

# APPEALS TO CIVIC VIRTUE VERSUS ATTENTION TO SELF-INTEREST: EFFECTS ON TAX COMPLIANCE

KATHLEEN M. MCGRAW  
JOHN T. SCHOLZ

In an influential article, Schwartz and Orleans (1967) reported that emphasizing normative obligations to obey the law had a greater positive impact on tax compliance than did emphasizing deterrence factors. In this article, we report the results of a conceptual replication of that earlier experiment, within the context of a dynamic model of tax schema change.

Prior to filing the 1987 tax year return (the first affected by the 1986 Tax Reform Act, TRA), taxpayers viewed one of two videotapes, one emphasizing the social consequences of the TRA and normative duties, the other emphasizing the personal consequences of the TRA and strategies to reduce tax liability. A control group of taxpayers did not receive information about the TRA. Data on attitudes, beliefs, and self-reported compliance were collected at three points during the tax season. Tax return data from the IRS were also obtained. The results indicate that the normative and personal consequences messages had predictable and persistent effects on tax-specific attitudes. However, these effects did not translate into changes in basic values or compliance, as reflected by either the self-report or the official tax return data. Possible explanations for the discrepancy between the results reported by Schwartz and Orleans (1967) and this experiment are considered.

Laws typically create a conflict between the civic duty to comply and the self-interested advantage gained by noncompliance. These different motivations are reflected in two competing images that appear in the compliance literature. The *normative* image views compliance behavior as determined by the response to the

---

A previous version of this paper was presented at the annual meeting of the Law and Society Association, Berkeley/Oakland, CA, 1990.

This research was funded by NSF Grant SES8710891 and by the Internal Revenue Service. Preparation of the manuscript was facilitated by the Dr. Nuala McGann Drescher Award to Kathleen McGraw. We are grateful to the many people who have provided us with assistance and advice throughout the course of this project. In particular, we thank Bill Lefbom and Pat Bennett at the Internal Revenue Service; John Boyle and his staff at Schulman, Ronca, and Bucuvalas, Inc., for implementing the survey component of the research; John Carroll and Bob Cialdini for their advice at various phases; Cheryl Di Carlo, Mary Ellen Kent, Ellen Mills, and Patricia Reisman for research assistance; and Neil Pinney and Marco Steenbergen for their assistance at all phases of the project.

LAW & SOCIETY REVIEW, Volume 25, Number 3 (1991)

question "What should I do?" or "What am I obligated to do?" whereas the *personal consequences* image assumes that the relevant question is "What will make me better off?" (Carroll 1987; McGraw and Scholz 1988). The normative image emphasizes moral reasoning about the appropriateness of norms and principles related to a perceived legal obligation. In contrast, the personal consequences image emphasizes cost-benefit calculations based on the expected utility of different courses of action.

Conflict has always been present between adherents of the two models, with some emphasizing the weakness of the normative perspective (e.g., Hyde 1983), and others pointing out the deficiencies of the currently more influential expected utility perspective (e.g., Etzioni 1988; Mansbridge 1990; Tyler 1990). We believe that both images are important for understanding citizen responses to legal obligations. In this article, we focus on the dynamic impact of exposure to information relating to norms and personal consequences on taxpaying behavior and related beliefs and attitudes. Income tax provides a useful arena for the study of compliance because of the clear contrast between self-interest and normative obligations. From a personal consequences perspective, tax compliance is viewed as an income-maximizing decision balancing the net gain of underreporting income or overclaiming deductions against the added risk of detection and penalization. The normative perspective on tax compliance, on the other hand, emphasizes moral reasoning and the processes of socialization and internalization of norms.

In 1967, Schwartz and Orleans published a classic study in tax compliance, an experiment that pitted the normative and personal consequences images against each other. Provided with Internal Revenue Service (IRS) tax return data for the set of taxpayers they studied, Schwartz and Orleans found that drawing attention to normative obligations to obey the law had a greater positive impact on tax compliance than did providing deterrence-oriented reminders about the risk of punishment. The Schwartz and Orleans experiment is generally viewed as a landmark, cited approvingly by many compliance and sociolegal scholars (e.g., Boruch 1989; Cialdini 1989; Roth, Scholz and Witte 1989; Tyler, Rasinksi, and Griffin 1986), in large part because it has been the only study in which a randomized experiment was used to test the implications of the two competing theories on actual taxpaying behavior.

Lempert (1989:155) has clearly articulated the need to replicate such important policy-relevant studies because "most studies are in some ways narrowly confined and leave open a number of credible threats to external validity" (p. 155). A National Science Council review of the tax compliance literature (Roth, Scholz, and Witte 1989) also called for a replication of the Schwartz and Orleans experiment because of controversies in interpreting and accepting their results (Friedman and Macaulay 1977) and recent ad-

vances in applications of cognitive and social psychology to compliance issues (Carroll 1987, 1989; Cialdini 1989; Smith and Kinsey 1987).

The experiment described in this article is a conceptual replication of the Schwartz and Orleans experiment. Taxpayers viewed one of two different videotaped messages in early 1988, just when they were receiving the new tax forms reflecting the first major changes introduced by the 1986 Tax Reform Act (TRA). One videotape emphasized the normative duties of citizenship and the social consequences of the TRA. The other emphasized the personal consequences of the TRA and strategies to reduce tax liability. Panel data on attitudes, beliefs, and self-reported compliance behavior were gathered from three interviews, the first immediately following exposure to the videotapes, the second about two weeks later, and the third six months later, after the filing of the 1987 tax year return. Moreover, because of unique cooperation from the IRS, we were able to examine the impact of the experimental messages not only on attitudes and self-reported behavior but also on officially documented taxpaying behavior.

### I. A DYNAMIC MODEL OF TAX SCHEMAS

The study rests on the assumption that compliance can be understood as a decision process in which a broad range of beliefs, attitudes, and values may play a role in determining behavior. Following Roth, Scholz, and Witte (1989), we refer to the sum of the relationships that exist among tax-relevant knowledge, beliefs, attitudes, and values as the taxpayer's *tax schema* (see also Sears and Citrin 1982). Despite considerable research effort, a clear mapping of the relationships among the concepts in tax schemas has been elusive (Lewis 1982; Roth, Scholz, and Witte 1989; Sears and Citrin 1982). Some relationships have proven to be somewhat robust across surveys: both normative concepts (e.g., the moral obligation to comply) and to a lesser extent self-interested deterrence concepts (e.g., fear of getting caught and punished for noncompliance) typically are related to self-reported compliance. These normative and deterrence factors are referred to as inhibitors, as they have been clearly linked to "inhibiting" noncompliance (Grasmick and Bursik 1990). Other relationships between tax policy preferences, knowledge, and evaluations of new laws on the one hand and compliance on the other are much less robust.

A particularly important concept in our model is fairness, because of the centrality of fairness in the rhetoric of the TRA political debates. The link between fairness and compliance is not as clear as the link between the inhibitors and compliance. In studies of tax compliance, the relationship appears to be indirect, mediated by feelings of moral obligation (Kinsey and Smith 1987; Scott and Grasmick 1981). Thus, judgments about the fairness of tax

laws appear to be more peripherally related to compliance behavior than the direct inhibitors. Fairness judgments may be less robust in their relationship to compliance because multiple standards exist in deciding "what's fair" (Hochschild 1981; Lane 1986). Although multiple standards exist, perhaps the most consistently replicated finding is that there is a self-serving bias in judgments of fairness (Reis 1984). Thus, those most hurt by a new law like the TRA are most apt to evaluate it as unfair. Ethical disagreements exist as to the appropriateness of this self-serving bias. From a cognitive perspective, however, it is reasonable. The personal consequences of a law will be relatively salient and accessible, providing an anchor for fairness judgments, whereas it is extremely difficult to assess changes affecting multiple and diverse social groups.

Although we presume that personal impact is a critical determinant of fairness judgments and view it as cognitively reasonable, numerous scholars, including economists (Harsanyi 1977), psychologists (Piaget 1965; Kohlberg 1984) and philosophers (Rawls 1971), have argued that the use of impersonal standards is the hallmark of true moral reasoning. The rhetoric of the 1986 TRA debates appealed to such impersonal standards with an emphasis on eliminating loopholes for wealthy individuals and corporations, eliminating taxes for the poor, and equalizing taxes for individuals at the same level of income. The normative experimental videotape provides a great deal of information about these shifts in the tax burden and explicitly links these shifts to enhanced fairness. We will test the hypothesis that providing information about broader social consequences increases the reliance on an impersonal standard, with a parallel attenuation of the reliance on personal impact, in fairness judgments.

### **The Influence of Norms and Personal Consequences on Compliance**

Models of moral reasoning (Rest 1984; Schwartz 1977) and formal models of compliance choices (Margolis 1982; 1991) have provided similar bases for understanding how normative concerns and personal consequences interact to influence compliance behavior. Analyses of moral reasoning have suggested a four-stage choice process: (1) interpret the situation as one that involves a moral principle; (2) formulate a plan of action applying the moral principle; (3) consider competing values or goals (Rest 1984) and costs (Schwartz 1977), and decide whether to act on the moral principle or in a manner consistent with the competing value or cost; (4) implement the chosen plan of action. At the critical third stage, a taxpayer who has recognized some obligation to report truthfully a particular item on the tax return may still consider competing moral principles as well as personal consequences before deciding what to do. If the system is perceived as unfair, or if the conse-

quences are considered too burdensome, the moral obligation to report truthfully may be “neutralized” (Thurman, St. John, and Riggs 1984), and moral obligation will be less likely to affect action. Similarly, Margolis (1982, 1991) has provided a formal analysis of these trade-offs in which individual differences in the propensity to weight obligations or personal consequences more heavily are represented as a factor determining the influence of two competing selves—the “altruistic” and the “selfish”—on an individual’s compliance behavior.

However, neither of these approaches considers processes that *change* the trade-offs between norms and personal consequences. A key question for our experiment is to discover how messages emphasizing either norms or personal consequences alter the trade-off, if indeed they do. What changes in the tax schema are induced by such messages, and how do these changes alter the factors affecting the balance between moral obligation and self-interest?

### The Dynamics of Tax Schema Change

The dynamics of tax schema change involve the complex, ongoing processes related to the preparation of tax returns that the taxpayer goes through during the tax season (Carroll 1989). Prior attitudes, habitual behavior, and new information interact with changes in the taxpayer’s environment in undoubtedly complex ways. The experimental interventions provided a very small amount of additional information in a persuasive setting to these many other changes that affected the participating taxpayers. We limit the discussion of dynamic change to those hypotheses relevant to the experimental manipulations. Our model relies heavily on the principle of cognitive consistency. Consistency theorists (e.g., Festinger 1957; Heider 1958; Newcomb 1968) presume that people desire consistency among related attitudes as well as between attitudes and behavior. The resulting “strain” toward consistency implies that a change in one component of the tax schema should be associated with systematic and predictable changes in other components of the schema structure.

Cialdini’s (1989) imaginative application of consistency research to tax compliance problems suggests that both kinds of consistency (attitude-attitude and attitude-behavior) provide the basis for understanding the impact of normative appeals on the tax schema. In the experiment, the normative appeal provided information about the social impact of the TRA, argued that these social consequences increased the fairness of the tax laws, and linked this information to the argument that good citizens should report their taxes honestly. If consistency is a driving force in schema change, the following testable hypotheses should hold. The normative information videotape should trigger the following sequence

of changes among cognitively linked concepts (relative to a control group): (1) increased awareness of social consequences; (2) increased reliance on social consequences (an impersonal standard) in fairness judgments; (3) increased perceptions of the fairness of the TRA; (4) increased levels of moral obligation to comply honestly; (5) increased levels of compliance behavior.

The change processes outlined above are logically derived from models of schematic processing. However, the principle of consistency also suggests cognitive inertia. Changes may be resisted, particularly for deeply held values such as moral obligation, which can be quite stable (Rokeach 1973), as well as for habitual behavioral routines (Carroll 1989). Moreover, experimental work suggests that consistency-related changes are most likely to occur for closely linked concepts (Judd et al. 1991). Accordingly, our second testable hypothesis is that information about the TRA is likely to have a stronger impact on specific policy evaluations than on more basic values and behavior. Media research reflects this specificity principle, in that media messages often influence knowledge and specific attitudes but rarely have an impact on values more directly linked to behavior (Kinder and Sears 1985; Tyler and Lavrakas 1985).

The same two-edged impact of consistency should hold for messages that emphasize personal consequences, although less attention has been paid to the possible impact of such messages on the tax schema. We should note that the personal consequence message in the experiment involved specific tax changes for individuals, with an accompanying emphasis on the use of aggressive strategies to minimize tax liability. We did not use the deterrence focus of the Schwartz and Orleans (1967) experiment because the TRA did not change penalties. Unlike that earlier experiment, therefore, we expected the personal consequences message to have a negative rather than a positive impact on compliance. Specifically, exposure to information about personal consequences should parallel the effects of exposure to normative information: (1) increased knowledge of consequences relating to individual (as opposed to the social) impact of the TRA; (2) increased reliance on personal consequences in fairness judgments; (3) decreased perceptions of the fairness of the TRA; (4) decreased levels of moral obligation to comply honestly; (5) increased avoidance and noncompliance behavior.

Avoidance and noncompliance may be facilitated by two processes. First, the tax schema of those made more aware of the personal impact of the TRA may be increasingly organized along self-interest lines, minimizing the salience of norms and principles. An increased impact of personal impact on judgments of the fairness of the TRA would provide evidence for this shift. Second, those who have more knowledge about personal impact may also

develop a more sophisticated understanding of different opportunities to minimize, legally or illegally, their tax liability.

## II. METHOD

### Participants

The participants were Long Island, NY, taxpayers. The sampling frame was provided by the IRS using the random selection techniques developed by the Taxpayer Compliance Measurement Program (TCMP) for drawing samples weighted by taxpayer characteristics. The IRS provided the names and addresses of a sample of 1,142 taxpayers who lived in zip code areas surrounding the State University of New York at Stony Brook (SUNY-Stony Brook), geographic proximity being necessary because participation required coming to the university. This sample was selected to reflect middle- and upper-income individuals who were most likely to be affected by the Tax Reform Act. Specifically, all members of the sample had completed itemized returns during the previous three years. Moreover, self-preparers were oversampled (two-thirds of the sample) relative to their proportion in the population, as were taxpayers filing Schedule C (business income, 25 percent of the sample) and Schedule E (supplemental income, 10 percent of the sample).

In December 1987, advance letters were sent to all taxpayers in the sample. These letters briefly described the nature of the study, the National Science Foundation's and SUNY-Stony Brook's sponsorship, offered \$50 compensation for participation in the research, and indicated that a project representative would contact the recipient to request participation. We then attempted to contact each taxpayer who had a listed telephone number (71 percent of the sample). In all cases, the taxpayer specified on the previous return was contacted, as opposed to an individual randomly selected within the household. In the case of joint returns, the individual who claimed to hold the primary responsibility for filing taxes was the designated respondent.

One hundred fifty-four of the 589 individuals (26 percent) who were successfully contacted and eligible participated in the research.<sup>1</sup> Seventy-six percent of the participants were men. The

---

<sup>1</sup> The complete sample disposition breaks down as follows: 29 percent had unlisted phone numbers and could not be contacted to participate; 8 percent were ineligible due to health or language problems, or lengthy travel plans; 11 percent of the cases were unresolved during the scheduled period of the first phase of the experiment (consisting of taxpayers that we unable to contact by phone despite numerous attempts or who agreed to participate but never showed up at the laboratory for their scheduled sessions); and 38 percent directly refused to participate. Because the sample was provided by the IRS and we did not have access to individual-level data, we know very little about the characteristics of the nonparticipants. From the aggregate tax return data provided by the IRS, we know that the nonparticipants had a substantially lower income (1987 income mean = \$47,372) than the participants (1987 income

median age was 47, with a range of 27 to 77. Ninety percent were married. The sample was extremely well educated, with 64 percent reporting at least a four-year college degree. Eighty-two percent were employed full time. The median reported total income was between \$50,000 and \$75,000, consistent with the mean income reported on their 1987 tax returns (mean = \$57,577, ranging from \$5,360 to \$359,067).

### Procedures and Data Sources

The experiment proceeded in several stages, with data obtained at several points in time.

#### 1. *Experimental Session*

The first phase of the study took place in the SUNY-Stony Brook Political Science Department's Media Research Laboratory in January 1988. On arrival to the laboratory, all participants read and signed a detailed informed consent form. They were randomly assigned to one of the three experimental conditions (normative TRA information, personal consequences TRA information, or a control group).<sup>2</sup> In the two TRA information conditions, each participant viewed a 30-minute videotape describing various aspects of the 1986 TRA.

Those randomly assigned to the *normative information* condition viewed a videotape we created that emphasized the normative principles underlying the tax reform movement and law. The tape

---

mean = \$57,577), probably a reflection of education differences (see main text). Our impression is that many of the participants were motivated out of respect for scholarly research rather than the \$50 compensation. Finally, self-preparers appeared particularly likely to agree to participate—78 percent were self-preparers, compared to 66 percent in the original sample.

<sup>2</sup> An explanation of the number of taxpayers in each experimental condition, as well as the number of cases used in the reported analyses, is in order. As noted earlier, 154 taxpayers agreed to participate and took part in the laboratory session. As is appropriate, these participants were randomly assigned to one of the three conditions. However, both time and response rate constraints required a modification of our random assignment procedures. Time was limited because the laboratory sessions had to be completed during January, before the pre-filing surveys were scheduled to begin. Because we were not as successful in persuading eligible taxpayers to participate in the study as we had originally hoped, the dwindling time made it necessary halfway through the experiment to modify the random assignment procedure. We allocated more respondents to experimental treatments relative to the control group because the critical predictions pertain to those two treatment groups. As a result, we obtained 60 subjects in the normative and 58 in the personal consequences conditions and 36 in the control condition. Minor sample attrition reduced these numbers. We report the attitudinal and self-report data for those taxpayers for whom there are complete data from all three interviews (total  $n=140$ , 55 in the normative condition, 54 in the personal consequences condition, and 31 in the control group). None of the results differ for analyses based on the total number of subjects available at each separate point in time. We requested tax return data from the IRS for the entire participating sample ( $n=154$ ), but because of missing data the number of cases reported in those analyses is slightly smaller. Those sample sizes are reported in Table 1.



began by describing public dissatisfaction with the previous tax system, which was attributed to perceptions of unfairness and needless complexity. A brief discussion of the politics of tax reform followed, with an emphasis on bipartisan support and the defeat of special interest groups. The third section considered in detail the impact of tax reform on the fairness of the tax system. Information about the reduction or elimination of special tax treatment items (e.g., tax shelters), the shift in the tax burden from the poor to the wealthy, and the minimum tax on corporate profits was presented and rationales provided. The tape concluded with public opinion data describing the importance Americans place on the norms of social responsibility and patriotism, emphasizing how these norms are related to tax compliance. The videotape itself was hosted and narrated by a middle-aged man, the setting was a book-filled office, and charts, figures, and news video clips were interspersed throughout the narrative.

Those randomly assigned to the *personal consequences information* condition viewed an edited version of a commercial tax planning videotape discussing the 1986 tax law changes. The tape opened with a quote from Judge Learned Hand: "There is nothing sinister in so arranging one's affairs to keep taxes as low as possible. . . . Nobody owes any public duty to pay more than the law demands." The use of aggressive legal strategies to minimize one's tax liability was emphasized throughout the tape. We selected information about strategies for dealing with changes due to the TRA that were most relevant to our subject population, including changes in common deductions, capital gains investments, aspects relevant to homeowners, retirement planning, small businesses, and tax shelters. Although it is unlikely that all of the strategies outlined in the tape were directly relevant to each participant's own tax situation, the tape did provide numerous examples of average taxpayers using aggressive strategies to minimize tax liability, thereby providing extensive social justification for that orientation. The tape ended with a discussion of the importance of aggressive tax planning and good recordkeeping. The format of this tape was identical to the normative tape: middle-aged speakers in offices, with a liberal use of charts and figures.

Immediately after viewing the videotape, the taxpayers in these two conditions completed a brief questionnaire that included knowledge tests of what they learned from the videotapes as well as indicators of initial attitude change. (See the Appendix for all variables used in the reported analyses.) Those assigned to the *control* condition completed the same questionnaire on arrival to the laboratory, without receiving any information about the TRA. They then participated in a simulation tax decision task that was not related to the TRA. Therefore, they serve as a control to the two experimental conditions because they are part of the sample willing to participate in the study, but they did not receive any informa-

tion in the course of the experimental session pertaining to the TRA.

The experimental laboratory session lasted approximately one hour. At the conclusion, the participants were paid \$50 in cash. Although they had been told in the initial contact letter that they could keep the money or donate it to charity, for a number of reasons the charity option was not made explicit at the time of payment. We suspect the vast majority kept the money, although some may have donated it to a charity on their own. It is noteworthy that six of the seven individuals who spontaneously requested that we donate the money directly to a charity in their name had viewed the normative information video, which concluded with a message about the importance of social responsibility.

## 2. Prefiling Interviews

Professional interviewers contacted the participating taxpayers and scheduled an in-home interview two weeks to one month after the laboratory sessions (primarily in February 1988). All of these interviews took place before the taxpayer had filed the 1987 return. All but one of the experimental participants were successfully interviewed at this time. The one-hour interviews covered numerous tax-related topics.

## 3. Postfiling Survey

In June and July of 1988 (approximately six months after the experimental sessions and after the 1987 returns had been filed), attempts were made to recontact all the participants for a final 20-minute telephone interview. A total of 140 of the original 154 taxpayers (91 percent) were successfully interviewed at this point, indicating remarkably little attrition in this study. Questions on the postfiling telephone survey included reactions to the TRA, information search strategies engaged in during the tax season, and self-report of compliance on the 1987 return.

## 4. Tax Return Data

Finally, we obtained aggregate-level tax return information for each experimental group from the IRS, following procedures designed to ensure that the IRS could not identify the participating individuals.<sup>3</sup> Data for both the 1985 and 1987 returns were made available, as well as the calculated differences between these two

<sup>3</sup> We provided the IRS with the social security numbers of the subjects in the three experimental groups as well as the social security numbers for three equal-sized groups of randomly selected *non*participants from the original sample. The use of the "decoy" groups served to insure that the IRS did not know which taxpayers participated in the experiment. The IRS in turn provided us with aggregate level data for the six (three experimental and three decoy) groups of taxpayers. We combined the three decoy groups into one, labeled "nonexperimental control" in Table 1.

years. The tax return information included summary information such as total reported income, adjusted gross income, and income tax, as well as more specific income and deduction line items. The data consist of group means and standard deviations, so that no individual return information was disclosed. The aggregate-level data meant, of course, that we could not link specific individual survey responses or individual characteristics to the taxpaying data.

### III. COMPARISONS WITH THE SCHWARTZ AND ORLEANS EXPERIMENT

To facilitate a comparison of our results with the Schwartz and Orleans (1967) results, we describe their experiment and clarify some key differences between the two studies. That earlier innovative field experiment involved randomly assigning respondents in a larger survey of taxpayer attitudes to three different questionnaires. A control group completed the basic interview, while the two treatment groups were asked additional questions, one set focusing on normative issues, the other set on deterrence issues. The normative questions were "designed to arouse motives for taxpaying ranging from guilt at violation to a patriotic desire to support the government in its most valued activities" (Schwartz and Orleans 1967:287), whereas the deterrence questions emphasized the punishments associated with noncompliance.<sup>4</sup>

The survey took place in March 1961, a month before the tax filing deadline. The IRS provided aggregate pre- and postsurvey tax return information for each of the three treatment groups, permitting an examination of the different questionnaire formats on taxpaying behavior. The normative group showed a significantly greater increase in reported income (in comparison to the control group), and a significantly greater increase in taxes paid (in comparison to both the deterrence and control groups). Thus, the results suggest that the additional "conscience" questions triggered or strengthened the normative bases of compliance, resulting in increased compliance. Or, to rephrase in our terms, activating the normative processing route was more effective in increasing compliance than activating the personal consequences processing route.

Like the Schwartz and Orleans experiment, our research is concerned primarily with understanding the relationship between

---

<sup>4</sup> The Schwartz and Orleans (1967) "conscience" manipulation consisted of seven questions, including reactions to education programs to remind citizen of their tax obligations, tax support for federal government activities, evaluations of tax "swindlers," comparisons of tax cheaters and draft dodgers, and a question about guilt. The "sanction" manipulation consisted of eleven questions, including experiences with the IRS, reactions to automatic reporting and the use of computers, reactions to various possible penalties for noncompliance, and estimates of the probability the IRS would examine the respondent's return for violations.

the two competing images of choice behavior and tax compliance. Moreover, we also have the unique opportunity to examine the impact of the experimental treatments on officially documented tax-paying behavior. However, the two experiments differ on several important dimensions. First, we have operationalized the independent variables or treatments quite differently by basing the normative and personal consequences appeals on the specific changes brought about by the TRA, because an underlying theme of our larger research program is how citizens adapt to legal changes. In addition, because the TRA did not substantially affect the penalties associated with noncompliance, the self-interest analogue focused on specific changes, providing information on how to take legal advantage of these changes to reduce one's tax liability. Unlike the deterrence-oriented questions in Schwartz and Orleans, which theoretically should have increased the amount of income and tax reported, our personal consequences message should reduce the amount of reported income and taxes paid. Moreover, whereas the Schwartz and Orleans normative manipulation consisted of a very diffuse set of questions (see note 4), the normative appeal in this experiment is more focused, emphasizing the positive social consequences and increased fairness resulting from the new law, messages that were emphasized consistently by proponents of tax reform. Finally, the format used to implement the independent variables differs considerably from that used by Schwartz and Orleans. Their open-ended questions embedded in the larger survey did not impart information about taxes, but rather encouraged the respondents to reflect on their own attitudes and values. Use of the detailed half-hour videotapes in this experiment reflects our concern with information acquisition and use in the context of legal change (see Scholz, McGraw, and Steenberg 1991 for a discussion).

#### IV. RESULTS

We begin by examining the impact of the experimental treatments on the official tax return data, as well as self-reported compliance behavior, in keeping with our goal of attempting to replicate the Schwartz and Orleans (1967) results. We then consider the impact of the treatments on knowledge of, and attitudes about, the TRA. Because the sample sizes are small and the predicted relationships "noisy," we report effects that are significant at the less stringent .1 alpha level.

##### **Behavioral Effects: The Tax Return Data and Self-reported Noncompliance**

Table 1 summarizes changes in tax return data for the three experimental groups, as well as the data for the randomly selected nonparticipants (see note 3). "Bottom line" tax items such as total

Table 1. Mean Differences, 1987-1985, in Tax Return Data

	Experimental Subjects			Nonexperimental Control*
	Normative	Personal Consequences	Control	
I. Income items				
Total income	\$6,972 (\$16,022)	\$6,912 (\$15,555)	\$11,377 (\$11,627)	\$8,908 (\$22,938)
Adjusted gross income	9,594 ( 16,778)	9,477 ( 15,014)	13,395 ( 11,518)	10,295 ( 21,722)
Other income	80 ( 429)	238 ( 2,119)	939 ( 5,097)	-18 ( 1,203)
Gross dividends	352 ( 1,067)	328 ( 1,111)	263 ( 1,483)	27 ( 806)
Interest	-10 ( 748)	47 ( 2,261)	227 ( 912)	76 ( 1,610)
II. Deduction items				
Employee business expenses	-346 ( 968)	-395 ( 2,555)	-408 ( 1,253)	-161 ( 1,112)
Cash contributions	169 ( 515)	247 ( 762)	449 ( 1,112)	241 ( 486)
Noncash contributions	56 ( 371)	23 ( 279)	-34 ( 251)	108 ( 692)
Medical deductions	34 ( 610)	-188 ( 1,638)	-356 ( 900)	-74 ( 1,208)
Miscellaneous deductions	-31 ( 1,477)	264 ( 3,006)	27 ( 1,200)	-245 ( 1,246)
Income tax	1,087 ( 3,785)	986 ( 2,877)	1,665 ( 2,840)	1,474 ( 7,490)

NOTE: The entries are the mean differences between the information reported on the 1987 tax return and the 1985 tax return. Amounts are rounded to the closest dollar. Positive values indicate larger amounts on the 1987 return. Standard deviations are in parentheses.

Cell sizes: Normative condition—from 53 to 58; personal consequences condition—from 52 to 57; experimental control—from 31 to 35; nonexperimental control—from 121 to 144.

\* See note 3 for explanation of this group of nonexperimental controls.

income, adjusted gross income, and total tax are reported (as in Schwartz and Orleans 1967), as well as more specific income and deduction items. The means of the differences between the two years (1985 and 1987) are reported, along with the standard deviation for each group. Because the experimental groups were randomly assigned, the groups would be expected to have the same preintervention (1985 tax season) values on the various line items. Analyses of the 1985 tax return data confirm that random assignment was successful, as there were initially no significant differences among the three experimental groups. All three groups would be expected to report approximately the same year-to-year average change in the *absence* of the experimental treatments. Thus, any difference between the groups in year-to-year changes is attributable to the impact of the experimental messages on compliance. If the normative video strengthened the normative bases of compliance, we would expect to see increases in reported income and taxes paid in that group, relative to the other two groups. Similarly, if the personal consequences videotape successfully provided strategies and justifications to minimize tax liability, we would expect to see decreases in income and taxes paid. According to the same logic, the normative condition taxpayers should have lower deductions and the personal consequences condition taxpayers higher deductions, relative to the control condition.

In fact, on none of the measures are the experimental groups in each row significantly different in the expected direction at the  $p < .05$  level (one-tailed test), the level at which the normative appeal had a significant effect on adjusted gross income and on taxes in the Schwartz and Orleans experiment (1967:Table 5). As expected, the personal consequences group reported less total income and adjusted gross income than the experimental control group at the marginal .1 level, the significance level associated with the comparable "sanction threat" group reported in Schwartz and Orleans. However, contrary to predictions, the taxpayers in the normative condition did not report more income. Rather, they reported amounts similar to those in the personal consequences condition, although the amounts for the normative condition do not differ from the control group. Moreover, the larger increase in income reported by the experimental control group relative to both treatment groups and the nonexperimental control group suggest that there may be some anomaly with the income reported by the control group, suggesting caution in interpreting the marginally significant finding as confirmation that the personal consequences condition led those individuals to report less income.

One could argue that the high variances and small sample sizes are responsible for the lack of significant results. However, the apparent volatility in income created comparably large variances in the Schwartz and Orleans experiment (although, of course, the income of our sample was much larger than the income

reported in that 1960s experiment). The volatility of tax reports underscores the need for individual-level tax data. In both experiments, only aggregate data were made available for analysis, which eliminated the possibility of data transformations and alternative analyses to understand and minimize the problems associated with the high variances within each group. However, the results in Table 1 suggest that large variances are not the only problem in finding support for the predicted behavioral effects. Only half (six of the eleven) of the differences between the means in the two experimental groups are even in the expected direction.

Further evidence that there were no significant differences in compliance levels as a result of the experimental treatments can be found in the self-reported compliance behavior data. The participants were asked a number of questions about previous non-compliance as well as about noncompliance on the 1987 tax return (see Appendix for question details). Those who gave at least one response that could be inferred to reflect noncompliance were classified as noncompliers, while those who gave no noncompliant responses were classified as compliers. The percentage of respondents who reported noncompliant behavior on the 1987 tax return did not differ as a function of experimental condition ( $\chi^2 < 1$ ). Indeed the pattern is contrary to expectations, with 30.9 percent of the normative condition taxpayers classified as noncompliant, in contrast to 24.1 percent and 21.5 percent of the taxpayers in the personal consequences and control conditions, respectively. Analyses of changes in compliance as a function of experimental condition similarly yielded null results.

In contrast to Schwartz and Orleans (1967), we found no evidence that the normative and self-interested treatments had an impact on taxpaying behavior, either as reported by the taxpayer or as evident in officially documented records. The converging null results for the self-report and IRS data strengthen our conclusion that the experimental interventions did not affect compliance.<sup>5</sup> We consider possible explanations for the contradictory sets of results below.

### Knowledge, Information Search, and Taxpaying Strategies

We consider next whether the experimental treatments had a systematic impact on a number of other important variables, such as tax-related knowledge and attitudes. Because an obvious explanation for the null behavioral results lies in the possibility of weak

---

<sup>5</sup> Since we have no means of directly linking the two measures in this study, we cannot comment on the validity of the self-report compliance data. The two types of compliance measures have yielded different results in other circumstances (Hessing, Elffers, and Weigel 1988). However, their officially documented variable was quite different from the one used in this study. Rather than using raw tax return data, as we have here, Hessing et al. (1988) had tax inspectors classify returns as reflecting evasion or nonevasion.

or ineffective manipulations, it is important to determine what, if any, impact the experimental treatments did have.

Table 2 presents summary data on the impact of the experimental messages on knowledge about the TRA as well as on behavior in preparing the 1987 tax return. Two types of knowledge were measured. The first was knowledge about specific changes on tax items (e.g., whether the TRA changed the deduction for interest on consumer debt and credit cards); the second was knowledge of the intended social consequences of the TRA (e.g., whether the TRA was intended to change the number of wealthy people who pay no tax). The data in Table 2 indicate that, as expected, taxpayers exposed to the normative videotape were more knowledgeable about the intended social consequences of the TRA, and that knowledge persisted. In contrast, the effects of the personal consequences videotape on knowledge were weaker. Immediately following exposure to the videotape, taxpayers in the personal consequences condition scored higher on knowledge of specific consequences as measured by the true-false test but that knowledge did not generalize to the other knowledge tests.

**Table 2.** Impact of Experimental Treatment on Knowledge and Tax Season Behavior

	Experimental Condition			Significance
	Normative ( <i>n</i> = 55)	Personal Consequences ( <i>n</i> = 54)	Control ( <i>n</i> = 31)	
<b>1. Knowledge of specific changes</b>				
Postexperiment true/false	.80 <sub>b</sub>	.98 <sub>a</sub>	.79 <sub>b</sub>	<i>p</i> < .001
Postexperiment	.51 <sub>a</sub>	.49 <sub>a</sub>	.45 <sub>a</sub>	N.S.
Prefiling	.30 <sub>a</sub>	.25 <sub>a</sub>	.26 <sub>a</sub>	N.S.
Postfiling	.61 <sub>a</sub>	.50 <sub>a</sub>	.56 <sub>a</sub>	N.S.
<b>2. Knowledge of social consequences</b>				
Postexperiment true/false	.87 <sub>a</sub>	.52 <sub>b</sub>	.46 <sub>b</sub>	<i>p</i> < .001
Postexperiment	.73 <sub>a</sub>	.58 <sub>b</sub>	.53 <sub>b</sub>	<i>p</i> < .001
Prefiling	.67 <sub>a</sub>	.58 <sub>b</sub>	.58 <sub>b</sub>	<i>p</i> < .1
<b>3. 1987 tax season behavior</b>				
Information search	1.76 <sub>b</sub>	2.09 <sub>a</sub>	1.68 <sub>b</sub>	<i>p</i> < .1
Discussion sources	2.18 <sub>ab</sub>	2.42 <sub>a</sub>	1.65 <sub>b</sub>	<i>p</i> < .1
Refigure taxes	.44 <sub>b</sub>	.71 <sub>a</sub>	.57 <sub>ab</sub>	<i>p</i> < .05

NOTE: The values in the two sets of knowledge rows are the mean percentage correct responses at the different assessment points (see Appendix for details about the measures).

The variables under the heading "1987 tax season behavior" are the number of media/print information sources consulted, the number of types of targets with whom the respondent discussed his/her tax return, and the percentage of respondents who to minimize their tax liability tried to figure out their taxes in several different ways. These three variables are from the postfiling survey.

The values in the final column are the significance levels of the differences among the three experimental conditions, according to an analysis of variance. N.S. = not significant.

Means in each row not sharing a common subscript differ from each other at least the *p* < .1 level, according to the Duncan multiple-range statistic.



Although the personal consequences message did not have any discernible long-term impact on knowledge, it did have an impact on the prefiling behavior. Two explicit messages provided in the personal consequences videotape were that (1) it is important to be aggressive in understanding the implications of the TRA as it applies to one's tax situation and (2) it is acceptable within legal limits to attempt to minimize one's tax liability. If the first message was received and accepted, we would expect to see evidence of increased information search prior to the filing of the 1987 return among the personal consequences condition taxpayers. Several indicators of information search and usage were included on the postfiling survey, including use of media, print, and professional information sources, as well as discussing taxes with family, friends, and coworkers (see Scholz, McGraw, and Steenberg 1991 for an extended discussion). Consistent with the videotape's message, the taxpayers in the personal consequences condition consulted more media and print sources of information, as well as engaging in more discussion about taxes, than did the taxpayers in the other two conditions. As noted above, this search did not produce greater knowledge about the TRA (at least as measured by our tests), perhaps because the search was focused specifically on issues related to the individual's return.

The personal consequences videotape also had an impact on strategies to minimize one's tax liability, consistent with the second message. In the postfiling survey respondents were asked, "In preparing your 1987 return, did you try to figure out your tax in several different ways so that you could pick out the way that cost you the least taxes?" As indicated in Table 2, 71 percent of those in the personal consequences condition said they did, in contrast to 44 percent of the normative condition taxpayers and 57 percent of the control condition taxpayers ( $\chi^2=7.32, p<.05$ ). In sum, taxpayers exposed to the personal consequences videotape sought out more sources of information prior to completing the 1987 return, and they were more likely to work through their return in different ways to minimize their tax liability.

### Attitudes about the TRA

The data in Table 3 report the perceived impact of the TRA on three target groups (oneself, the poor, and the very wealthy) as well as perceptions of the overall fairness of the TRA and the fairness of the taxpayer's own particular situation. Recall that the central themes of the normative information videotape were that the TRA would reduce and simplify taxes for the poor and close tax loopholes used by the rich, and that as a result the system would be more fair. Those messages were accepted by the taxpayers who viewed that videotape. Subjects in the normative condition, more than subjects in the other two conditions, believed that the tax lia-

**Table 3.** Impact of Experimental Treatment on Perceptions of the TRA: Impact and Fairness

	Experimental Condition			Significance
	Normative ( <i>n</i> = 55)	Personal Consequences ( <i>n</i> = 54)	Control ( <i>n</i> = 31)	
<b>Impact</b>				
1. Personal impact				
Postexperiment	3.29 <sub>a</sub>	3.35 <sub>a</sub>	3.37 <sub>a</sub>	N.S.
Prefiling	3.24 <sub>a</sub>	3.46 <sub>a</sub>	3.48 <sub>a</sub>	N.S.
Postfiling	3.55 <sub>a</sub>	3.78 <sub>a</sub>	3.65 <sub>a</sub>	N.S.
2. Impact on the poor				
Postexperiment	1.65 <sub>a</sub>	2.07 <sub>b</sub>	2.67 <sub>c</sub>	<i>p</i> < .001
Prefiling	1.89 <sub>a</sub>	2.39 <sub>b</sub>	2.29 <sub>b</sub>	<i>p</i> < .01
Postfiling:	N.A.			
3. Impact on the rich				
Postexperiment	3.33 <sub>a</sub>	2.79 <sub>b</sub>	3.07 <sub>ab</sub>	<i>p</i> < .1
Prefiling	3.80 <sub>a</sub>	3.59 <sub>a</sub>	3.69 <sub>a</sub>	N.S.
Postfiling:	N.A.			
<b>Fairness</b>				
1. Fairness of TRA for respondent's taxes				
Postexperiment	3.04 <sub>a</sub>	2.57 <sub>b</sub>	2.67 <sub>b</sub>	<i>p</i> < .05
Prefiling	2.82 <sub>a</sub>	2.43 <sub>b</sub>	2.51 <sub>ab</sub>	<i>p</i> < .1
Postfiling	2.76 <sub>a</sub>	2.37 <sub>b</sub>	2.32 <sub>ab</sub>	<i>p</i> < .1
2. Overall TRA fairness				
Postexperiment	3.73 <sub>a</sub>	2.91 <sub>b</sub>	3.20 <sub>b</sub>	<i>p</i> < .001
Prefiling	3.40 <sub>a</sub>	2.89 <sub>b</sub>	3.19 <sub>ab</sub>	<i>p</i> < .01
Postfiling	3.09 <sub>a</sub>	2.67 <sub>b</sub>	2.61 <sub>b</sub>	<i>p</i> < .1

NOTE: Entries are means scores for each experimental condition. For impact judgments, higher values reflect perceptions of increased tax liability due to the TRA; for fairness judgments, higher values reflect greater perceived fairness.

The values in the final column are the significance levels of the differences among the three experimental conditions, according to an analysis of variance. N.S. = not significant.

Means in each row not sharing a common subscript differ from each other at least the *p* < .1 level, according to the Duncan multiple-range statistic.

N.A. = not available (question not asked in postfiling survey).

bility of the poor had decreased and, less strongly, that the tax liability of the rich had increased. In contrast, there were no differences among the three experimental groups in perceptions of their own tax liability (the aggregate mean of approximately 3.5 on that scale represents judgments between "no change" and paying "somewhat more").

The normative condition taxpayers also believed that the TRA resulted in a fairer tax system, and those beliefs persisted over the six months of the study period. The message of a fairer tax system also influenced perceptions of the fairness of their own tax situation. In contrast to those who viewed the personal consequences tape, the normative condition taxpayers believed that the amount of taxes they personally paid was fairer as a result of the TRA. These judgments also persisted.

Changes in some important perceptions occurred over time. As is evident in Table 3, perceptions of personal liability increased

over the course of the tax season for taxpayers in all three conditions. In contrast, perceptions of fairness, both personal and overall, diminished over time. Repeated-measures analyses of variance provide statistical support: the personal impact judgments increased over time ( $F(2,274)=6.76, p<.002$ ), whereas perceptions of fairness decreased over time ( $F(2,274)=8.81, p<.001$ , and  $3.55, p<.05$ , for overall and personal fairness, respectively). There were no interactions with experimental condition, indicating that these trends occurred among all three groups of taxpayers. Pessimism appears to overcome initial optimism as taxpayers work their way through their tax returns.

To determine whether the standards used to evaluate the fairness of the TRA were influenced by the experimental messages, we also examined the relationship between impact (personal and on the poor) and fairness perceptions. The zero-order correlations in the first two panels of Table 4 support our expectation that the experimental messages would substantially change the criteria used to evaluate the fairness of the TRA. Consider first the correlations

**Table 4.** Determinants of Fairness Perceptions: Personal and Impersonal Standards

	Experimental Condition		
	Normative ( <i>n</i> =55)	Personal Consequences ( <i>n</i> =54)	Control ( <i>n</i> =31)
Personal Impact—Fairness Correlations			
Postexperiment	-.15	-.43***	-.57***
Prefiling	-.19	-.52***	-.16
Postfiling	-.36**	-.57***	-.16
Impact on the Poor—Fairness Correlations			
Postexperiment	-.35**	-.13	-.30
Prefiling	-.30*	-.12	-.09
Postfiling: N.A.			
Regression Model (Dependent Variable=TRA Fairness) <sup>a</sup>			
Personal impact	-.19	-.52***	-.15
Impact on the poor	-.40**	-.11	-.07
Adjusted <i>R</i> <sup>2</sup>	.17**	.25***	.00

NOTE: The values in the top two panels are the zero-order correlations between impact judgments and evaluations of fairness of the TRA. Negative coefficients mean that the TRA is viewed as more fair to the extent that tax liability of each target (self and poor) has decreased. The data in the bottom panel are the results of separate regression equations (in each column), where the dependent variable is TRA fairness judgment. The entries are standardized regression coefficients (betas).

<sup>a</sup> Based on prefiling survey data.

\* $p<.05$

\*\* $p<.01$

\*\*\* $p<.001$

between personal impact and fairness. The relationship was generally much stronger among the personal consequences taxpayers, suggesting that exposure to that videotape strengthened the cognitive link between personal impact and fairness. In contrast and as expected, the normative videotape strengthened the link between the impersonal standard of the tax liability of the poor and fairness.

The multivariate regression results reported in the bottom panel of Table 4 underscore the results from the correlational analysis. In the control group, representing the “natural” state of affairs, fairness beliefs are more strongly determined by self-interest than the impersonal criterion. This self-interest bias was strengthened among those taxpayers who viewed the personal consequences videotape. However, this pattern was reversed for the normative condition taxpayers, whose evaluations of the fairness of the TRA were largely determined by judgments of its impact on the poor.<sup>6</sup>

### Moral Obligation to Comply

According to the logic of the schematic model developed earlier, not only should the experimental messages have had an impact on TRA-specific knowledge and evaluations, but they also should have influenced concepts linked with these in the tax schema, such as a moral commitment to fully comply with the law. However, our results yielded no evidence that any such effects occurred. In the pre- and postfiling surveys the respondents were asked, “How much of a moral obligation—that is, an obligation based on your own personal feelings of right and wrong—do you feel to be completely honest in filling out your tax return?” Analyses of these responses revealed no differences due to experimental condition at the prefiling survey ( $F < 1$ ), and only a very weak trend ( $p < .20$ ) for the expected higher levels of moral obligation among the normative condition subjects at the postfiling survey.

We also conducted a multivariate analysis regressing postfiling moral obligation on prefiling moral obligation, postfiling evaluations of the TRA (personal impact, personal fairness, and overall

---

<sup>6</sup> Although separate regression equations are reported in Table 4 for purposes of clarity, an analysis of the experimental sample as a whole, with dummy-coded terms for experimental condition and the appropriate multiplicative interaction terms between condition and the two impact standards, indicated that the observed condition differences are significant. The predictors of the TRA fairness judgments in that equation were: personal impact, impact on the poor, normative condition, personal consequences condition, interaction between normative condition and impact on the poor, and interaction between personal consequences condition and personal impact. The weight accorded to the impact on the poor was significantly stronger in the normative condition (when contrasted with the other two experimental groups,  $p < .05$ ), whereas the weight accorded to personal impact was significantly stronger in the personal consequences (again, when contrasted with the other two conditions,  $p < .05$ ).

fairness), and dummy codes for the experimental conditions. Again, the experimental treatments did not have an impact on postfiling moral obligation. The analysis indicated that moral obligation was quite stable over the six-month period between surveys (zero-order correlation is .53), consistent with the view that moral values are particularly enduring standards of behavior (Rokeach 1973). The only other variable attaining significance ( $p < .05$ ) in that analysis was personal fairness, with moral obligation enhanced among those taxpayers who believed that their own taxes were more fair as a result of the legal change. Neither perceptions of the overall fairness of the TRA nor perceptions of personal liability had an impact on moral obligation.

In short, the attitudinal data indicate that the experimental treatments had a systematic impact on policy-specific evaluations, knowledge, and search behavior, but that these changes did not extend to more basic moral values or behavior. A final regression analysis serves to clarify why the experimental messages did not have their expected impact on behavior. Table 5 reports regression coefficients for the critical variables considered in this study, with self-reported noncompliance on the 1987 tax return as the dependent variable. The analysis is not intended to reflect a complete model of noncompliance—only the variables most relevant to this study are included. Inclusion of the prefiling noncompliance measure serves to control for many of the excluded measures found to impact on compliance in other studies (Roth, Scholz, and Witte 1989). We therefore analyze the impact of the relevant predictors on compliance controlling for prior reported behavior.

Table 5. Determinants of Self-reported Noncompliance

	<i>B</i>	(S.E.)	Beta
Previous noncompliance	.12	(.11)	.09
Moral obligation	-.16	(.05)	-.26**
Personal impact	.03	(.04)	.06
Fairness of own taxes	.09	(.06)	.21
TRA fairness	-.06	(.04)	-.14
Normative condition	.12	(.10)	.14
Personal consequences condition	.04	(.10)	.05
Adjusted $R^2 = .06^*$ $n = 140$			

NOTE: The dependent variable is reported noncompliance on the 1987 tax return. All the predictors, other than previous noncompliance, are postfiling indicators. Higher values reflect noncompliance, higher levels of moral obligation, increased taxes due to the TRA, and increased fairness. The two condition variables are dummy-coded terms for the experimental treatments.

\* $p < .05$

\*\* $p < .01$

Reinforcing the null conclusions reported earlier, the experimental messages did not have a direct impact on reported taxpaying behavior. Nor did evaluations of the TRA, either liability or fairness. Rather, even controlling for prior behavior, noncompliance was more likely among those taxpayers who expressed lower levels of moral obligation. Estimation of various interactive models

indicated that the coefficients for the determinants of reported noncompliance did not differ as a function of the experimental treatments. In sum, these data suggest a disjunction between policy-specific attitudes, which were influenced by the experimental messages, and more basic values and behavior, which did not change as a result of the experimental treatments.

## V. DISCUSSION

This experiment confirms that both normative and personal consequences information have some systematic and predictable effects on an important set of tax-related attitudes and behavior. Taxpayers who viewed a videotape emphasizing the positive normative implications of the 1986 TRA (1) were more knowledgeable about the intended social consequences of the TRA, including being aware of shifts in the tax burden for the poor and wealthy; (2) believed that both the system overall and their own personal tax situations were more fair; and (3) adopted an impersonal, rather than self-interested, standard to evaluate the fairness of the TRA. In contrast, taxpayers who viewed a videotape emphasizing the legitimacy of self-interest and focusing on specific individual changes resulting from the TRA (1) were more likely to seek out information and discuss taxes during the tax season; (2) were more likely to work through their return in different ways to minimize their tax liability; and (3) exhibited the strongest self-interest bias in evaluating the fairness of the new law. The experimental messages had theoretically meaningful impacts that persisted beyond the immediate postexperimental session. The impact on knowledge and evaluations of the tax law did not, however, translate into changes in values or taxpaying behavior. Neither the self-report nor the official tax return data provided compelling evidence that the normative or personal consequences videotape affected compliance. Moreover, neither the level of moral obligation to fully comply with the law nor the link between moral obligation and compliance changed in response to the different messages. Thus, the experimental impacts were limited to those concepts in the tax schema that are more peripheral to compliance behavior.

What can we conclude about this failure to replicate the positive impact of normative appeals on actual compliance behavior found in the Schwartz and Orleans (1967) experiment? Given that the experiment did find expected impacts on TRA-specific attitudes and beliefs, we do not believe that our results can be attributed to weak operationalizations of the normative and personal consequences concepts. The clearest implication of the combined results is that impacts on compliance are at best difficult to achieve. Even Schwartz and Orleans found no effects for their deterrence appeal, and only marginally significant impacts for the normative appeal. Moreover, it is not at all clear which specific aspects of the mul-

tidimensional normative appeal in the Schwartz and Orleans experiment were responsible for the positive impact on compliance.

The numerous differences between the two experiments, including time (the 1960s vs. the 1980s), location (Chicago suburbs vs. Long Island), sample characteristics, and the political and cultural *Zeitgeist*, make it difficult to point conclusively to the reasons for the discrepant results. Indeed, it would be impossible to "prove" that any of those factors are responsible. However, if one accepts the validity of the results from both experiments, we can point to two methodological factors that may be important. The first is the timing of the intervention. Our intervention took place over three months prior to the tax filing deadline, whereas the Schwartz and Orleans survey preceded the deadline by only one month. The empirically observed rates of the temporal decay of induced attitude change are variable and contingent upon a number of complex factors (McGuire 1985). The data from this experiment indicate that to the extent that attitude and belief changes did occur (for the fairness and impact judgments), there was a moderate degree of decay over time. It is reasonable to assume that to the extent that changes in compliance were likely to occur, they would be most evident soon after the experimental intervention. The natural decay of message retention and associated attitude change, as well as the added complexity and "noise" associated with adapting to a new set of tax regulations, could only serve to attenuate any value and behavioral effects.

The second important factor is the specific implementation of the intervention. Our intervention focused on the transmission of detailed information about the new tax law that had implications for either normative or self-interest concerns. In contrast, the parallel Schwartz and Orleans interventions required the participating taxpayers to actively reflect on their own norms and values (in the "conscience" condition) or their own fears about formal penalties (in the "sanction" condition) by responding to open-ended probes in the survey. Actively engaging the individual through "value self-confrontation" techniques has produced belief, attitude, and behavior change in a variety of field settings (Rokeach 1973; Ball-Rokeach, Rokeach, and Grube 1984; Schwartz and Inbar-Saban 1988).<sup>7</sup> Formally, the value self-confrontation technique involves making behavior-relevant values salient, focusing attention on the link between these values and behavior, and pointing out possible discrepancies between endorsement of the value and undesirable or counternormative behavior. According to this line of reasoning, the Schwartz and Orleans "conscience" manipulation may have been more effective because the respondents were re-

---

<sup>7</sup> The effectiveness of the value self-confrontation method is variable, as it has produced significant behavior change in about half of its published applications (Schwartz and Inbar-Saban 1988).

quired to reflect on their own tax-relevant values and to consider the ways in which these values were inconsistent with illegal tax-paying behavior.

If this analysis is correct, then normative appeals made closer to the filing deadline (or repeated during the tax season) that actively engage the individual in moral reasoning and that connect taxpaying with values important to the individual may produce significant behavior change. In line with the logic of value self-confrontation, Cialdini (1989) has suggested that a "National Tax Test" television program aired immediately prior to the April 15 filing deadline and designed to actively engage the audience in thinking about their values and taxes might have a positive impact on compliance.

As a final point regarding the tax return data, our efforts demonstrate the possibility of cooperation between a government agency such as the IRS and the academic research community. Access to sensitive data bases can be achieved in a manner that protects the privacy and confidentiality of research participants (Boruch 1989; Boruch and Cecil 1979; Roth, Scholz, and Witte 1989). However, the variability of tax return data underscores the limitations of using only aggregated data. Alternative methods of access that can help mitigate the large within-group variances are needed to provide sufficiently sensitive tests for changes in tax behavior in future field experiments. We encourage future cooperative exchanges that will permit more precise specification and empirical tests of the relationships among attitudes, information, and compliant behavior.

In conclusion, we believe that understanding the structure of tax schema remains a useful conceptual framework for dynamic analysis. Within this schematic framework we specified a dynamic process of change: information about the TRA that emphasized either normative or personal consequences would first influence knowledge, information search behavior, and policy-specific evaluations, which in turn would influence basic core values such as moral obligation, and consequently influence the decision to comply. Although we did not replicate the effects on behavior found in Schwartz and Orleans (1967), we did find support for many of our expectations concerning the impact of normative and personal consequences information. Clearer specification of the nature and strength of schematic linkages remains an exciting challenge for future work. Continued theoretical development and empirical verification through the use of field experiments or panel studies will result in the development of a dynamic, cognitively informed model of compliance with tax and other laws.



## APPENDIX

## VARIABLES USED IN THE ANALYSES

1. **Knowledge of specific changes:** sum of correct responses to two part questions such as, "As a result of the TRA, did the deductions for state and local sales taxes: change, remain the same, or not sure?" If respondent answered "change," he or she was required to indicate whether the provision had increased, been reduced or restricted, or eliminated. Both parts of the question had to be answered correctly to be scored as a correct response. The same ten knowledge questions were asked at the postexperiment and prefilling surveys; a subset of five questions were asked at the postfiling survey.
2. **Knowledge of social consequences:** Same two-part format as above; six questions such as, "Did the law intend to change the number of poor people who pay income taxes?"
3. **1987 Tax season behavior:**
  - a. **Information search:** sum of use of following sources of information: newspapers, magazines, and television programs; income tax guides sold in stores; IRS publications; direct IRS assistance; paid tax professionals.
  - b. **Discussion sources:** sum of the following different targets with whom the respondent discussed taxes: spouse; family; close friends; people in your line of work; people in clubs and organizations you belong to.
  - c. **Refigure taxes:** "In preparing your 1987 return, did you try to figure out your tax in several different ways so that you could pick the way that cost you the least taxes?"
4. **Impact:**
  - a. **Postexperiment and prefilling:** "How much more or less in taxes do you think the *poor*/ the *extremely wealthy*/ *you*/ will have to pay in taxes as a result of the new tax laws?"
  - b. **Postfiling personal impact question:** "Do you think the new law increased, decreased, or did not change the amount of taxes you had to pay this year?"
5. **Fairness:**
  - a. **Personal (Postexperiment and prefilling):** "How would you rate the fairness of the taxes you will pay under the new law?" (Postfiling): "How would you rate the fairness of the taxes you had to pay under the new law?"
  - b. **Overall TRA:** "Overall, do you think that as a result of the TRA of 1986 the tax laws are a great deal fairer, somewhat fairer, somewhat less fair, a great deal less fair, or do you think there is no real difference in fairness?"
6. **Moral obligation:** "How much of a moral obligation—that is, an obligation based on your own personal feelings of what's right or wrong—do you feel to be completely honest in filling out your tax return?"

7. **Self-reported compliance:** Italicized responses counted as non-compliance; those who indicate at least one instance of non-compliance classified as noncompliant.
- a. **Previous noncompliance** (responses in reference to previous three tax returns):
    - i. **deductions:** "Do you think that the average amount of the deductions that you claimed were *definitely more* than what you were entitled to, *probably more*, probably less, definitely less or were they exactly the amount to which you were entitled?"
    - ii. **income:** For each of five different sources of income the respondent indicated having: "How likely is it that you left any income from these sources—even a small amount—off your federal tax return? Would you say that you *definitely have*, *probably have*, probably have not, or definitely have not?"
    - iii. **Overall:** "Do you think that the average amount of tax you paid was definitely greater than the amount you legally owed, probably greater than the amount you owed, *probably less*, *definitely less*, or exactly what you owed?"
  - b. **1987 tax return compliance:**
    - i. **deductions:** as above; plus, if respondent indicated taking any of five common deductions, "Suppose that you were absolutely certain that your 1987 tax return was going to be thoroughly audited by the IRS. Would you have claimed a *much smaller* deduction, a *slightly smaller* deduction, or exactly the same deduction?"
    - ii. **income:** as above; plus, if respondent indicated reporting any of six income sources, "Suppose you were absolutely certain that your 1987 tax return was going to be thoroughly audited by the IRS. Would you have reported a *much larger* amount, a *slightly larger* amount, or exactly the same amount?"

## REFERENCES

- BALL-ROKEACH, Sandra, Milton ROKEACH, and J. W. GRUBE (1984) *The Great American Values Test*. New York: Free Press.
- BORUCH, Robert F. (1989) "Experimental and Quasi-experimental Designs in Taxpayer Compliance Research," in J. A. Roth, J. T. Scholz, and A. D. Witte (eds.), *Taxpayer Compliance: An Agenda for Research*. Philadelphia: University of Pennsylvania Press.
- BORUCH, Robert F., and Joe S. Cecil (1979) *Assuring the Confidentiality of Social Research Data*. Philadelphia: University of Pennsylvania Press.
- CARROLL, John (1987) "Compliance with the Law: A Decision-making Approach to Taxpaying," 11 *Law and Human Behavior* 319.
- (1989) "A Cognitive-Process Analysis of Taxpayer Compliance," in J. A.

- Roth and J. T. Scholz (eds.), *Taxpayer Compliance: Social Science Perspectives*. Philadelphia: University of Pennsylvania Press.
- CIALDINI, Robert B. (1989) "Social Motivations to Comply: Norms, Values, and Principles," in J. A. Roth and J. T. Scholz (eds.), *Taxpayer Compliance: Social Science Perspectives*. Philadelphia: University of Pennsylvania Press.
- ETZIONI, Amitai (1988) *The Moral Dimension: Toward a New Economics*. New York: Free Press.
- FESTINGER, Leon (1957) *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.
- FRIEDMAN, Lawrence M., and Stewart MACAULAY (1977) *Law and the Behavioral Sciences*. Indianapolis: Bobbs-Merrill.
- GRASMICK, Harold G., and Robert J. BURSIK, Jr. (1990) "Conscience, Significant Others, and Rational Choice: Extending the Deterrence Model," 24 *Law & Society Review* 837.
- HARSANYI, John C. (1977) "Morality and the Theory of Rational Behavior," 44 *Social Research* 623.
- HEIDER, Fritz (1958) *The Psychology of Interpersonal Relations*. New York: Wiley.
- HESSING, Dick J., Henk ELFFERS, and Russel H. WEIGEL (1988) "Exploring the Limits of Self-Reports and Reasoned Action: An Investigation of the Psychology of Tax Evasion Behavior," 54 *Journal of Personality and Social Psychology* 405.
- HOCHSCHILD, Jennifer L. (1981) *What's Fair? American Beliefs About Distributive Justice*. Cambridge, MA: Harvard University Press.
- HYDE, Alan (1983) "The Concept of Legitimation in the Sociology of Law," 1983 *Wisconsin Law Review* 379.
- JUDD, Charles M., Roger A. DRAKE, James W. DOWNING, and Jon A. KROSNICK (1991) "Some Dynamic Properties of Attitude Structures: Context Induced Response Facilitation and Polarization," 60 *Journal of Personality and Social Psychology* 193.
- KINDER, Donald R., and David O. SEARS (1985) "Public Opinion and Political Action," in G. Lindzey and E. Aronson (eds.), 2 *Handbook of Social Psychology*. New York: Random House.
- KINSEY, Karyl A., and Kent W. SMITH (1987) "Income Tax Cheating: Opportunities and Preferences." Presented at the Midwestern Sociological Meetings, Des Moines.
- KOHLBERG, Lawrence (1984) *Essays on Moral Development, Vol. 2: The Psychology of Moral Development*. New York: Harper & Row.
- LANE, Robert E. (1986) "Market Justice, Political Justice," 80 *American Political Science Review* 363.
- LEMPERT, Richard (1989) "Humility Is a Virtue: On the Publicization of Policy-relevant Research," 23 *Law & Society Review* 145.
- LEWIS, Anthony (1982) *The Psychology of Taxation*. New York: St. Martin's Press.
- MANSBRIDGE, Jane J. (1990) *Beyond Self-Interest*. Chicago: University of Chicago Press.
- MARGOLIS, Howard (1982) *Selfishness, Altruism, and Rationality*. New York: Cambridge University Press.
- (1991) "Incomplete Coercion: How Social Preferences Mix with Private Preferences," in K. Monroe (ed.), *The Economic Approach to Politics*. New York: Harper/Collins.
- MCGRAW, Kathleen M., and John T. SCHOLZ (1988) "Norms, Social Commitment, and Citizen Adaptation to New Laws," in P. J. van Koppen, D. J. Hessing, and G. van den Heuvel (eds.), *Lawyers on Psychology and Psychologists on Law*. Amsterdam: Swets & Zeitlinger.
- McGUIRE, William J. (1985) "Attitudes and Attitude Change," in G. Lindzey and E. Aronson (eds.), 2 *Handbook of Social Psychology*. New York: Random House.
- NEWCOMB, T. M. (1968) "An Approach to the Study of Communicative Acts," 64 *Psychological Bulletin* 393.
- PIAGET, Jean (1965 [1932]) *The Moral Judgment of the Child*, trans. M. Gabain. Glencoe, IL: Free Press.

- RAWLS, John (1971) *A Theory of Justice*. Cambridge MA: Harvard University Press.
- REIS, Harry T. (1984) "The Multidimensionality of Justice," in R. Folger (ed.), *The Sense of Injustice: Social Psychological Perspectives*. New York: Plenum Press.
- REST, James R. (1984) "The Major Components of Morality," in W. L. Kurtines and J. L. Gewirtz (eds.), *Morality, Moral Behavior, and Moral Development*. New York: Wiley.
- ROKEACH, Milton (1973) *The Nature of Human Values*. New York: Free Press.
- ROTH, Jeffrey A., John T. SCHOLZ, and Ann Dryden WITTE (1989) *Taxpayer Compliance: An Agenda for Research*. Philadelphia: University of Pennsylvania Press.
- SCHOLZ, John T., Kathleen M. McGRAW, and Marco STEENBERGEN (1991) "Legal Change and Social Context: Citizen Adaptation to the 1986 Tax Reform Act." Paper submitted for publication. (Unpublished, SUNY at Stony Brook).
- SCHWARTZ, Richard D., and Sonya ORLEANS (1967) "On Legal Sanctions," 34 *University of Chicago Law Review* 274.
- SCHWARTZ, Shalom (1977) "Normative Influences on Altruism," in L. Berkowitz (ed.), 10 *Advances in Experimental Social Psychology*. New York: Academic Press.
- SCHWARTZ, Shalom, and Netta INBAR-SABAN (1988) "Value Self-Confrontation as a Method to Aid in Weight Loss," 54 *Journal of Personality and Social Psychology* 396.
- SCOTT, Wilbur J., and Harold G. GRASMICK (1981) "Deterrence and Income Tax Cheating: Testing Interaction Hypotheses in Utilitarian Theories," 17 *Journal of Applied Behavioral Science* 395.
- SEARS, David O., and Jack CITRIN (1982) *Tax Revolt: Something for Nothing in California*. Cambridge MA: Harvard University Press.
- SMITH, Kent W., and Karyl A. KINSEY (1987) "Understanding Taxpaying Behavior: A Conceptual Framework with Implications for Research," 21 *Law & Society Review* 639.
- THURMAN, Quint C., Craig ST. JOHN, and Lisa RIGGS (1984) "Neutralization and Tax Evasion: How Effective Would a Moral Appeal Be in Improving Compliance to Tax Laws?" 6 *Law & Policy* 309.
- TYLER, Tom R. (1990) *Why People Obey the Law*. New Haven, CT: Yale University Press.
- TYLER, Tom R., and Paul J. LAVRAKAS (1985) "Cognition Leading to Personal and Political Behaviors: The Case of Crime," in S. Kraus and R. M. Perloff (eds.), *Mass Media and Political Thought: An Information-processing Approach*. Beverly Hills, CA: Sage Publications.
- TYLER, Tom R., Kenneth RASINSKI, and Eugene GRIFFIN (1986) "Alternative Images of the Citizen: Implications for Public Policy," 41 *American Psychologist* 970.