliance in certain circumstances. In addition to rational justifications for enforcing law through sanction, certain nonrational elements may also enter. The frustration-aggression hypothesis of the learning theorists may be interpreted as giving one such reason.\(^3\) Violation of

\(^2\) Fuller, The Morality of Law 3-32 (1964).
\(^3\) Dollard, Frustration and Aggression (1939). The original hypothesis asserted that frustration was a necessary and sufficient condition for aggression. It was later modified...
socially approved norms may be assumed to act as a frustration to many in the society. The imposition of legal sanction may serve as the most convenient socially approved expression of aggression against the source of the frustration. Sanction, in these terms, appears to be a more gratifying response that the provision of reward to non-offenders or to the offender when he complies with the standard. Another explanation, proposed by Freud in *Totem and Taboo*, interpreted the punitive response to illegality as reflecting in part an attempt by sanctioners to repress their own impulses for comparable deviant activity. Exploring the subject in another context, Freud suggested that sanction tendency has a second powerful, unconscious source:

> In our unconscious we daily and hourly deport all who stand in our way, all who have offended or injured us... Indeed our unconscious will murder even for trifles... it knows no other punishment for crime than death: and this has a certain consistency, for every injury to our almighty and autocratic ego is at bottom a crime of lese-majeste.\(^3\)

Ranulf argued that sanction is often irrational in the sense that it exceeds the social needs for effective control and that this reflects an inclination toward "disinterested cruelty."\(^3\) This impulse he found particularly characteristic of the middle class; he suggested, by way of explanation, the distinctively insecure social position of that stratum.

Whatever the merits of such explanations, and of many more like them, they have in common the view that punishment does not inevitably serve its manifest purpose. Yet the argument can only be sustained in the final analysis if alternative means of securing compliance can be demonstrated to be at least as functional as punishment. Only when such information is available can the persistence of punishment be legitimately scored as gratuitous and hence irrational.

II

If we take seriously the task of developing a model to predict the effects of law, one of our first jobs must be to study closely the impact of sanctions on legal compliance. Accordingly, we need to examine such questions as: (1) Does the threat of punishment deter? (2) How

---

\(^3\) Ranulf, *Moral Indignation and Middle Class Psychology* (1938).
On Legal Sanctions

and when does it accomplish its effects? (3) What side effects does it have? and (4) How does it compare with alternative reactions?

In order to approach these issues empirically, it was proposed some years ago that a field experiment be conducted on motivational factors affecting compliance with federal income tax laws.33 Recently it has become possible to carry out such a study.

Tax compliance was selected for several reasons: (1) The payment of taxes is one of the most widespread of all serious legal obligations in American society. More than a third of the total population is required by law to file federal income tax returns. Of those obligated, the overwhelming majority do so. As a result, it is possible to sample diverse groups in the population to examine the effect on compliance of various background characteristics.34 (2) Taxpaying is required at least annually, so that changes over time may be sequentially observed. (3) Compliance with tax law varies from full or even excessive compliance to serious evasion. It has been established that at least one-fourth of all taxpayers evade taxes to some degree.35 Frequency of violation has practical and methodological implications. In policy terms, it means a serious loss of revenue. According to the IRS sample audit of 1948, 1.4

34 Schmölders inferred from interviews in Germany that 90% of his subjects were "unfamiliar" with the possibility of tax dodging, businessmen and professionals being overrepresented among those who reported knowledge of such a possibility. Schmölders, Fiscal Psychology: A New Branch of Public Finance, 12 NAT'L TAX J. 340 (1959). He speculated that attitudes toward taxpaying may be related to the terms used in different language to describe tax: Steuer (support) in Germany and skat (common treasure) in Denmark may make for or reflect complaint attitudes; imposto (imposition), the opposite in Italy. Schmölders is now engaged in comparative research on these matters in Italy, Spain, England, and France.


The relationship between disincentive effects and compliance remains to be explored. The one empirical study of disincentive effects known to the authors was carried out in England where compliance appears high. Nevertheless, the disincentive effect of a sharply graduated income tax was virtually balanced by incentive effects, even among a segment of the population (solicitors and accountants) where tax awareness was doubtless unusually great. Break, Income Taxes and Incentives to Work, 47 AM. ECON. REV. 529 (1957). For a theoretical discussion of disincentives and other issues of tax policy, see Blum & Kalven, The Uneasy Case for Progressive Taxation (1953); cf. Strumpel, The Disguised Tax Burden: Compliance Costs of German Businessmen and Professionals, 19 NAT'L TAX J. 70 (1966).

35 Violation rates appear to be higher in the payment of state taxes, especially where the possibility of detection is slim. Groves, Empirical Studies of Income Tax Compliance, 11 NAT'L TAX J. 291 (1958).
billion dollars were lost in that year because of underestimations on personal returns of taxes owed to the Government. If the present study reveals legitimate and practical devices for increasing tax compliance, the government's gain could be substantial. Reduction in evasion may, furthermore, serve to distribute the tax burden more equitably and add to the sense of legitimacy on which the legal order presumably rests. From a methodological point of view, the pervasiveness of violations creates an opportunity for study. It means that, in principle at least, violators can be compared with those who comply, to determine the differences between them. It also opens up the possibility, exploited in this study, of examining the efficacy of different motivations through exposing taxpayers to various stimuli and observing the increases toward full compliance which result. (4) Taxpaying is behavior which can be described in detail by quantitative indices. As a result, increases in taxpaying can be rigorously described, and related they can be through powerful statistical tests to the independent variables. (5) Because income taxation is practiced in many countries and in localities within the United States, results of this study can be replicated for comparative purposes in other settings.

The tax project was explicitly designed to study theoretically significant variables. The primary objective was to determine the effect of sanction threats and to compare them with appeals to conscience as determinants of legal compliance.

The technique here employed for examining tax compliance was the field experiment. This method, more widely praised than used in the

---

86 Farioletti, Some Results from the First Year's Audit Control Program of the Bureau of Internal Revenue, 5 Nat'l Tax J. 65 (1952); see also BITTKER, FEDERAL INCOME ESTATE AND GIFT TAXATION 899 (3rd ed. 1964).

87 Generalizations arising from studies of tax compliance cannot, of course, be freely generalized to other areas of behavior. Taxpaying differs from other forms of legal compliance in the population affected, the predominance of rational motivation, the regularity and discreteness of the act of compliance, the publicity given to its enforcement, and so forth. The findings in this sphere should therefore be examined for other kinds of legal compliance before they are used as a basis for a general theory. The need for extensive replication is increased even further if the objective is a general theory of compliance extending beyond legal to non-legal control.

88 This choice was dictated by our inability to provide realistic incentives other than those that might arise from moral consideration.

89 For a detailed discussion of theoretical considerations in the design of the study, see Schwartz and Skolnick, Televised Communication and Income Tax Compliance, in TELEVISION AND HUMAN BEHAVIOR 155 (Arons & May eds. 1963).

90 Field experimentation appears to be a far more appropriate method for studying tax compliance than alternatives, such as analytic induction, survey method, and laboratory experimentation. Schwartz, Field Experimentation in Sociolegal Research, 13 J. LEGAL ED. 401, 403-8 (1961).
social sciences, is designed to bring to natural situations some of the precision of the laboratory. In pure form, it includes the full application of the classic experimental model. This means differential application of one or more independent variables to experimental groups and the non-application of these variables to one or more control groups. The experimental and control groups are selected in such a way as to be initially comparable with each other in every way. The experimental groups are then subjected to stimuli which constitute the independent variables. The observer must be in a position to observe changes in the behavior of all groups, so as to be able to appraise the effects of each of the independent variables.

The field experiment differs from laboratory procedure, however, in that it is carried out in a situation which exists without instigation by the experimenter. Subjects are studied in a natural context where they would have been even if no study had been conducted. Care is taken, moreover, to avoid any awareness on the subjects' part that they are participants in an experiment. To the extent that these conditions can be maintained, it is thought that this procedure avoids a major source of distortion, and therefore permits results to be generalized with greater confidence to the "real world."

The key to this experiment lay in obtaining data on the actual tax-paying behavior of those studied. With the cooperation of the Internal Revenue Service, we obtained figures for groups of taxpayers. These groups comprise the experimental treatment groups and the controls. By giving us distributions for entire groups, the IRS complied with the statutory provision that no individual returns be disclosed.

Assignment to experimental groups followed conventional techniques. Subjects were matched individually for precision control using residential criteria from census tracts. The sample was drawn from areas where income was generally above 10,000 dollars to minimize the number of subjects whose compliance was assured by full withholding and short form filing. While this criterion prevented coverage of lower income families, it increased the chances of obtaining differential effects from the varying treatments. Subjects were then assigned to experimental or control groups on a random basis. Treatment groups consisted of 92 taxpayers in the placebo, 92 in the conscience, and 89 in the sanction group. They were interviewed during the month prior to filing.

41 A strong case for the use of field experiments in research on legal policy questions is stated by ZEISEL, KALVEN & BUCHHOLZ, DELAY IN THE COURT (1959).
42 We acknowledge with thanks the cooperation of the Internal Revenue Service, particularly Bertrand M. Harding, William H. Smith, J. R. Turner, and Howie Wilson.
43 INT. REV. CODE of 1954, § 7218.
their returns on their attitudes toward political and civic issues, with particular emphasis placed on tax policy.

In each of the experimental groups, questions were included which were intended to accentuate certain motives for tax payment. The "sanction-treated" group was asked questions such as the following: "A jail sentence of three years could be imposed for willful failure to pay tax on interest." Under what conditions do you think the Government should impose a jail sentence? A series of similar questions

Questions specified the payment of taxes on interest for two reasons. The Revenue Service had mounted a campaign during the year of the interviews to publicize the requirement of reporting interest, and this background provided a rationale for the interviewers' interest in interest. Second, we concentrated on a particular type of compliance for theoretical reasons. We predicted a relatively higher increase in compliance in the reporting of interest than in other types of compliance, especially in the group threatened with sanction. The latter effect has not occurred in findings to date, suggesting that taxpayers did not distinguish between income by source.

The questions aimed at inducing sanction threat are given below. Italicized portions comprise instructions to the interviewer. The interviews were carried out by a professional research firm, the Research Guild, using ten experienced interviewers who were randomly assigned. We are grateful to Hugh Edwards of the Research Guild for the adroit execution of an unusually complex field operation.

Sanction Questions

One important area of government which is of concern to many people is the Federal Income Tax. We would like to know how people like you feel about the way the Government goes about enforcing the income tax.

1. First of all, we are interested in experiences people may have had with the Internal Revenue Service. Has your return ever come up for investigation? If No, Go On To 2. If So: 1a. Briefly, do you recall what happened in the investigation? 1b. How did you feel about the experience? (PROBE. For example, did you feel that you were treated fairly, was the experience anxiety-producing?)

2. In the last session of Congress a bill was passed which requires banks and other savings institutions to report the names of people who have interest payments of $10.00 or more. Do you feel that it is right for the government to have this kind of information about people?

2a. Why do you feel this way?

3. The Internal Revenue Service is installing high speed computers like Univac. Do you think that these electronic computers should be used to check up on whether people report all their interest?

4. These computers together with the information from the banks would make it possible for the Internal Revenue Service to check up on as much as 70% or more of the people who have interest, with little effort and expense. Do you think the government should investigate the reported income from interest of so many people?

4a. Why do you feel that way?

5. Would you be in favor of the Internal Revenue Service making public the names of people who wilfully did not report the full amount of their interest?

5a. Why do you feel that way?

6. Failure to report all interest would involve "subscribing to a false tax return." The fine for such an offense is $5,000. Do you think this fine is set so high to deter people from evading their tax on interest, or to make up for the money lost due to tax evasion, or for some other reason?

7. A jail sentence of three years could be imposed for wilful failure to pay tax on interest. Under what conditions do you think the Government should impose a jail sentence?

8. Would you approve of a plan to investigate the previous returns of individuals who did not report their interest on this year's return?

8a. Why is that?
emphasized the severity of sanctions available to the government and the likelihood that tax violators would be apprehended. The "conscience" group, on the other hand, was exposed to questions accentuating moral reasons for compliance with tax law. These included questions such as: "Would you consider a citizen's willful failure to pay tax on interest an indication that he is unwilling to do something for the country as a whole?"

Similar questions posed to this group emphasized the severity of sanctions available to the government and the likelihood that tax violators would be apprehended. The "conscience" group, on the other hand, was exposed to questions accentuating moral reasons for compliance with tax law. These included questions such as: "Would you consider a citizen's willful failure to pay tax on interest an indication that he is unwilling to do something for the country as a whole?"
sized noncontroversial uses for which tax money is employed, citizen obligation to government,\footnote{47} and the value of personal integrity.

In addition to the two treatment groups, a placebo group was selected by the same technique of precision control and random assignment. This group was given the same basic interview without any accentuation questions.\footnote{48} A fourth group served as an untreated control.

6. Basic research in nuclear physics
7. Development of modern weapons systems
8. Civil defense
9. The space program
10. The Peace Corps
11. The foreign aid program
12. Building roads and highways
13. Area redevelopment
14. Flood control projects
15. Urban renewal programs
16. National recreation areas
17. Other

5. Some people think that it is "fair game" to swindle the United States Government, by not reporting all their income from interest, for example. What is your opinion of such an attitude?
   5a. Do you think this kind of attitude (that is, fair game to swindle the government) indicates weakened respect for law?
   5b. Why do you feel that way?
   5c. Do you feel that such an attitude indicates disrespect for the government?
   5d. Why do you feel that way?

6. It has been suggested that there is not much difference between draft dodging and tax evasion.
   6a. Do you think that failure to pay tax on interest and draft dodging both indicate that evaders and dodgers put their own self interest above the interest of others?
   6b. Do you think that failure to pay tax on interest and draft dodging are similar in that both indicate a failure to take seriously the responsibilities of citizenship?
   6c. Why do you feel that way?
   6d. When you consider all relevant factors, including the number of people involved, do you think that the people who fail to pay all the tax they owe on interest hurt the country about as much as, more or less than the men who illegally evade the draft?
   6e. Why do you feel that way?

7. How guilty do you think a person should feel who knowingly does not pay his tax on the interest which he earned?

\footnote{47} An appeal to conscience could, of course, heighten the determination to resist taxes. See \textsc{Wilson}, \textit{The Cold War and the Income Tax: A Protest} (1963). To judge from the results reported subsequently, the sample included a minority of citizens—not unlike Thoreau, Edmund Wilson, or J. Bracken Lee—for whom reflection on the intended uses of taxes might decrease the tendency to comply.

\footnote{48} The interviewed or placebo control group was provided in order to rule out a number of "threats to validity." This group was initially as similar to the two experimental groups as random assignment could make it. It was subjected to an interview which was the same as that received by the experimental groups, except of course, for those questions which constituted the two experimental treatments. The history and motivation of all three groups after the interview and prior to their payment of taxes was presumably the same, except insofar as the treatments made the experimental groups behave differently. The effects on taxpaying behavior were determined by the same measurements, and the impact of these observations was minimized by virtue of the subjects' lack of awareness that changes in taxpaying conduct were under scrutiny. Regression effects were constant, since
Its purpose was to determine whether the basic interview, with or without accentuation, affects taxpaying behavior.49

Two kinds of data may be reported at present: interview responses and actual tax returns. The first is derived from an internal analysis of the interviews themselves. Toward the close of the interview, after the experimental treatment was given, all three groups were asked an open-ended question: "What reasons do you think taxpayers might have for reporting all the interest they earn on their tax returns?" Answers to this question were content-analyzed to determine whether normative reasons were the first given. The results of this analysis indicate that several social variables affect the manner in which taxpayers are oriented toward paying taxes and/or the manner in which they respond to different motivational appeals.50

all three groups were drawn at random from the same initial segment of the population. Dropout rate following the initial interview was constant and below 5% for all three groups. All of these considerations lead to the conclusion that the differences between the groups were internally valid; i.e., that the experimental treatments did in fact account for the observed differences among the treatment groups and the interviewed, or placebo, control. For a detailed authoritative discussion of these threats to validity, see Campbell and Stanley, Experimental and Quasi-Experimental Designs for Research, in HANDBOOK OF RESEARCH IN TEACHING 171 (Gage ed. 1963) (also available as a 1966 monograph).

49 The untreated control was included to test the effects of the interview through a comparison with the placebo, or interviewed, control group. Its value for this purpose was diminished substantially by a tactical decision which is subject to criticism. Members of this group were not contacted to determine whether they were available for interview. Accordingly, the untreated control includes individuals similar to those who were dropped from the treated group. If those who would have been unavailable for interview differed from those interviewed, the untreated control is not comparable to the other three groups.

On the other hand, even preliminary contacting might have diminished the value of this group in determining the pretest effect. One solution, still possible if the relevant data can be obtained, is to compare the 120 in each of the four groups selected. While this would diminish the magnitude of effect, by adding untreated dropouts to the treated and placebo groups, it would avoid the problem of bias resulting from dropout.

As matters now stand, however, the design fulfills Campbell and Stanley's criteria as a true experimental design. It falls into their category No. 4, as a "pretest-posttest control group design." It short of their category No. 5, "The Solomon Four-Group Design," not only because of the flaw mentioned above, but also because there is no treatment group which was not initially observed. Given the interview technique for introducing experimental variables, this feature of the Solomon design was impossible.

The actual design, strong on "internal validity" as described in footnote 48 supra, loses somewhat in external validity; i.e., capacity for generalization to other populations, settings, treatment, and measurement variables. Its external validity is also diminished by the omission of the lowest socio-economic bracket from the groups studied, a measure explained in the text. Increased external validity was obtained, on the other hand, by the use of presumably nonreactive techniques in introducing the experimental variables and the use of an unobtrusive measure of effect. For further discussion of reactivity in experimental designs, see WEBB, UNOBSERVABLE MEASURES (1966); Campbell and Stanley, supra note 48, at 171; Campbell, Factors Relevant to the Validity of Experiments in Social Settings 54 PSYCHOL. BULL. 297 (1957).

50 In this analysis, only first order relationships are presented. That is, the groups are
The first of these variables is social class or socio-economic status. Using the Hollingshead two-factor index of occupation and education, we found no relationship between initial normative orientation, as inferred from responses in the placebo interview group, and social class (Table I, column 1). The success of conscience appeals did not vary systematically by the dimension of social class, being greater in classes II and IV than in classes I and III (Table I, column 2). There was, however, a direct relation between social class and the capacity of sanction threats to elicit normative content. Column 3, Table I, shows a remarkable regularity in the relationship between normative responses after sanction threat and socio-economic status. Comparing

not broken down in a manner which permits the examination of the effect of the treatments on normative attitudes by two independent variables simultaneously. This multivariate method will be employed at a later stage of the research. Preliminary inspection, however, gives no special reason to suspect that first order effects will disappear under second order analysis. For one thing, there is a relatively even distribution of variables by other categories. Where asymmetry exists, its control appears likely to accentuate first order relations. For instance, Catholics do not respond to sanction normatively, yet are over-represented in the class I, sanction-treated group, which does. (Cf. Tables I and IV and associated analyses.)

The Hollingshead Index of Social Position identifies five classes, of which four are represented in our sample. These four classes range from the top executives and professionals of class I through the upper middle class II, middle and lower middle class III, to working or upper lower class IV. Only the unskilled lower class are unrepresented in this sample. For a detailed description of the characteristics of these classes as described for New Haven (where the scale was validated), see Hollingshead & Redlich, Social Class and Mental Illness 66-186 (1958). The techniques for generating and validating the scale are described id. at 387-407.

### TABLE I

| Socio-Economic Status and Normative Orientation toward Taxpaying (Percentage Giving Primary Normative Response) |
|---|---|---|---|
| Placebo | Conscience | Sanction |
| **Class** | **%** | **N** | **%** | **N** | **%** | **N** |
| I (highest) | 56 | 15/23 | 71 | 17/24 | 82 | 22/27 |
| II | 60 | 15/25 | 87 | 13/15 | 67 | 16/24 |
| III | 56 | 15/27 | 71 | 25/35 | 59 | 16/27 |
| IV (lowest)** | 53 | 9/17 | 88 | 14/16 | 50 | 5/10 |
| Total | 57 | 52/92 | 75 | 69/92 | 67 | 59/88 |

* The N column shows the number giving the primary normative response over total respondents in the category. Numbers in the three groups vary slightly in each table because those who did not answer are omitted.

** Hollingshead's lowest socio-economic category was unrepresented in the sample because of a restriction to persons filing long form returns.

---

51 The Hollingshead Index of Social Position identifies five classes, of which four are represented in our sample. These four classes range from the top executives and professionals of class I through the upper middle class II, middle and lower middle class III, to working or upper lower class IV. Only the unskilled lower class are unrepresented in this sample. For a detailed description of the characteristics of these classes as described for New Haven (where the scale was validated), see Hollingshead & Redlich, Social Class and Mental Illness 66-186 (1958). The techniques for generating and validating the scale are described id. at 387-407.
columns 1 and 3 in Table I, we find an increase of 26% in normative orientation following the sanction threat for class I as against 7%, 3% and -3% in classes II, III, and IV respectively.\(^{62}\)

This is the first instance of a phenomenon which occurs repeatedly in the data. For class I, it appears that, while direct appeals to duty do not create a greatly increased sense of obligation, threat of sanction does. This may well indicate a major mechanism of social control. Sanctions may be most effective in preventing violations where they are converted into a sense of moral obligation. Whether this "induced morality" actually increases compliance, however, will only become apparent when it is related to behavior such as the actual taxpaying of the subjects.

Another point of interest emerges from the normative reactions of class IV as seen in Table I, row 4. Respondents in class IV show the highest gain in normative content following the conscience appeal (35%). By contrast, however, they show a slight decline following threat. In effect, they will go along with reminders of the obligation to pay taxes but will not be pushed into that position by warnings of potential punishment. The response to conscience appeals in the working class group suggests a combination of patriotism and an appreciation of the benefits provided by a welfare state. Indifference to sanctions in this stratum may relate to the reality of tax enforcement policy; i.e., the probability of tax investigation is, in fact, an increasing function of income. In addition, the working class may be somewhat more prone to view the world as dominated by luck and fate, and thus not expect a close relationship between conduct and consequence.

If such differences affect actual compliance, the practical implications for tax policy may be extremely important. Accentuation of threat, through publicized prosecutions and the like, would be called for against the rich but not against the working class. For the latter, educational campaigns emphasizing patriotic considerations appear to be more promising.

These differences by social class are consistent with reactions by educational background. Initial differences in normative content, as shown in the placebo group (Table II, column I) are not significant,\(^ {62}\)

\(^{62}\) We have used arithmetic differences in percents to compare the placebo and experimental groups. An alternative method is to consider the experimental group scores as a percent of the placebo. This method controls for difference in base scores as indicated by the placebo and accentuates the magnitude of difference. Arithmetic difference was adopted to keep the discussion consistent with the tables. The basic results are not affected by the different modes of calculation. For a discussion of alternative treatments of this problem, see Webb, Unobtrusive Measures 6-8 (1966), and the sources cited therein.
though slightly higher for college than high school educated. Under the treatments, however, the least educated show a marked response to conscience appeals (36% increase), but no increase at all under sanction threat. College graduates, by contrast, show some evidence of threat-induced conscience (10%), even though this increase is less than the 18% produced by conscience appeals.

Taking the two variables of class and education together, what possible explanations of these findings emerge? Less educated, working class people may be more prone to respond to conscience appeals because of greater piety or naïveté, or because of a conviction that government action is needed for the solution of social problems. Alternatively, the better educated, upper class respondents may already have been exposed to such reasons for taxpaying and either accepted them—as the slightly higher proportion of normative responses for upper against lower categories in the placebo in Tables I and II suggests—or be resistant to attitude change when it is implicitly urged on them. In regard to sanction, less educated, working class individuals may be inclined to discount the prospect of prosecution being directed against them, be less worried about the experience if it should occur, and be less likely to convert the fear of sanction into normative reasons for acquiescence.

Another factor which relates to the impact of given motivation is the individual's employment status. As seen in Table III, the self-employed were slightly more normative in the placebo condition than those working for others. Subjected to conscience appeals, however, the self-employed showed a 38% increase in normative content, whereas the self-employed were only moderately moved (12%). The effects of the conscience appeal on those working for others overshadows the limited effects of sanction threat on either group. The explanation of this finding is not immediately apparent. Perhaps those who are employed in an organization become accustomed to responding to exhortations that they comply for the good of the organization, and they ex-
On Legal Sanctions

TABLE III

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Conscience</th>
<th>Sanction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Self-employed</td>
<td>59</td>
<td>34/58</td>
<td>71</td>
</tr>
<tr>
<td>Employed by others</td>
<td>52</td>
<td>16/31</td>
<td>90</td>
</tr>
</tbody>
</table>

...tend this to the government when urged. Whatever the explanation, the phenomenon is intrinsically interesting. It suggests that the increasing pervasiveness of corporations as a source of employment helps to explain tax compliance, if not the fulfillment of the general obligation of citizenship, in complex societies. Along the same lines, it may help to explain the resistance to taxation among small businessmen and farmers, as in the Poujadist movement.

Religion also emerges as a correlate of taxpaying orientation. In the placebo group, Catholics showed the most frequent normative orientation, followed by Protestants and Jews (Table IV, column 1). Sub-

TABLE IV

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Conscience</th>
<th>Sanction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Catholics</td>
<td>67</td>
<td>16/24</td>
<td>64</td>
</tr>
<tr>
<td>Protestants</td>
<td>52</td>
<td>22/42</td>
<td>83</td>
</tr>
<tr>
<td>Jews</td>
<td>48</td>
<td>10/21</td>
<td>76</td>
</tr>
</tbody>
</table>

jected to the experimental interviews, however, Catholics showed little normative reaction to sanction threat (2%) and actually declined slightly (3%) in these responses upon being confronted with considerations of conscience. Perhaps their normative orientation is so well set by initial training or so exclusively related to an authoritative church that they do not respond to appeals or threats from other sources. By contrast, Protestants and Jews increased their normative statements to a point exceeding the Catholics under conscience appeals. Protestants showed an increase of 31% in normative content under conscience appeals, and Jews were almost equally responsive to the conscience appeal (28%). Jews showed a slightly greater normative reaction (19%) than Protestants (15%) to sanction threat, the effects of which brought Jews and Protestants virtually up to the position of the Catholics.

These diverse reactions to motivational appeals must be viewed with caution. They are based on relatively small groups which were selected...
from relatively well-to-do segments of the population. The correlations reported, moreover, have not, because of small numbers, been subjected to multivariate analysis to check the independence of effects. Nevertheless, the marginal analysis which has been done suggests some definite variations in initial orientation to taxpaying and in responsiveness to various motivational appeals. At a minimum, the findings suggest the following relationships:

1. Cultural and social structural factors correlate with legally relevant attitudes.
2. Conscience appeals often have a greater effect on groups with low initial normative orientations than on those who are initially high in normative orientation.
3. Sanction threat has a mixed effect on normative orientations, being capable of "inducing morality" among several categories.

It remains to be seen how these reactions relate to tax compliance. At this point, we are unable to report changes in payment in relation to social characteristics or attitudinal responses. We have received preliminary data, however, for changes in tax payments for the gross experimental and control groups. These data permit comparisons of the two experimental groups with the control group in regard to changes in tax payment on returns filed before the interview (for fiscal 1961) and after the interview (for fiscal 1962). Since the groups were randomly assigned (after matching by residence), differences in increase of taxpaying are presumably attributable to the experimental treatments.

A caveat must be entered at this point. The results obtained from the experiment are not of a magnitude which uniformly produces statistically significant differences. Some of the results reported, especially if taken separately, could well be attributed to chance. This could be the consequence of small samples, weakness of the experim...
mental treatments, a limited amount of cheating relative to the large base of compliance, or the fluctuation of income found in all groups. Nevertheless, the results do fall into a pattern which is highly suggestive. As indicated in Table V, column Ia, those threatened with sanction declared a mean increase in adjusted gross income of 181 dollars as compared with a mean decrease of 87 dollars for the placebo control (and a mean decrease of 13 dollars for the untreated control). It should also be noted, as seen in Table V, column IIIa, that income tax after credit shows a mean increase for the sanction-threatened group, compared with a decrease for the two controls.

If we look at the number of individuals in the sanction group who increased their adjusted gross income, this pattern is repeated. In the sanction-treated group, as shown in Table VI, 49 of 87 taxpayers, or 56%, exceeded the median dividing line set by the placebo group. The number of sanction-threatened individuals also slightly exceeded the placebo median on income tax paid.

It is equally important to note that the conscience appeal had a stronger effect on income reported than did the threat of sanction. Adjusted gross income reported by the conscience-appeal group showed a mean increase of 804 dollars compared with 181 dollars for those threatened with sanction (Table V, column Ia). The mean increase in AGI in the conscience group is statistically significant when compared with the placebo. Income tax payments for this group are also appreciably higher than for the sanction group or the controls. The number of conscience group members who exceeded the placebo median on adjusted gross income reported is 56 of 88 or 64% (Table VI, column I.) From this it may be inferred that more than one-fourth of the conscience group who would—in the absence of the appeal—have been below the median, were moved above it on AGI reporting: (The number of individuals who increased their taxes—Table VI, column III—also exceeded the median, although this effect is slight and statistically not significant.) These results tend to confirm the proposition that conscience appeals can be more effective than sanction threats.

Various interpretations of these results are possible. Despite efforts to avoid threats, some implication of punishment may have been conveyed in the conscience interview. If so, the greater reporting of income by this group might have resulted from a subjective summation of threat

more so for the conscience group. On the other hand, fewer than half of the taxpayers in the experimental groups exceeded the median deduction of the placebo, particularly in the sanction group. This suggests that the increase in deductions noted in the $t$ test analysis resulted from a small number of person deducting a large amount.

The advice and assistance of Donald G. Morrison on statistical aspects of the analysis is gratefully acknowledged.
<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>I. Adjusted gross income</th>
<th>II. Total deductions</th>
<th>III. Income tax after credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Sanction threat</td>
<td>87</td>
<td>$181</td>
<td>$3,481</td>
<td>$273</td>
</tr>
<tr>
<td>B Conscience appeal</td>
<td>88</td>
<td>804</td>
<td>4,007</td>
<td>177</td>
</tr>
<tr>
<td>C Placebo control</td>
<td>88</td>
<td>-87</td>
<td>2,858</td>
<td>132</td>
</tr>
<tr>
<td>D Untreated control</td>
<td>111*</td>
<td>-13</td>
<td>3,510</td>
<td>320</td>
</tr>
</tbody>
</table>

Comparison of increase, 1961-62, t Test**

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>t</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanction (A) v. Conscience (B)</td>
<td></td>
<td></td>
<td>1.35(.10) B &gt; A</td>
</tr>
<tr>
<td>Sanction (A) v. Placebo (C)</td>
<td>1.09(N.S.)</td>
<td>.55(N.S.)</td>
<td>.39(N.S.)</td>
</tr>
<tr>
<td>Conscience (B) v. Placebo (C)</td>
<td>.55(N.S.)</td>
<td>1.30(.10) A &gt; C</td>
<td>1.87(.05) B &gt; C</td>
</tr>
<tr>
<td>Placebo (C) v. Untreated control (D)</td>
<td>1.68(.05) B &gt; C</td>
<td>1.81(.05) D &gt; C</td>
<td>.14(N.S.)</td>
</tr>
</tbody>
</table>

* Adjusted gross income tabulations are based on 110 taxpayer returns. Data for one taxpayer showing a very large decrease in adjusted gross income were excluded by the Revenue Service from the tabulations to avoid possible identification.

TABLE VI  
NUMBERS OF TAXPAYERS IN EACH EXPERIMENTAL GROUP  
EXCEEDING MEDIAN INCREASE, 1961-62, OF CONTROL GROUPS IN VARIOUS ELEMENTS  
OF TAX REPORTING; SIGN TEST FOR SIGNIFICANCE*  

<table>
<thead>
<tr>
<th>Comparison</th>
<th>I. Adjusted Gross Income</th>
<th>II. Total Deductions</th>
<th>III. Income Tax after Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N &gt; Median of</td>
<td>N &gt; Median of</td>
<td>N &gt; Median of</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>C(144)</td>
<td>%</td>
</tr>
<tr>
<td>Sanction threat(A) v. Placebo(C)</td>
<td>87</td>
<td>49</td>
<td>56</td>
</tr>
<tr>
<td>Conscience appeal(B) v. Placebo(C)</td>
<td>88</td>
<td>56</td>
<td>64</td>
</tr>
<tr>
<td>Combined experimental (A + B) v. Placebo(C)**</td>
<td>175</td>
<td>105</td>
<td>60</td>
</tr>
</tbody>
</table>

* Brownlee, Statistical Theory and Methodology in Science and Engineering 180-84 (1960). This test uses the placebo to establish a median and then computes the probability of at least the number which exceeds that median doing so by chance.

** A sign test of the placebo against the medians of the untreated control shows that the two groups are very similar on AGI (P = .54, median of D = 114), but very different for deductions (P = .000002, median of D = 238) and somewhat so for tax (P = .084, median of D = -68). These observations reinforce the conclusion, indicated earlier, that the untreated control is not dependable in the absence of further data. In substance, this means that the bias resulting from dropouts or the interview made a difference, particularly in deductions. The similarity of AGI's of the placebo and the untreated control strongly suggests that the smaller deductions in the two experimental and the placebo groups resulted from the interview rather than from dropout bias.
and conscience motivations. Moreover, threats conveyed under the guise of an appeal to conscience may have led to much greater compliance because of the mechanism of threat-induced moralism noted in the interview results. On the other hand, it is possible that the subjects of these interviews were simply more responsive to appeals to conscience than to the kind of threats which apparently made at least some sanction group members dig in their heels.

There is additional evidence which suggests that some who are induced to comply by threat find ways of expressing their resentment. Deductions taken by the sanction-threatened group showed a mean increase as compared with those who were given the conscience appeal. As shown in column IIa of Table V, those threatened with sanction had a mean increase in deductions of $273, compared with $177 for the normative group. It is tempting to interpret this difference as meaning that the threatened group said: "You may beat me into admitting higher income, but I'll find a way of getting it back." One might also infer that the conscience-appeal group kept its deductions low in recognition of the importance of tax payments for the welfare of the country.54

These interpretations lose force, however, from the position of the two control groups on deductions. For one thing, the placebo control shows an even lower mean increase in deductions than the conscience group, a result which is not attributable to induced motivations. A second confounding fact is that the untreated control shows even higher deductions than the sanction treated group.55 A third point to be noted is that a minority of members of both experimental groups increased their deductions over the median provided by the placebo (Table VI). This shows that the increase in aggregate deductions by the experimental groups is attributable to a minority of the members of each group, the minority being even smaller in the sanction than in the conscience group. If there is a heel dragging effect, then, it is found in a relatively small group. It remains to be seen whether this group can be identified as having distinctive social characteristics and attitudes.

These three findings make it prudent to interpret the suggestive

---

54 Enrick reports some sentiment along these lines. As one subject put it: "Taxes are sure high. But of course we must be thankful for what we have, and if taxes are needed to keep the country what it is, we'll just have to pay them." Enrick, A Pilot Study of Income Tax Consciousness, 16 NAT'L TAX J. 169, 172 (1963).

55 Deductions are larger in the untreated control than in the placebo by a statistically significant margin (Table V, column II). This reinforces the conclusion that the interview led to a general decrease in deductions.
differences between the two experimental groups with considerable caution. The results, nevertheless, give some evidence to support the following propositions:

1. Compliance can be increased by threat of punishment.
2. Appeals to conscience can be a more effective instrument than sanction threat for securing compliance.

Further analysis is needed to determine how these experimentally induced motives affected subgroups whose verbal reactions to the interviews varied so widely. Those results are now being prepared by the Internal Revenue Service. When completed, it is hoped that they will contribute to a fuller understanding of the motivations affecting legal compliance.

CONCLUSION

This paper has reviewed some widely held assumptions concerning the operation of legal sanctions. These were then subjected to empirical examination. The study reported was a field experiment aimed at determining the effectiveness of sanction, as compared with an appeal to conscience and with a placebo control, in increasing normative sentiments about compliance and in heightening actual compliance in the payment of federal income taxes. The findings indicate that motivations of various kinds make a difference in taxpaying. They suggest that the two types of appeal affect normative orientation differently according to the status of those subjected to the appeals. Sanction threat increases normative orientation most markedly among the upper class, the better educated, and non-Catholics. Appeals to conscience change attitudes toward tax compliance most among the best and least well-educated, those employed by others, and Protestants and Jews. As to actual changes in tax compliance, returns currently available for the gross treatment groups suggest that conscience appeals are more effective than sanction threats, though both have some effect. The conscience appeal appears, moreover, to produce less loss through resistance, an effect implied by the increase in aggregate deductions noted in the sanction-treated group.

These results must, of course, be viewed with caution. They were obtained from a population in a given geographical area, by a novel (i.e., otherwise untested) method, in an examination of a largely well-to-do population. The behavior examined, moreover, may be atypically sensitive to conscience considerations, particularly in the American urban setting. Before these results can be confidently used as a girder
in the construction of a theory of legal compliance, they must be joined with studies of the same problem which use different methods, populations, and types of legal compliance. To build an adequate theory of legal compliance, much more work is needed both at the drawing board and in the field.

Nevertheless, the results of the study carry several implications for a theory of sanction. They suggest that the threat of sanction can deter people from violating the law, perhaps in important part by inducing a moralistic attitude toward compliance. This mechanism seems particularly significant when those subject to sanction threat are not trained by, and associated with, an authoritative institution other than the state. The threat of punishment appears, however, to produce some resistance to compliance. Such resistance can be minimized through alternative techniques of securing compliance, such as the utilization of appeals to conscience and to a sense of civic responsibility, motives which can be more powerful than sanction threat in increasing compliance with the law.