AN EXPERIMENTAL STUDY OF THE IMPACT OF RULE COMPLEXITY ON RULE COMPLIANCE

By

DOUGLAS MARK LAUFER

Bachelor of Science
University of Denver
Denver, Colorado
1974

Master of Taxation
Colorado State University
Fort Collins, Colorado
1980

Submitted to the Faculty of the Graduate College
of Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
Doctor of Philosophy
July 1985

AN EXPERIMENTAL STUDY OF THE IMPACT OF RULE COMPLEXITY ON RULE COMPLIANCE

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

I would like to express my appreciation to the members of my committee, Dr. Milton F. Usry, Dr. P. Larry Claypool, Dr. Orley M. Amos, and especially Dr. Lawrence H. Hammer. Their insights and expertise made this study possible. I also appreciate the assistance received from other faculty at Oklahoma State University.

I am also indebted to my fellow doctoral students. The support and friendship of Rick Crosser, Shirley Dale, Curt DeBerg, Joel Hausler, and Ed Scribner is appreciated. A special debt is owed to Larry Watkins whose guidance and alliance were of immense value.

Additionally, I would like to thank my running partners, Jerry Crockett, Dick Eggerman, and Rex Finnegan who helped me keep the proper perspective of the program and enhanced my appreciation and enjoyment of Oklahoma.

Most of all I wish to thank my family, Amanda, Dustin, and my wife, Pam. Their support and understanding were essential to undertaking and completing the doctoral program. Additionally, Pam's administrative services were invaluable and made it possible for me to receive this degree without ever learning to spell.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE RESEARCH PROBLEM

Introduction

The purpose of this chapter is to develop an overview of the tax compliance problem, provide a description of the purpose of the research, and to describe the organization of the remainder of the dissertation.

United States Tax Law--An Overview

The current United States tax structure is directly traceable to 1913 and passage of the 16th Amendment. This amendment made constitutional a tax on incomes--a form of tax in dispute up until that time.

The 16th Amendment to the U.S. Constitution reads,

> The Congress shall have power to lay and collect taxes on income, from whatever source derived, without apportionment among the several states and without regard to any census or enumerations.

The intent of the amendment was simply to allow the government to assess an income tax for revenue raising purposes. However, this simple beginning has evolved into a complex maze of rules and procedures whose components include the Internal Revenue Code, Tax Regulations, Revenue Rulings, and common law provided by the courts in interpreting the tax laws.

Until 1939, the income tax laws were relatively simple and taxes were levied on just 6 percent of the population [Goode, 1964]. Raising revenue to operate the government appears to be the primary objective of the tax law during this time period. During the time period 1939 to 1954, the scope of the tax law expanded rapidly. By 1945, 74 percent of the population had to pay federal income tax [Goode, 1964]. Many changes in the law were an outgrowth of equity considerations because the increasing complexity created a need for more rules and clarification of said rules [Sommerfeld, Anderson, and Brock, 1982]. Finally, in 1954, Congress passed the Internal Revenue Code of 1954 which was a comprehensive revision and reorganization of the 1939 Revenue Code. The recodification was an effort to delete obsolete material and make the Code more understandable [U.S. Congress, House, H. Rept. 1337]. All revenue measures which changed the tax law since 1954 have been incorporated into that codification.

It appears that the Internal Revenue Code revisions made during the period between 1954 and 1969 were primarily designed to achieve social and economic objectives. Congress utilized the tax law to manipulate the private sector by offering incentives to taxpayers to engage in activities deemed good for the public. For example, the investment tax credit (ITC) was implemented to encourage private sector purchases of Section 38 property (basically tangible personal property used in a trade or business or for the production of income). After initial introduction in 1962, the ITC provisions were frequently modified as Congress continued to manipulate investment spending and saving activity [Sommerfeld, Anderson, and Brock, 1982].

Since 1969, public concern about abusive tax avoidance schemes has grown [Sommerfeld, Anderson, and Brock, 1982]. Legislative emphasis has been directed at tax reform, tax equity, and fiscal responsibility. Many tax shelters have been eliminated. The Internal Revenue Service (IRS) has been granted more extensive power to enforce the law, and taxpayer and preparer penalties have been greatly increased. The 1976 Tax Reform Act introduced several statutory provisions regulating the conduct of tax return preparers. Additional sanctions have been implemented since 1976 [Laufer, 1980]. A shift in emphasis is apparent considering the fact that prior to 1976 the Internal Revenue Code provided few provisions which affected the conduct of tax return preparers. Although new tax legislation is often directed by social and economic objectives (as evidenced by the 1981 Economic Recovery Tax Act which was predominantely an economic incentive tax bill), it appears that at present there is an overriding concern with assuring compliance with the existing tax laws. This is evidenced by the "compliance gap" hearing held by the subcommittee on oversight of the Internal Revenue Service of the Senate Finance Committee [U.S. Congress, Senate Hearings, 1982].

Tax Law Compliance

The Internal revenue laws in the United States impose a tax on the income of individuals and various organizations. The taxes are levied and collected under what is predominantly viewed as a self-assessment system. Additional taxes may not be assessed by the government until numerous costly legal procedures have been taken. The economic efficiency of the U.S. taxing system is dependent on a high level of

voluntary compliance. If voluntary compliance declines, the government must increase its audit, assessment, and collection activities to prevent a decline in tax revenues. This, in turn, would result in an increase in the cost of administering the tax law, which means that the share of the tax revenues available for other government functions would decline.

The widespread belief in the country today is that there is a decline in voluntary compliance [U.S. Congress, Senate Hearings, 1982]. Verification of this belief is a diffucult task due, in part, to divergence of opinion as to what exactly constitutes noncompliance and how it is best measured. If the assumption is made that for each taxpayer there exists a "correct" tax liability, then noncompliance would be the reporting of a tax liability in some amount other than the correct one. The cause of the discrepancy could be due to misstatement or omission of items of income or expense or an error in the calculation of the liability, given the correct income and expense data. Such misstatements could be intentional or unintentional.

The working definition of noncompliance utilized by the IRS under the Taxpayer Compliance Measurement Program (TCMP) is basically as detailed above. That is, noncompliance is the difference between the reported and correct tax liability. The objective of the TCMP is to provide information on taxpayer compliance characteristics for the purposes of optimizing enforcement activity. Based on information gathered through the TCMP, an estimate of taxpayer compliance is made. This measurement, called the voluntary compliance level (VCL), provides a percentage of the proportion of total tax liability voluntarily

reported by the taxpayer to total correct tax liability as determined by the government, such that:

$$V_i = \frac{X_i}{Y_i} \times 100 = \frac{X_i}{X_i + (Y_i - X_i)} \times 100$$

where: V_i = Voluntary compliance level

X_i = Tax liability reported by taxpayer

Y_i = Total correct tax liability

Note that it is possible for X_i to be greater than Y_i. This would cause V_i to be greater than 100 percent. The noncompliance in such a case would result in an overreporting of tax liability.

Numerous factors are viewed as contributing to noncompliance, not the least of which is tax law complexity:

> The precise reasons for the decline in voluntary compliance cannot be easily identified; however, a number of factors may contribute to the problem.. For example, the complexity of the tax code and frequent changes in its provisions may contribute to higher levels of taxpayer misunderstanding than existed in earlier times. This higher level of misunderstanding would lead to an increase in inadvertent noncompliance. [U.S. Congress, Senate Hearings, 1982, p. 39]

In an attempt to deal with this "inadvertent" noncompliance, there have been numerous calls for simplification. The legislation which dealt most directly with this problem was the Tax Reduction and Simplification Act of 1977 [U.S. Congress, Public Law, 1977]. Committee reports to the bill reveal that a primary purpose behind the act was simplification of the individual income tax [U.S. Congress, S. Rept. 66]. To the extent that the number of computations required by an individual can be reduced, the tax law will be more simplified [U.S. Congress, S. Rept. 66]. Accordingly, in theory, compliance should be enhanced through simplification.

Any change in the tax law which enhances compliance would improve the economic efficiency of the tax system by reducing the relative cost of tax assessment and would promote equity within the tax system, or at least the perception of equity. However, changes in the structure of the tax system are costly to both the public and private sector.

Within the public sector the costs are manifested in the form of direct monetary expenses for printing of items such as hearing proceedings, committee reports, House and Senate bills, Public Laws, revisions to the tax code and regulations. There is also a significant opportunity cost. Members of the Congress and their staff must spend time researching the social and economic ramifications of proposed tax law changes. If tax bills become law, then the Treasury Department and IRS will incur costs for updating or promulgating new regulations and procedures and retraining personnel. There is also the potential for an increased case load on the tax courts if the new law is not interpreted in the same way by the government and taxpayer. Accordingly, unnecessary changes will interfere with more efficient use of public funds. If unnecessary proposals can be avoided, Congress can spend more time on other pertinent issues. The IRS can direct more resources to enforcement or other activities and, hopefully, additional tax court overload can be avoided.

The cost in the private sector may be even more significant. Tax law changes create the need to use resources in the private sector in a manner similar to that of the Treasury Department and IRS. Tax practitioners must research the new law, retrain staff, prepare relevent publications, and communicate and review the implications of the change with their clients. Litigation with regard to disputes with the IRS

over the interpretation of a new tax law may divert private sector economic resources from more productive use.

On a broad scale, tax laws appear to have an impact on financial decisions and resource allocation. A change in the tax system would create temporary economic disequilibrium and, if ineffective, would waste private sector resources. Accordingly, before changes are made to simplify the tax law for the purpose of increasing taxpayer compliance, a more rigorous examination of the relationship between complexity and noncompliance should be undertaken. If complexity has a significant negative impact on noncompliance, then policy makers should explicitly examine the level of complexity each tax provision creates and make adjustments when necessary to simplify the law. If complexity is not a significant factor in noncompliance, tax simplification should not be attempted or it must be justified on some other grounds.

Purpose

Congress, the IRS, and the media all attribute a portion of taxpayer noncompliance to the level of complexity in the tax law. The purpose of this study was to determine empirically if the level of complexity in a self-assessment system has a significant impact on the level of compliance within such a system. This study distinguishes between two forms of complexity in an effort to provide clear empirical evidence as to whether or not a relationship exists between complexity and level of noncompliance. A distinction was made between complexity in the form of more detailed and numerous computations (computational complexity) versus complexity in the ambiguity or subjectivity of rules

(rule complexity). The study examined the impact on noncompliance of both forms of complexity.

Due to regulations protecting the IRS and the privacy of taxpayers, actual tax records were not available for this study. Therefore, the present study was conducted as a laboratory experiment utilizing students as surrogates for taxpayers and self-grading of examinations as a task surrogate for tax return filing.

Organization of Remaining Chapters

Chapter II provides a review of the relevant tax compliance and behavioral experimentation literature. Chapter III details the methodology employed in the study. Chapter IV offers an analysis of the results of the study. Chapter V contains a summary and conclusions of the study.

CHAPTER II

LITERATURE REVIEW

Introduction

In this chapter, the literature pertinent to the present study is reviewed. An overview of the tax literature is presented and then the behavioral literature is reviewed.

Tax Literature

There has been little published research in the tax compliance area. The research which has been published has investigated tax evasion from a theoretical and empirical perspective. The theoretical studies have utilized microeconomic models for the most part (for example, see [Allingham and Sandmo, 1972]; or [Fishburn, 1981]). The empirical studies have utilized different methodologies in attempting to measure the impact of various factors on the tax evasion decision. Friedland, Maital, and Rutenberg [1978] used a simulation approach to determine the sensitivity of tax evasion to changes in variables such as tax rate, magnitude of fines, audit frequency, and socio-economic variables. The experiment provided subjects with a "salary" and required that the subjects report the amount of salary earned and the corresponding tax liability. The subject's objective was to maximize net income (salary minus tax liability). Monetary rewards were

distributed proportionately based on net income at game's end. There were four rounds with ten salary payments per round. The tax rate, audit frequency, and penalty level variables were preannounced and varied over the pay periods. The researchers calculated the frequency and magnitude of the difference between reported income and earned income. Zero order and multiple regression correlation coefficients were calculated so to compare the relative impact of the variables. Based on the results of the simulation, the rate of tax was found to be the most important determinant of the probability of evasion. Additionally, large fines were found to be more effective deterrents than frequent audits. These conclusions were drawn based on the size of the beta coefficient.

Mork [1975] utilized data obtained from the Norwegian Occupational Life History study in an effort to assess the impact of level of income on tax evasion. The life study questionnaire included a request that the respondents state their income. Based on the response to the questionnaire, respondents were divided into income classifications. The class intervals were N.Kr. 3000. For the same respondents income data was taken from their tax declarations. The researcher assumed that the income stated on the questionnaire by the respondents was their true income and that by a comparison of the data inferences could be drawn with regard to tax evasion. The average income reported to the tax authorities was determined for each income classification and compared to the interval midpoint for each class. (The researcher was not able to calculate the average income by classification based on questionnaire response and, therefore, used interval midpoint as a proxy.) The income reported as a percent of true or stated income was calculated

for each class interval. There was a steady decline in the percent of income reported to true income as level of income increased. The author felt the study provided "soft" evidence of the existence of tax evasion and that tax evasion may increase as income levels increase. The study did not provide any statistical analysis beyond the percent relationships. The author stated, "interpretation of our empirical result is quite complex and should be left to the reader" [Mork, 1975, p. 74]. Based on this reader's interpretation, the empirical results appeared inconclusive at best. Deficiencies in the design, data collection and assumptions make rigorous statistical analysis of the data a fruitless endeavor. For example, accepting a questionnaire response to the question of income level as true income and comparing this to income as determined within the guidelines of the tax law appears to be an invalid comparison.

A macro-level empirical study was conducted by Clotfelter [1983]. This study involved analysis of the relationship between tax evasion and marginal tax rates and utilized aggregate tax data. A model for tax evasion was developed based on the difference between reported income and the amount of income that the IRS examiners determined to be due. The model has the following form:

$$X_i = Y_i - V_i - U_i$$

where

X_i = Individuals' reported income

Y_i = Individuals' true taxable income

V_i = Measure of deliberate evasion

U_i = Error term reflecting mistakes in the calculation of taxable income

This equation can be rearranged so that the difference between true and reported income is the sum of evasion and error: $Y_i - X_i = V_i + U_i$. The difficulty with this model is that it is impossible to identify individual tax evasion. However, if mistakes were random, then the U_i's will have an expected mean value of zero, and, it would be possible to measure the evasion component within the population. Aggregate data reported by the IRS makes it clear that there is a significant tendency toward underreporting of taxable income [IRS, 1969]. By application of econometric models to aggregate TCMP data, the author found that this tendency was sensitive to marginal tax rates. Econometric estimates of the elasticity of underreported income with respect to the marginal tax rates were positive and generally significant. The author also found that the presence of several different tax forms was associated with increased underreporting among non-business returns. However, there was no significant effect for business returns. The problem with this study is that the variable V_i may be measuring not only deliberate evasion but also variances due to the ambiguity or subjectivity of tax laws which allow advantageous interpretation of the law by the taxpayer.

The only other aspect of the compliance problem which has been the subject of rigorous research is the measurement of unreported income, from legal and illegal sources, at the macro-level. Studies in this area include: Tanzi [1982], Gutman [1977], and IRS [1979]. The procedure utilized in most of these studies has been to estimate the magnitude of tax evasion by the currency-based "cash footprint" method. Tanzi utilized an econometric model which explained the ratio of currency holdings to money supply as a function of interest rates,

share of wages and salaries in personal income, per capita income, and average tax rates. The ratio was viewed as an indicator of unreported transactions. The underlying rationale is that the demand for money is roughly proportional to the flow of total economic activity. By statistically controlling the variables and estimating the velocity of underground money a regression model was employed to estimate the volume of unreported taxable income corresponding to the estimated excess currency. A positive correlation between tax rates and currency holdings was found which the author interpreted as support for the theory that evasion rises with marginal tax rates.

Gutman [1979] examined the change in the ratio of currency in circulation to demand deposits and attributed the increase in the ratio over the last 35 years to increased underground activity. Gutman reasoned that the incentive to engage in off-the-book work, which allows tax evasion, has increased over the past 40 years; and since currency is widely used to facilitate this activity, the increase in currency ratio must be attributed to increased underground activity. For 1976, Gutman's estimate for underground GNP is $176 billion. There are some serious flaws with currency-based estimates. The currency equations implicitly assume all money demand comes from domestic sources, but there is a strong possibility that a large proportion of currency in circulation outside the banking system is held by foreign investors [Henry, 1983]. Additionally, the use of economic time series analysis is statistically flawed by the high degree of correlation between variables. There is no strong underlying theory to support the relationship between tax rates, noncompliance, and currency demand which are the basis for Tanzi's contentions. Gutman's model, which is

more simplistic than Tanzi's, does not allow control for variables such as interest rates, price levels or per capita income. Additionally, the Gutman model assumes the money velocity of the underground economy is equal to that of the legal economy. Evidence exists to the contrary [Henry, 1983]. In summary, accurate measurements of the money supply are difficult to obtain and seem to pick up much activity which has little to do with noncompliance [Henry, 1983].

The IRS has the only truly direct measurement technique, the previously mentioned TCMP. In a 1979 study the IRS estimated the amount of unreported income (from legal and illegal sources) for 1976 was between $100 and $135 billion. The service does not disclose the details of the procedure utilized to arrive at this estimate; consequently, the reasonableness of this estimate is not determinable.

There are numerous articles, papers, and reports of a heuristic or intuitive nature which present an argument for a complexity-compliance relationship. The IRS Commissioner, Roscoe Egger, stated that he believes that simplifying the tax code would be the most effective way to enhance compliance [Wall Street Journal, 1984]. In 1985 the Congress, in reaction to the complexity problem, is going to further study simplification of the tax code and will seriously examine several bills which address comprehensive income tax reform. The tax reform proposals include the Treasury Department Report [U.S. Treasury, 1984] and the Bradley-Gephart Fair Tax Plan [U.S. Congress, S. Rept. 1421]. In general, these reform proposals seek to greatly simplify the tax law by broadening tax bases, lowering tax rates, and eliminating numerous tax preferences.

Many of the current concepts for reform are derived from an extensive study of tax reform conducted by the Treasury Department in 1977 [U.S. Treasury, 1977]. The private sector has also been active in analyzing the noncompliance issue. It is often an agenda topic at the tax conferences of professional organizations, as evidenced by the 1983 American Bar Association Invitational Conferences on Income Tax Compliance. The lead paper at this conference [Henry, 1983] reviewed the various definitions and estimates of noncompliance, pointing out that there is no consensus on cause or amount of tax revenues lost due to noncompliance. Henry concludes that the real growth in noncompliance is uncertain. He recommends nonvoluntary assessment and collection mechanisms (i.e., withholding at the source) as well as indirect taxes as potential solutions. Additionally, Henry feels continued attempts to produce noncompliance estimates and make changes which seem to promote increased compliance are desireable.

From the period 1967 through 1977 there have been only five doctoral dissertations which address tax compliance and just two tax research endeavors utilizing the laboratory experiment methodology [Brighton, Michaelsen, and Willis, 1978]. Only one of these doctoral dissertations has relevance to the compliance issue [Marquardt, 1975]. In addition, a recent doctoral dissertation endeavored to study complexity within an experimental setting [Raabe, 1980]. These two works are discussed below.

Raabe [1980] attempted to determine the feasibility of an "elective filing system." Raabe theorized that a degree of complexity could be eliminated through an elective filing system which would greatly reduce the taxpayers' compliance efforts and costs. Under this system

the IRS would use "its massive data processing capabilities to coordinate the various information returns that it receives annually. The IRS could prepare a tentative computation of each taxpayer's adjusted gross income, taxable income, and tax liability. Upon receipt of this tentative return, the individual would decide whether to accept its determination or to override the proposed return." The study was conducted as a laboratory experiment using both student and nonstudent subjects. The subjects were provided with a set of financial and related information for a hypothetical taxpayer and computer-printed tentative tax returns from the IRS (hypothetically). The subjects' task was to determine if the tentative tax return was complete and correct based on their information set and to submit changes if necessary. Subjects' responses were statistically analyzed to determine if the completeness of the tentative returns affected compliance. The results of the experiment indicated that subjects whose tentative returns overstated taxable income were more likely to submit revisions than those subjects whose tentative returns understated taxable income. It appears that the results do not necessarily indicate that an elective system would enhance compliance.

For the other relevant dissertation, [Marquardt, 1975], the researcher interned with the National Office of the IRS which provided him with the opportunity to obtain data not typically available to other researchers. Eight attributes potentially affecting voluntary taxpayer compliance were selected. These attributes were: complexity, tax form difficulty, inadequacy of the instructions, level of judgment necessary, effort required to verify information, math computations required, underlying records needed, and changes in law. Each

attribute was divided into three categories for the purpose of ordinal measurement. The attributes and the categories were both subjectively determined by the researcher. Some attributes were not precisely defined. Marquardt utilized the Goodman-Kruskal index of order-association to measure the relationship between compliance and the various attributes. The statistical method was applied to the available data base information. The research findings indicated the existence of a relationship between five of the attributes and compliance for seven classes of tax return preparers. The five attributes were ranked as to relative importance; however, no data were provided with regard to the impact on compliance of varying the requirements or level of difficulty within each attribute.

Behavioral Literature

There is an extensive body of compliance theory within the psychology and social psychology literature. The compliance problem has been examined from several perspectives. Studies have explored the behavior aspects of task performance. In many situations the task performance variable is synonymous with, or at least similar in nature to, compliance. The subject's task performance is measured against some norm or standard. The independent variable task level represents a form of complexity but does not capture the essence of the complexity variables utilized in the present study. The tasks were typically manual in nature and related to job activities. A majority of the research falls within the performance analysis paradigm. Within this paradigm a widely utilized model for performance analysis is the multiplying model: Performance = Motivation x Ability [Heider, 1958].

Using this basic framework, research efforts have addressed such questions as the impact of goal setting or other incentive systems on task performance with conflicting conclusions regarding the association between level of performance and incentives. Within these studies goals represented a budgeted level of output. The incentives were either a piecemeal rate or bonus system. Findings include: high goals contribute to high levels of performance [Locke, 1966], high goals in the absence of feedback have little effect on task performance [Locke, 1968], and goals set too high have negative effect on performance [Forward and Zander, 1971]. The intent of the research was to evaluate how best to promote compliance with an assigned or budgeted standard (i.e., encourage performance which complys with the norm). The incentive schemes utilized to motivate compliance do not appear to lend themselves to application in the taxpaying economy.

Another fertile ground for compliance research is the effect on compliance of task difficulty and task ordering. This appears to be directly related to the present study; however, investigation of the literature revealed an orientation towards specific job applications. Within this paradigm, there has been extensive study of the effectiveness of the two procedures for increasing compliance. These procedures are the foot-in-the-door (foot) and the door-in-the-face (face) procedures. In the "foot" technique individuals are asked to first comply with simple tasks followed by progressively harder tasks. The "face" approach is just the opposite, first seeking compliance with extremely hard tasks followed by less difficult tasks, with the latter tasks typically being the ones with which actual compliance is desired. Empirical evidence supports both techniques as effectively increasing

compliance [Freedman and Fraser, 1966], [Cialdini and Ascani, 1976]. The most effective approach of the two was an initial easy task followed by a hard task [Goldman, Gier, and Smith, 1981]; however, the success of either procedure is conditional upon the level of difficulty of the tasks [Goldman, Gier, and Smith, 1981]. This last finding appears to support the generally held intuitive belief that significant increases in complexity will result in declining compliance. The research does imply that the learning process occurs faster when the initial tasks are simple and level of difficulty is progressively increased. It is difficult to ascertain whether this phenomenon can be utilized to enhance tax compliance. It does not appear feasible to educate taxpayers by beginning with simple tasks and proceeding to add complexity step-by-step.

Another aspect of compliance theory research explores the possible effects of negative sanctions (penalties, etc.) on producing compliance with rules or norms. The research within this paradigm has generated different, often conflicting, results. Early studies by Ball [1955] suggested punishment was of minor importance as a behavioral influence. Other research concluded that an increased probability of imprisonment is correlated with lower crime rates [Tittle, 1969]. Debate and research continues with regard to sanctions, although partial reconciliation has been achieved via the wide acceptance of the deterrence doctrine [Erickson, Gibbs, and Jensen, 1977]. Simply stated, this doctrine provides that the probability of getting caught is the most important attribute in deterring noncompliance behavior. With regard to the tax economy, this would seem to indicate that the perceived probability of being audited and having additional taxes and penalties

assessed is of greater importance than the penalty structure. Unfortunately, lack of additional resources greatly curtail the ability of the IRS to significantly expand audit activity [IRS, 1979].

Summary

This chapter presented an overview of prior research in tax compliance. Rigorous empirical investigation of the inadvertent noncompliance issue has been limited. The few tax studies that have been conducted did not conclusively establish the link between complexity and compliance. The behavioral studies suggest that such a link exists in a repetitive learning environment, but it is not clear from such research that such a link exists in non-learning environments such as the taxpaying economy. The current study seeks to bridge this gap by conducting a laboratory experiment designed to determine whether the level of compliance is a function of the level of complexity in a non-learning environment.

CHAPTER III

METHODOLOGY

Introduction

In this chapter the experimental design and the specific structure of the experiment are explained. Additionally, the statistical model utilized to analyze the data is presented.

The Experiment

A laboratory experiment was utilized to allow manipulation of the independent variables while facilitating control of the extraneous variables. The subjects were 111 college students enrolled in three different sections of Principles of Accounting. The experimental task was designed to simulate individual taxpayer activities. The subjects were provided materials which allowed for self-assessed examination grading on each of four different examinations administered during the semester. The subjects' task was to determine their examination score and communicate the information on the appropriate form to the experimenter. The subjects' grades in the course were based solely on examination scores.

The hallmark of laboratory experimentation is abstraction, i.e., the deliberate manipulation of one or more crucial variables, the deliberate control of many others and the precise measurement of one or

more variables of interest [Swieringa and Weick, 1982]. To produce a desired phenomena in a laboratory experiment, a variety of procedures which enable the researcher to manipulate, control and measure variables are required. Numerous articles and books have detailed the factors which can jeopardize validity of an experiment and how the problems may be avoided. Campbell and Stanley [1963] discuss eight extraneous variables which pose a threat to internal validity if not controlled in the experimental design. These effects are: History--events occurring during the experiment; Maturation--change in subjects due simply to passage of time (i.e., growing tired); Testing--effects of prior tests on subsequent tests; Instrumentation--potential changes in the measuring instrument may produce changes in measurements obtained; Selection--bias in selection of respondents; Experimental mortality--loss of respondents during experiment; Selection-maturation interaction--may be confounded with effect of the experimental variable [Campbell and Stanley, 1963, p. 5]. This experiment was patterned after the counter-balanced design discussed by Campbell and Stanley [1963, pp. 50-52]. The design is discussed more fully in the experimental design section of this chapter. Utilization of this experimental design was critical in assuring that the internal validity of the experiment was sound.

Unfortunately, external validity (generalizability) cannot be achieved with any certainty in a laboratory experiment [Cherulnik, 1983]. This is because it is not possible to capture every causal factor in the real-world environment in a laboratory setting [Cherulnik, 1983]. For example, in the income tax content it is not possible in a laboratory experiment to duplicate the phenomenon of an

individual taxpayer incurring daily financial transactions and then summarizing the activity, applying the relevant tax law to complete and file a tax return reporting an actual tax liability. It is also not possible to observe the outcome of this process in the taxpaying economy. Therefore, with regard to external validity, the attempt should be to highlight selected behavioral processes and certain conditions related to these processes [Swieringa and Weick, 1982].

Dickhaut, Livingstone, and Watson [1972], in the AAA Committee report on Behavioral Accounting Research, established the following argument in support of the existence of external validity in a behavioral experiment.

i) Experimental reality is a necessary (but not sufficient) condition for internal validity.

ii) Internal validity is a necessary (but not sufficient) condition for external validity

Since the factors which can jeopardize internal validity can be controlled in the experimental design, it is on the issue of reality that determination of fulfillment of both necessary conditions hinges. Experimental reality exists when subjects are "involved" in the experiment [Aronson and Carlsmith, 1968]. It is not the lack of involvement per se which threatens experimental realism but rather the lack of a particular type of involvement. For example, the use of a mental task in the laboratory to simulate a physical task in the real world can introduce an unintended high level of subject interest [Birberg and Nath, 1968].

Swieringa and Weick [1982] state that experimental realism (whether laboratory events are taken seriously) is more important than

mundane realism (whether laboratory events are similar to real world events) because the focus is more on the effects of stimuli and less on the effects on the setting in which the stimuli is typically encountered. There appears to be misplaced concern about mundane realism on the part of accounting experimenters. Swieringa and Weick state that if laboratory experiments have experimental realism, there appears to be little, if any, need to strive for mundane realism. If the experimental situation is meaningful to the subjects, then the impact on and the credibility for the subjects should be such that a behavior similar to that found in the real-world setting is evoked and the experimental procedures should capture the variable of interest. Mundane realism may make it more difficult to learn from the experiment [Swieringa and Weick, 1982, p. 34]. An experimental setting within a context which seeks to maintain mundane realism seeks to recreate a real-world situation within an experimental framework. However, the subjects typically have little incentive to react and respond to the stimuli as they would in the actual situation. Additionally, since the subjects may face the situation in their daily activities, they may override the experimental manipulation by utilizing skills and routines they already possess. This problem is evident within the behavioral research in audit decision making [Joyce and Libby, 1982]. Within this paradigm it is not unusual to create a fact pack, with regard to a particular audit decision, which replicates the facts as they might be found in an actual audit setting. The subjects, who often are practitioners, in responding to the question of how they would handle the particular situation probably will not have an incentive to be certain the answer is in fact in agreement with what would actually occur. The possibility exists

that an incentive may be present to respond differently than they would in the actual situation. Also, the subjects' backgrounds in the area may cause them to utilize their preexisting knowledge in lieu of specific experimental information.

There were three separate surrogate utilizations in the present study:

1) Determination and reporting of examination score as task surrogate for determination and reporting of tax liability,

2) Students as subject surrogate for taxpayers, and

3) Classroom setting as surrogate for the "real world".

Utilization of generic tasks in behavioral studies can enhance experimental realism [Aronson and Carlsmith, 1968]. The surrogate task utilized in the present study is generic in nature, (i.e., subjects were asked to grade and report their own examinations rather than compute and report their own taxable income). An attempt was made to achieve a high level of experimental realism without building mundane realism into the experiments. The surrogate task utilized was designed to achieve a high level of correlation between the subjects' task involvement in the laboratory setting and the real world phenomenon. The effort to accomplish this is evidenced by the similar nature of the task and the instruments employed to complete the task. To grade the exams the subjects were required to read and interpret instructions, perform mathematical computations, and report results on the appropriate forms. This procedure is very similar to that facing taxpayers in the tax economy.

There has been extensive research to investigate the implications of utilization of students as subject surrogates. Such research has

been conducted within the accounting discipline, as well as in other social science areas (see, for example, Abdel-Khalik [1974] and Ashton and Kramer [1980] or for research outside the accounting discipline see Zelditch and Evans [1962]). There is no consensus about the impact of using student surrogates in an experiment; however, explicit consideration of the impact of using students must be made on each experimental factor. The design of the present experiment matched the students' incentives with those of taxpayers, i.e., both have something of value at risk. Since both have something at risk, the probability of similar involvement by students in determining their examination score with that of taxpayers in determining their tax liability is increased.

It is doubtful that a laboratory or classroom can simulate real-world conditions. However, Zelditch [1969] argues that the experimenter only need create those aspects relevant to some theory. There is no need to attempt to create a completely "real" instance in the laboratory. To the extent that subjects are functioning in an environment which necessitates self-imposed compliance with instructions and rules that involve varying levels of complexity, it would seem the relevant aspects of the taxpaying economy have been replicated.

The notion that degree of complexity has an impact upon level of compliance appears to be accepted within the Congress and Treasury Department as if it were a positive model. A result of this has been a call for simplification of the individual income tax. The IRS has defined noncompliance to be any variation between true tax liability and reported tax liability. The Service views the degree of complexity to be a function of the number of computations made by the taxpayer and

the intricacy of the tax forms which the taxpayer must utilize [U.S. Congress, Simplification, 1977].

In analyzing the impact of complexity on compliance, it appears there has been a failure to separate rule complexity from computational complexity despite the fact that they represent two different phenomena. The present study was designed to measure the impact on compliance of both forms of complexity. To evaluate the effects of rule complexity on compliance the rules utilized in grading exams were manipulated. Some subjects received very precise instructions about how to determine points earned on a problem, whereas others received more general guidelines which required subjective interpretation. To capture the effects of computational complexity on compliance, some subjects were required to perform only a few, straightforward computations and to record their exam score on a simple form, while others were faced with more difficult and numerous calculations which were to be reported on a more intricate set of forms. The following three null hypotheses were tested:

H_{o_1}: As complexity increases compliance will not be diminished.

H_{o_2}: As computational complexity increases compliance will not be diminished.

H_{o_3}: As rule complexity increases (becomes more subjective) compliance will not be diminished.

A fourth research hypothesis was tested to determine which form of complexity was of greater consequence in the study:

H_{o_4}: Compliance level for computational complexity treatments is equal to compliance level for rule complexity treatments.

In devising tax policy and procedures to address the noncompliance problem, it may be important to distinguish between the kind of complexity to be eliminated. Potentially, the costs of making a change could outweigh any benefits gained.

Theoretically, inadvertent noncompliance due to complexity should be a zero sum game. The direction of errors should be randomly distributed with a mean value of zero. That is, if the errors are truly inadvertent and random, the number of and amount of errors in the taxpayer's favor should be equal to those favoring the government. Other factors, most significantly the overt underreporting of tax liability for self-enrichment, obviously affects compliance. Any measurement of the impact of complexity on compliance is greatly impeded by the confounding effect of the other factors, such as cheating. The design of the present study provided an opportunity to observe compliance behavior directly with factors other than complexity held constant across the sample groups.

To the extent that effects other than complexity are eliminated and errors are strictly inadvertent, the above hypotheses would not be rejected because the mean value across the different levels of complexity would all approximate zero. However, the distribution of the errors should be increased as the level of complexity increases. Accordingly, the present study tested the following additional null hypotheses:

$H_{O_{1A}}$: As degree of complexity increases variability of differences will not change.

$H_{O_{2A}}$: As computational complexity increases variability of differences will not change.

$H_{O_{3A}}$: As rule complexity increases variability of differences will not change.

$H_{O_{4A}}$: Variability of differences due to computational complexity is equal to variability of differences due to rule complexity.

Administration of the Experiment

Four examinations were administered to each subject during one semester. Each examination was representive of activity for one tax year. Examinations were collected upon completion in the normal manner. At the class meeting following the examination each student received his/her examination, a set of grading instructions, and forms on which to report their examination score. (Appendix A provides an overview of the experimental procedures. Appendixes B, C, and D offer examples of the instruments). Grading instructions and forms were designed to capture differing levels of complexity. More precisely, subjects received one of four different grading packets: Simple/Objective, Simple/Subjective, Complex/Objective, or Complex/Subjective. Figure 1 presents a graphic illustration of the treatment matrix.

GRADING INSTRUCTIONS	COMPUTATIONS: SIMPLE	COMPLEX
SUBJECTIVE	SS	SC
OBJECTIVE	OS	OC

Figure 1. Treatment Matrix

Simple represents computational simplicity, minimal calculations, and basic reporting form [see Appendix C, p. 67]. Complex refers to computational complexity; that is, required calculations were more detailed and numerous and forms were more intricate than in the simple setting [see Appendix C, pp. 68-70]. The grading instructions represent rule complexity. Objective instructions were very exact in providing information on how to determine points earned for each exam question; while subjective instructions were more general, less specific, and to that extent, more ambiguous or complicated. Appendix B includes examples of objective and subjective treatments. The students were required to report (file) their examination score within a set period of time on the forms provided. Students turned in only the forms; they kept their examinations.

Students were assessed a penalty for late filing of reported examination scores. Additional enforcement procedures were utilized and applied uniformly across the sample. The procedures involved audit verification of examination scores and penalties for noncompliance. Audit implementation and selection techniques were designed to replicate many of the facets of the IRS process. The examinations were stratified into groups by score and then selected for audit randomly. The subjects received letters advising them that they had been selected for audit [see Appendix D, p. 75]. A few of the audits required review of the entire examination. The results of these audits were used to determine areas of emphasis in future audits. This approach appears to replicate the IRS TCMP process. For the remaining audits, only selected sections of the examinations were reviewed. This is similar to the situation faced by most taxpayers selected for audit

investigation. The penalty structure was provided to the students in detail as a supplement to the course syllabus [see Appendix D, pp. 72-74].

For purposes of compliance measurement, copies of all examinations were made. Every examination was graded by the instructor to determine the correct score which was compared to the reported score. This phase of the experiment was not disclosed to the subjects until the experiment was completed. Additionally, a select sample of examination pages was copied for utilization in the audit verification phase of the experiment.

Subjects' Reward Structure

Laboratory experiments are often criticized for failure to provide incentives that adequately motivate subjects. In the present study an attempt was made to overcome this perceived deficiency through the reward and penalty structure utilized. All subjects were enrolled students in principles of accounting. In order to receive credit for the course, the students had to earn a passing grade. The students' grades in the course were based exclusively on performance on five examinations. Four of the examinations were self-graded. The students were required to report an examination score to get any credit at all for the examination. They faced the possibility of audit and, if they reported an incorrect score, loss of examination points. This reward and penalty structure in essence captures the attributes of the self-assessing system of individual income taxation utilized in the United States.

Postexperiment Questionnaire

A questionnaire was utilized to determine the perceived salience of various experimental factors to the subjects. Each subject completed the questionnaire after the experiment was completed and before the aspects of the experimental procedures were explained. The questionnaire consisted of sixteen structured (closed) questions with Likert attitude scale responses. The subjects also had the opportunity to offer general comments. The responses were anonymous. For analysis the scale was quantified as follows: definitely yes, 5; yes, 4; not applicable, 3; no, 2; definitely no, 1. The questionnaire and the mean response and standard deviation for each item are detailed in Appendix E (pp. 77-79).

Experimental Design

The experimental effect of interest was the impact of the experimental treatment (complexity) on compliance. However, other possible experimental effects could occur due to group assemblage and the sequencing of treatments. These could pose a threat to internal validity of the experiment. To avoid these problems, the students were assigned randomly to one of four groups and all students were subjected to all treatments. This design is referred to as a counterbalanced or rotating design [Underwood, 1949]. Within this framework, a Latin-square design was utilized. Figure 2 illustrates the design graphically.

	Exam 1	Exam 2	Exam 3	Exam 4
Group A	X_1O	X_2O	X_3O	X_4O
Group B	X_2O	X_4O	X_1O	X_3O
Group C	X_3O	X_1O	X_4O	X_2O
Group D	X_4O	X_3O	X_2O	X_1O

where:

X_1 = Simple Computations, Objective Instructions

X_2 = Simple Computations, Subjective Instructions

X_3 = Complex Computations, Objective Instructions

X_4 = Complex Computations, Subjective Instructions

O = Measurement of Noncompliance (Reported Score - Actual Score)

Figure 2. Latin-Square Counterbalanced Design

Utilization of this design made each classification orthogonal, thereby permitting between group comparisons while avoiding correlations and interactions arising from sequencing, maturation and practice effects. This made it possible to statistically analyze the main effect using F-tests from the ANOVA procedure.

The ANOVA

The analysis of variance (ANOVA) is a statistical procedure which is utilized in the analysis of experimental data. ANOVA provides a method for dividing the variation observed in experimental data into different parts. Using ANOVA it is possible to assess the relative magnitude of variation resulting from different sources and compare

the variation with expectations under the null hypothesis [Ferguson, 1976].

Distributional Assumptions

There are three distributional assumptions made when utilizing F-tests within an ANOVA framework to test hypotheses about means. One assumption is normality of the dependent variable distribution. Unless there is reason to suspect a fairly extreme departure from normality, it is probable that the conclusion drawn from the data using an F-test will not be seriously affected [Ferguson, 1976]. Although the effect of departure from normality is to make the results appear more significant, it is reasonable to accept this assumption without specific testing if the F-tests do not indicate significant differences in treatment means.

A second assumption is independence of observations. This assumption is more difficult to evaluate. However, in general, non-independence of observations increases the probability that treatment effects will be indicated for non-efficacious treatments. For an in-depth discussion of analysis of the normality and dependence assumptions see Glass, Peckham, and Sanders [1972].

The final assumption is homogeneity of variances; that is, variances in the population from which the samples are drawn are assumed to be equal (i.e., $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2$). There is a good deal of evidence that the ANOVA is virtually unaffected by violations of this assumption if the samples are the same size, which they were in the present study [Box, 1954]. However, in the present study the equivariance hypothesis is of interest in its own right. Accordingly, in

addition to testing the hypothesis concerning treatment effects the Levene test was utilized to test for homogeneity of variance.

The Levene Test

The Levene's test procedure involved performing an analysis of variance on the absolute deviations of differences from the treatment group means. The mean for each of the treatment groups is first calculated, then the difference between reported and actual exam score is subtracted from the calculated mean for each observation within the treatment group. The absolute values of these deviation differences was the data for computation of an ANOVA. This permitted the use of F-tests of the hypothesis that the mean absolute deviation from the mean is constant across treatments (i.e., the second set of hypotheses presented above). ANOVA procedures allow the treatment means to be analyzed for significant differences. The Levene test provides a method to test for significant differences in variability present for each treatment. Studies of the Levene test indicate that it is relatively insensitive to nonnormality of underlying distributions and at least as powerful as Bartlett's test [Dayton, 1970].

Summary

This chapter presented the design and procedures utilized in conducting the experiment. The statistical models utilized were also discussed. In conducting laboratory experiments external validity cannot be achieved with any certainty. Apparently, however, the presence of experimental realism does enhance the external validity of the experiment. In the present study the generic task utilized allows a

high degree of experimental realism to be achieved. The probability of subject involvement similar to that of taxpayers determining their tax liability enhances the generic task experimental design.

For the purpose of determining the impact of complexity on compliance, the counterbalanced Latin-square design was utilized. This design should enhance the experiment's internal validity and allow statistical analysis of the effect of the complexity treatment on compliance using F-test and ANOVA procedures. The Levene test procedure was also utilized to permit comparison of treatment group variances. In the absence of overt noncompliance, analysis of variability present in each treatment provides a method to compare the impact of differing levels of complexity on compliance.

CHAPTER IV

ANALYSIS OF DATA

Introduction

In this chapter, the results of the laboratory experiment are presented. The results of the experiment are analyzed and discussed. An interpretation of the results is provided.

Results of Analysis of Data from Experiment

The objective of the study was to determine if complexity plays a significant role in compliance with rules. Towards this end, an experiment was designed which allowed for manipulation of independent variables so that the experimental effect of the treatments could be measured. The dependent variable in the model was the difference between reported exam scores and actual exam scores (reported score - actual score = difference). The counterbalance research design for the model contained three independent variables: occasions (examinations), groups, and treatments (complexity level). The main effect independent variable of concern was the treatments. The four treatments are: simple computational and objective complexity; complex computational and objective rule complexity; simple computational and subjective rule complexity; and complex computational and subjective rule complexity. The results of the ANOVA for the Latin-square arrangement are presented in Table I.

TABLE I

ANALYSIS OF VARIANCE STATISTICS--
FULL MODEL

Source of Variation	Degrees of Freedom	F Ratio	Observed Significance Level
Complexity	3 , 434	.35	0.7928
Computation Complexity	1 , 440	0.74	0.3905
Rule Complexity	1 , 440	0.16	0.6901
Interaction	1 , 434	0.13	0.7151
Groups	3 , 434	1.83	0.1389
Examinations	3 , 434	2.18	0.0880

This arrangement keeps the examination and group main effects from contaminating the main effect of the complexity treatments. Based on data collected and procedures utilized, the null hypothesis of no treatment effect cannot be rejected because the F ratio of .35 has an observed significance level of .7928 with 3 and 434 degrees of freedom. The lack of significant differences between the complexity treatments indicates that the level of compliance was not affected by complexity. The presence of a strong compliance incentive may be a cause of this lack of treatment effect. The penalty and audit structure which was applied uniformly across all sample cells was intended to simulate many of the features of the penalty and audit procedures provided for in the

have been more influential and restrictive than that of the tax law and, therefore, compliance was enhanced. Several factors support this possibility. Responses to the post-experimental questionnaire indicate that 92% of the subjects expected to be audited at least once. Eighty-seven percent of the subjects stated that the possibility of audit and penalties increased their effort to comply. Additionally, 78% of the subjects felt that had the instructor graded the examinations, no difference in score would have occurred. Based on these responses, it appears that the students, on the average, expected to be audited. This, coupled with the penalty structure, suggests that the enforcement structure substantially eliminated overt noncompliance. It is possible that inherent differences between the three class sections could dilute the treatment effect. The mean differences between actual and reported score for the three sections were 0.206, -0.041 and -0.147. These differences are less than the treatment mean differences (see Table II, pg. 41), which were not significant. Accordingly, it does not appear that class sections had an impact on the treatment effect.

The group classification was fully controlled by assigning each person to a group independently and at random. Accordingly, this source of main effect and interaction was removed. The inability to reject a null of no effect due to groups at a 10% alpha level is predictable assuming that overt noncompliance was discouraged by the perceived level of audit and that inadvertent noncompliance is a zero-sum game.

It would not have been unusual for occasions (examinations) to produce an effect due to repeated testing, maturation, practice and cumulative carry-over. Therefore, it is not abnormal that this

variable, with an observed significance level of .088 was more significant than the treatment (complexity) and group variables.

The results of ANOVA designed to test the two forms of complexity are also presented in Table I. The purpose was to compare the impact of the two types of complexity on compliance. The treatment matrix, Chapter III, Figure 1, details the configuration of these complexity components. The null hypothesis of no treatment effect for computational complexity cannot be rejected because an observed significance level for differences is .3905 for an F ratio of .74 with 1 and 434 degrees of freedom. Similarly, the null hypothesis of no treatment effect for rule complexity cannot be rejected because the observed significance level for the difference is .6901 for an F ratio of .16 with 1 and 434 degrees of freedom. The computational complexity, rule complexity interaction also is not significant; the observed significance level is .7151 for an F ratio of .13 with 1 and 434 degrees of freedom.

These results provide weak support for the hypothesis that computational complexity is a greater cause of noncompliance than rule complexity. However, there does not appear to be sufficient support for the contention that either form of complexity has any significant effect on the overall level of compliance.

The analysis of variance presented above addresses the question of variation between treatment means. The means and standard deviations for the difference between reported and actual scores are reported in Table II.

TABLE II

MEAN AND STANDARD DEVIATIONS FOR THE DEPENDENT VARIABLE--SCORE DIFFERENCES-- BY TREATMENT

Treatment	Observation	Mean Difference	Standard Deviation
Simple, Subjective	111	0.6667	7.1723
Simple, Objective	111	0.5405	6.5001
Complex, Objective	111	-0.3153	8.1664
Complex, Subjective	111	-0.3423	11.1799
Total	444	0.0158	8.4277

The ANOVA does not capture the differences in the degree of variability within treatments. Inspection of the standard deviations listed in Table II above suggests that the variances may not be homogeneous. This observed effect is not measured with the ANOVAs performed and presented in Table I. To remedy this problem, the Levene test, a procedure which allows examination of the equi-variance hypothesis, was used to analyze the data. The results are presented in Table IV. The source of variation due to complexity is comparing the differences in variability between each of the four treatments: simple, subjective; simple, objective; complex, objective; complex, subjective. The source of variation due to computational complexity has two treatment levels. The simple, subjective and simple, objective are combined and the variability compared to that of the combined complex, objective and

complex, subjective. Similarly, rule complexity also has two treatments which are compared for differences in variability. For rule complexity the simple, subjective and complex, subjective treatments are combined and compared to that of the simple, objective and complex, objective treatments.

TABLE III

ANALYSIS OF VARIANCE STATISTICS--
LEVENE'S TEST

Source of Variation	Degrees of Freedom	F Ratio	Observed Significance Level
Complexity	3 , 440	3.46	0.0164
Computational Complexity	1 , 440	5.57	0.0187
Rule Complexity	1 , 440	2.06	0.1519

The test of the null hypothesis that no difference in variability due to complexity treatment effect exists, provided an F ratio of 3.46 with 3 and 440 degrees of freedom. This test statistic corresponds with an Observed Significance Level of 0.0164.

Splitting the complexity treatment into its two components provides two additional null hypotheses to test: (1) there is no

difference in variability due to computational complexity; and (2) there is no difference in variability due to rule complexity. The test of the former hypothesis resulted in an F ratio of 5.57 with 1 and 442 degrees of freedom, and a corresponding observed significance level of 0.0187. The latter hypothesis had an observed significance level of 0.1519 based on an F ratio of 2.06 with 1 and 442 degrees of freedom. The statistics in Table III indicate that variability in compliance behavior is sensitive to the degree of computational complexity present. Rule complexity does not appear to be as significant. The null hypothesis, i.e., there is no difference in variability due to rule complexity, cannot be rejected at alpha levels up to 0.15. These results would indicate that as the level of computational complexity increases there will be a corresponding decrease in compliance due to an increase in the frequency and degree of errors. With regard to rule complexity it is possible that the objective and subjective treatments used in this experiment were not sufficiently different to cause a significant difference in compliance.

Summary

Based on analyses of the results it appears that differing levels of computational complexity will effect the degree of varability present in compliance. However, the results of the experiment are not able to support the contention that rule complexity has a similar effect on compliance. Neither form of complexity appeared to have any significant effect on the overall level of compliance in the study. The different levels of rule complexity (objective and subjective) may not have differed significantly because the two treatments may not have

been sufficiently different. The presence of strong compliance incentives may have inhibited any differences in the overall level of compliance.

CHAPTER V

SUMMARY AND CONCLUSIONS

Introduction

The objective of this chapter is to provide a brief overview of the research, provide conclusions drawn from the study, discuss the limitations of the study and to offer recommendations for future research.

Overview

The purpose of this research was to determine if complexity has a significant impact on the compliance with a self-assessing and reporting system. Two forms of complexity were examined: computational complexity and rule complexity. This relationship is of interest because the perceived increase in taxpayer noncompliance is, in part, attributed to tax law complexity. Legal and practical constraints do not allow a direct study of the compliance, complexity relationship that exists in the actual economy. Therefore, this research project made use of a generic task laboratory experiment methodology.

The experiment was structured so that undergraduate students were surrogates for taxpayers. The students graded their own examinations. Towards this end, the students were provided with instructions and forms (rules) to utilize in determining and reporting their examination score. This represented the task surrogate for determining and

reporting of tax liability. The instructions and forms were construct ed to capture two levels of computational complexity (simple and complex) and two levels of rule complexity (objective and subjective). The examination scores reported by the student-subjects were utilized in determining their course grade. The experimentor also determined an exam score for each student (i.e., the correct score) which was compared to the reported score. Students were not specifically aware of this fact. The differences between the reported scores and correct scores were statistically analyzed. Additionaly, a post experiment questionnaire completed by the subjects was analyzed to ascertain the perceived salience of various experimental factors by students.

Conclusions

The statistical analysis generated two conclusions which merit further discussion and interpretation. These are:

(1) Overall level of noncompliance is not significantly affected by complexity.

(2) Variability in compliance among individuals is affected by computational complexity.

Within the tax structure there is a tendency to under report tax liability [IRS, 1969]. The lack of that tendency in the results of the present study has policy implications from two standpoints. First, if noncompliance is defined as the difference between reported and actual tax liability in the aggregate, then based on the results of this study complexity does not appear to be a significant attribute in affecting compliance because the overall level of noncompliance was constant across treatments. Second, the lack of noncompliance in the aggregate

provides some insight as to attributes which may have significant influence on the level of compliance. The sum of reported exam scores for the 444 observations exceed the sum of actual exam scores by seven points, or an average difference of 0.015 points per exam. This fact indicates a much greater effort to comply, with rules, by the subjects than may be present within the self-assessing income tax system of the United States. The presence of a strong compliance incentive may be a cause of the lack of a treatment effect. The attributes which created this strong incentive warrant further investigation because they may have the potential to significantly affect the level of compliance. In the present study the enforcement structure may have been largely responsible for the high degree of compliance.

Based on the analysis of the results, there appears to be a significant relationship between the degree of computational complexity and the variability between reported and actual scores. That is, the frequency and degree of errors increase as complexity increases. This finding has intuitive appeal. As a task becomes more difficult, one would expect more errors. However, the overall level of compliance is not affected, as discussed in the previous chapter. This would indicate that a change in tax law which decreased the level of complexity would cause a corresponding decrease in errors (variability) but not significantly affect the overall level of compliance. Accordingly, if the goal of the policy makers is to reduce the occurences of inadvertent noncompliance, then reduction in the complexity level is warranted. However, the type of noncompliance which appears to concern the Treasury and the IRS is "negative" noncompliance (i.e., underpayment of true tax liability) [IRS, 1983]. Results of this experiment appear to

indicate that this form of noncompliance is not of an inadvertent nature. Therefore, it appears that reduction of the complexity level will not significantly alter overall negative noncompliance.

It may still be desirable to seek simplification of tax law for reasons other than affecting the overall level of compliance. It is possible that simplification could increase the equity of the tax law, or at least the perception of equity. Also, decreased complexity might reduce compliance cost and result in a more efficient use of resources in the private sector. The findings of this research tend to indicate that the heuristic argument supporting a strong relationship between complexity and the overall level of tax compliance is erroneous. Accordingly, if the government wishes to enhance the overall level of tax compliance, Congress should consider alternatives other than tax simplification.

Limitations

Behavioral experimental studies raise questions with regard to internal and external validity by their very nature. The methodology and design of the present study were such that the internal validity should be insured. The key to the existence of external validity is experimental realism. All factors indicated that the subjects were sufficiently involved in the task to insure such realism. The explicit reward/penalty structure greatly enhanced the realism. The incentive (student's grades) was significant. However, the ability to generalize the results is limited because the presence of no variation between reported and correct scores in the aggregate may have been due to a high level of motivation towards compliance--a level which may not be

representative of the typical taxpayer. There are several possible explanations for this phenomenon.

One possible explanation is the existence of a corresponding condition. Compliance with rules is increased when there is rule maker-individual interaction. One empirical study found that there was a difference in compliance when correspondent versus non-correspondent conditions exist [Thibaut, Freidlan and Walker, 1974]. In this study a correspondent condition also existed since the experimenter was also the classroom teacher. The condition was constant across the sample.

Another possible source of the strong effort to comply may have been the deterrents to noncompliance. The subjects may have viewed the magnitude and probability of loss (ranging from loss of points on examinations to an F in the course) in a different perspective than a taxpayer's view of noncompliance sanctions. The apparent lack of significant differences in the variability of the rule complexity treatments may be inherent in the treatment design utilized in the experiment. It is possible that the objective and subjective treatments used in this experiment were not sufficiently different to cause a significant difference in compliance.

Future Research

The government has expressed concern about inadvertent noncompliance. However, the greater concern is actually with negative noncompliance (actual tax liability greater than reported liability) [IRS, 1983]. Based on the results of the this study, it appears that complexity is not a significant cause of negative noncompliance. Additionally, it appears that the audit and penalty structure within the

experiment had significant impact on the students' attempts to comply with the rules provided. This observation is supported by the post-experiment questionnaire and also by informal student feedback. It follows that penalty and audit structure may constitute a significant compliance attribute and a more effective tool in curbing and controlling the apparent rise in noncompliance than tax simplification.

SELECTED BIBLIOGRAPHY

Abdel-Khalik, A. R. [1974], "On the Efficiency of Subject Surrogation in Accounting Research," The Accounting Review, (October 1974), pp. 743-50.

Allingham, M. G., and Sandmo A. [1972], "Income Tax Evasion: A Theoretical Analysis," Journal of Public Economics, (November, 1972), pp. 323-38.

Aronson, E. and Carlsmith, V. M., "Experimentation in Social Psychology," Handbook of Social Psychology, 2nd ed., Addison-Wesley Publishing Co., 1968.

Ashton, R. H. and Kramer, S. S. [1980], "Students as Surrogates in Behavioral Accounting Research: Some Evidence," The Accounting Review, (Spring 1980), pp. 1-15.

Ball, J. C. [1955], "The Defference Concept in Criminology and Law," Journal of Criminal Law, Criminology and Policy Science, (September, 1955), pp. 347-54.

Birberg, J. G. and Nath, R. [1968], "Laboratory Experimentation in Accounting Research," The Accounting Review, (January 1968), pp. 468-79.

Box, G. E. P., "Some theorems on quadratic forms applied in the study of analysis of variance problems," Annals of Mathematical Statistics, (Vol. 25, 1954), pp. 290-302.

Brighton, G. E., Michaelsen, R. H., and Willis, E. [1978], "The Academic Accountant's Role in Tax Research," working paper, University of Illinois, 1978.

Campbell, D. T. and Stanley, J. C. [1963], Experimental and Quasi-Experimental Designs for Research, (Boston: Houghton Mifflin Company, 1963).

Cherulnik, P. D. [1983], Behavioral Research--Assessing the Validity of Research Findings in Psychology, (New York: Harper & Row, 1983).

Cialdini, R. B. and Ascani, K. [1976], "Test of a concession procedure for inducing verbal, behavioral and further compliance with a request to give blood," Journal of Applied Psychology, (June, 1976), pp. 295-300.

Clotfelter, C. T. [1983], "Tax Evasion and Tax Rates: An Analysis of Individual Returns," The Review of Economics and Statistics, (August 1983), pp. 363-73.

Dayton, C. M. [1970], The Design of Educational Experiments, (New York: McGraw-Hill, 1970).

Dickhaut, J. W., Livingstone, Watson, "On the Use of Surrogates in Behavioral Experimentation," Accounting Review (Supplement, 1972), pp. 455-71.

Erickson, M. L., Gibbs, J. P., and Jensen, G. F. [1977], "The Deterrence Doctrine and the Perceived Certainty of Legal Punishments," American Sociological Review, (April 1977), pp. 305-17.

Ferguson, G. A. [1976], Statistical Analysis in Psychology & Education, 14th Edition, (New York: McGraw Hill Book Company, 1976).

Fishburn, G. [1981], "Tax Evasion and Inflation," Australian Economic Papers, (December, 1981), pp. 325-32.

Freedman, J. L. and Fraser, S. C. [1966], "Compliance without pressure: The foot-in-the-door technique," Journal of Personal and Social Psychology, (August, 1966), pp. 195-202.

Friedland, N., Maital, S., and Rutenberg, A. [1978], "A Simulation Study of Income Tax Evasion," Journal of Public Economics, (August, 1978), pp. 107-16.

Forward, J. and Zander, A. [1971], "Choice of unattainable group goals and effects on performance," Organizational Behavior and Human Performance, (May, 1971), pp. 184-199.

Glass, G. V., Peckman, P. D., and Sanders, J. R. [1972], "Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance," Review of Educational Research, (Summer, 1972), pp. 237-288.

Goldman, M., Gier, J. A., and Smith, D. E. [1981], "Compliance as Affected by Task Difficulty and Order of Tasks," The Journal of Social Psychology, (June, 1981), pp. 75-83.

Goode, R. [1964], The Individual Income Tax, (Washington, D.C.: The Brookings Institution, 1964).

Gutman, P. [1977], "The Subterranean Economy," Journal of Financial Analysis, (33, Nov/Dec 1977).

Henry, J. S. [1983], "Noncompliance with U.S. Tax Law--Evidence on Size, Growth and Composition," Income Tax Compliance: A Report of the American Bar Association Section of Taxation Invitational Conference on Income Tax Compliance, March, 1983, pp. 15-112.

Heider, F. [1958], The Psychology of Interpersonal Relationships, (New York: Wiley, 1985).

Internal Revenue Service [1969], Taxpayer Compliance Measurement Program, (Washington, D.C.: Government Printing Office, 1971).

Internal Revenue Service [1979], Estimates of Income Unreported on Individual Income Tax Returns, (Washington, D.C., Government Printing Office, September 1979).

Internal Revenue Service [1983], Office of the Assistant Commissioner of Planning and Research Division, 1983 Income Tax Compliance Research--Estimate for 1973-81 (July 1983).

Joyce, E. S. and Libby, R. [1982], "Behavioral Studies of Audit Decision Making," Journal of Accounting Literature, (Spring, 1982), pp. 103-123.

Laufer, D. M. [1980], "Potential Statutory Liabilities and Penalties of the Tax Return Preparer," Masters Thesis, Colorado State University, 1980.

Locke, E. A. [1966], "The relationship of intentions to level of performance," Journal of Applied Psychology, (February, 1966), pp. 60-66.

_____ [1968], "Toward a theory of task motivation and incentives," Organizational Behavior and Human Performance, (March, 1968), pp. 157-189.

Marquardt, J. D. [1975], "Income Tax Compliance--An Examination of Critical Attributes," Ph.D. dissertation, University of Illinois, 1975.

Mork, K. A. [1975], "Income Tax Evasion: Some Empirical Evidence," Public Finance, (Spring, 1975), pp. 70-6.

Raabe, W. [1980], "Income Tax Simplicity: An Examination of an Elective Filing System," Ph.D. dissertation, University of Illinois, 1980.

Runkel, P. J. and McGrath, J. E. [1972], Research on Human Behavior, (New York: Holt, Rinehart and Winston, Inc., 1972).

Sommerfield, R. M., Anderson, H. M., and Brock, H. R. [1982], An Introduction to Taxation, (New York: Harcourt Brace Jovanovich, 1982).

Swieringa, R. S. and Weick, K. E., "An Assessment of Laboratory Experiments in Accounting," paper presented at Current Research Methodologies in Accounting Conference, University of Chicago, April 29 and 30, 1982.

Tanzi, V. [1982], "Underground Economy and Tax Evasion in the United States: Estimates and Implications," The Underground Economy in the United States and Abroad, Lexington Books, 1982.

Thibaut, S., Friedland, N., and Walker, L. [1974], "Compliance with Rules: Some Social Determinants," Journal of Personality and Social Psychology, (December, 1974), pp. 792-801.

Tittle, C. R. [1969], "Crime Rates and Legal Sanctions," Social Problems, (Winter, 1969), pp. 409-23.

U.S. Congress. House. Committee on Ways and Means. H. Rept. 1337, 83d Cong., 2d Sess. 1 (1954).

U.S. Congress. Public Law 95-30, 95th Congress, H.R. 3477, May 23, 1977.

U.S. Congress. S. Rept. 66, 95th Cong., 1st Sess.; H. Rept. 27, 95th Cong., 1st Sess.

U.S. Congress. Joint Committee on Taxation. Issues in Simplification of the Income Tax Laws. 95th Congress, 1st Session, 1977.

U.S. Congress. Senate. Committee on Finance. Hearings on Compliance Gap. 97th Congress, 2d Session, 1982.

U.S. Congress. S. Rept. 1421, 99th Cong., 2d Sess., 1984.

U.S., Constitution. Amendment XVI.

U.S. Department of Treasury [1977]. Blueprints for Basic Tax Reform, (Washington, D.C., Government Printing Office, January, 1977).

U.S. Department of Treasury [1984]. Tax Reform for Fairness, Simplicity, and Economic Growth, (Washington, D.C., Government Printing Office, November, 1984).

Underwood, B. J. [1949], Experimental Psychology, (New York: Appleton-Century Crafts, 1949).

Wall Street Journal [1984], "IRS is Losing Battle Against Tax Evaders Despite Its New Gear," 10 April 1984, p. 1.

Zelditch, M. [1969], "Can you Really Study an Army in the Laboratory," in A. Etzioni (ed.), A Sociological Reader on Complex Organizations, 2nd ed. (Holt, Rinehart & Winston, Inc. 1969), pp. 533-39.

_____ and Evan, W. M. [1962], "Simulated Bureaucracies: A Methodological Analysis," Simulation in Social Science Readings, edited by H. Geutzkow (Prentice-Hall, 1962), pp. 48-60.

APPENDIXES

APPENDIXES

The following appendixes provide examples of the various instruments utilized in conducting the experiment. The intent is to provide sufficient information here so that the experiment can be replicated. Accordingly, a discussion of the content of the appendixes is provided below.

Appendix A includes a flow chart of the experiment procedures and an Examination Format Summary. The flow chart provides an overview of the sequence of events during the experimental process. The examination format was consistent for all four examinations. The 160 total examination points were utilized in an effort to avoid anchoring on the traditional grading scale that many student-subjects associate with a 100-point examination. Additionally, grading curve information was not provided to the subjects until examination scores had been reported.

Appendix B contains samples of examination questions and corresponding answers and instructions. Examples are presented for fill-in-the-blank problems, journal entry problems, and accounting number problems. These three types of questions were designed to facilitate application of the rule complexity treatments. This was achieved by providing either subjective or objective information in the answer sheet and grading instructions. The examination questions were varied between the two treatments to negate student collusion which would have contaminated the results. Note that the last sentence of grading instructions informs the subject which form or schedule to report their

score on. These instructions facilitate the application of the computational complexity treatment.

Appendix C provides copies of the reporting forms utilized for the experiment. Form One S was used for simple computations. Form One and related schedules were required for the complex computations treatment. Each subject utilized the simple reporting form twice and was also required to employ the more complex form on two occasions. These computational complexity treatments were matched with the rule complexity treatments. The sequencing on a group-by-group basis is detailed in Figure 2 (p. 33).

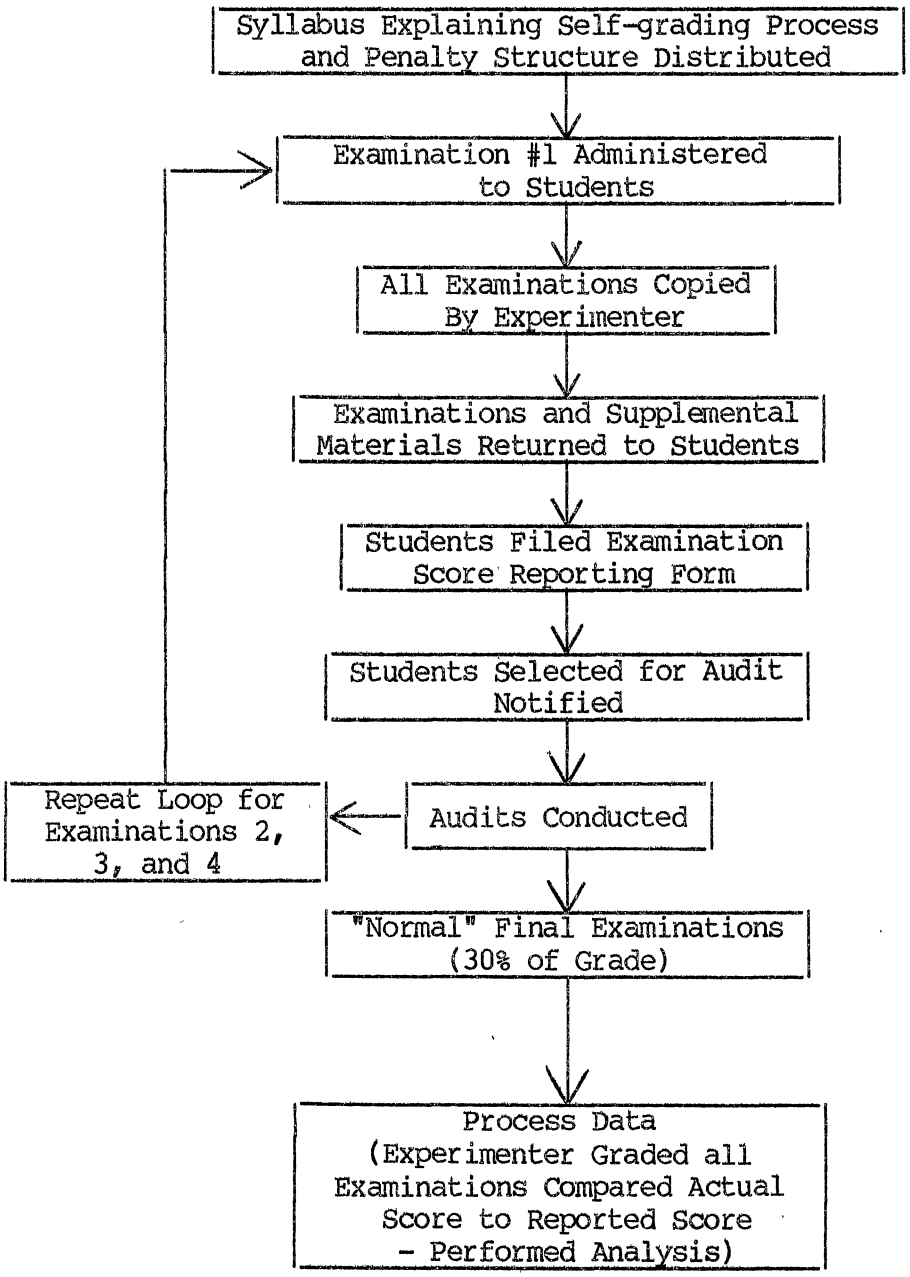
Appendix D presents the Syllabus Supplement and the audit notification letter. The syllabus supplement was distributed the first day of the semester. The procedures to be followed during the semester were described in the supplement. The penalty provisions contained therein were employed uniformly throughout the semester. The audit letters were distributed to the subjects, either in class or by U.S. mail, within a week of filing their reporting form.

Appendix E contains the post-experiment questionnaire and results which were administered upon completion of the experiment, at the end of the semester.

APPENDIX A

EXPERIMENTAL PROCEDURES OVERVIEW

STEP-BY-STEP--EXPERIMENT PROCEDURES



EXAMINATION FORMAT

The examination design was the same throughout the semester. Each examination consisted of five sections as follows:

1.	20 Multiple Choice Questions	60 points
2.	10 True/False Questions	10 points
3.	10 Fill-in-the-blank Questions	30 points
4.	Journal Entry Problems	20 points
5.	Accounting Number Problems	40 points
	Total	160 points

The examinations were designed so that the content tested was equal between treatments. However, the actual questions varied between treatments to negate student collusion.

When the students had their examinations returned they also received:

1. Answer Sheet
2. Grading Instructions
3. Reporting Forms

The version they received was a function of treatment class.

APPENDIX B

ILLUSTRATIVE EXAMINATION QUESTIONS, ANSWERS, AND INSTRUCTIONS

ILLUSTRATIVE FILL-IN-THE-BLANK PROBLEM

SUBJECTIVE TREATMENT

Exam Question

1) The _____ is the financial statement reporting the profitability of a business entity during a specific period of time.

Answer Sheet

In this section the correct answer which is most commonly utilized is listed below. In some instances more than one word or phrase have the same meaning. If your answer is synonomous with the one listed below then it is correct. Your answer must have exactly the same meaning as the answer given below to be considered correct.

1) Income Statement

Grading Instructions

Each correct answer is worth three points. Total points earned in this section is equal to number of correct answers multiplied by 3. Record that number on Form One S, Line 3.

OBJECTIVE TREATMENT

Exam Question

1) The_____ is the financial statement that provides information about an entity's resources and claims against those resources as the end of a specific period of time.

Answer Sheet

In this section it is possible that different words or phrases exist which carry the same meaning. Accordingly, for some questions more than one answer is listed below. To be correct your answer must match up with one of the answers provided. The use of common abbreviations is acceptable.

	Correct Answer	Acceptable Alternatives
1)	Balance Sheet	Statement of Financial Position

Grading Instructions

Each correct answer is worth three points. Enter the total number of correct answers on Schedule FB, Line 1.

ILLUSTRATIVE JOURNAL ENTRY PROBLEM

SUBJECTIVE TREATMENT

Exam Question — The examination questions gives two transactions and requires journal entries be made.

Answer Sheet

For this problem the most common account title is utilized for journal entry purposes. In some instances a different or modified version of title provided could be correct. Your account title will be considered correct if it communicates the proper meaning, with regard to the transaction, within the scope of generally accepted accounting terminology

Cash	6,000	
Common Stock		6,000
Accounts Payable	800	
Cash		800

Grading Instructions

Problem 2 - Each line of each journal entry is worth two points—one for account title and one for dollar amount. If the correct account and amount are selected but recorded improperly (i.e., debited when a credit was appropriate, or visa versa) then only one point is earned, instead of two. If reasonable journal entry format is not utilized then no points are earned. Sum up points earned and enter on Schedule JE, Line 2.

OBJECTIVE TREATMENT

Answer Sheet

Cash	6,500	
Accounts Receivable		6,500
Cash	4,000	
*Common Stock		2,000
**Paid-in Capital in excess of par, common stock		2,000

*Capital, capital stock or stock would all be considered acceptable substitutes for the account title "common stock."

**Paid-in capital is an acceptable shortened account for the answer provided above

Note that the use of common abbreviations (such as A/R for Accounts Receivable) is acceptable

Grading Instructions

Problem 2 - Each line of each journal entry is worth two points--one for account title and one for dollar amount. If the amount is correct but the incorrect account title is used, or visa versa, then only one point is earned. If the correct account and amount are selected but recorded improperly (debit entry made when credit entry was appropriate, or visa versa) then only one point is earned, instead of two. No points are earned if answers are given in such a way that they cannot be identified as representing journal entries (such as listing account names and dollar amounts in one column. Variations on standard format such as using brackets to indicate credit amounts will be tolerated. Use of common account title abbreviations is acceptable. Sum up points earned and enter on Form One S, Line 5.

ILLUSTRATIVE ACCOUNTING NUMBER PROBLEM

SUBJECTIVE TREATMENT

Exam Question The examination question provides beginning of the month account balances and six events (transactions) which occur during the month. The students are required to determine end of month account balances for:

A) Cash
B) Loan Payable
C) Common Stock
D) Accumulated Depreciation

Answer Sheet

A)	$22,000	C)	$20,000
B)	$ -0-	D)	$11,000

Grading Instructions

Problem two - Each correct answer is worth five points. If an incorrect answer is given, partial credit of two points may be earned as follows. If you failed to consider the impact of some, but not all, of the transactions on the account balance, then for every transaction for which you properly adjusted the account balance in question one point is earned. If the impact of all the transactions on an account balance was made in arriving at the answer three points are earned. Sum up points earned and enter on Line 7 of Form One S.

OBJECTIVE TREATMENT

Interest Payable
Paid-In Capital
Depreciation Expense
Cash

Answer Sheet

A)	$ -0-	C)	$ 3,000
B)	$ 5,000	D)	$21,000

Grading Instructions

Problem one - Each correct answer is worth five points. Partial credit for incorrect answers may be earned as follows:

A) Interest Payable - If your incorrect answer was the result of failure to adjust the account balance for the impact of the Jan. 1 transaction then one point partial credit is earned. If your incorrect answer was the result of a math error made in determining the account balance after the Jan. 1 transaction, then three points partial credit are earned.
B) No Partial Credit
C) Depreciation Expense - If your incorrect answer was the result of failure to adjust the account balance for the Jan. 31 event then one point partial credit is earned.
D) Cash - There are several transactions which effect the cash account. The Jan. 1 transaction causes a decrease (credit) in cash of $10,500. The Jan. 16 transaction causes an increase of $55,000, the Jan. 22 transaction a $4,000 decrease, and the Jan. 26 transaction a $35,500 decrease. For each of these handled correctly one point is earned. If you handled all transactions correctly but made a math error in arriving at the account balance as of Jan. 31 then 3 points are earned.

Sum up points earned and enter on Line 7 of Form One S.

APPENDIX C

EXAMINATION SCORE REPORTING FORMS

REPORTING FORM FOR SIMPLE COMPUTATIONAL TREATMENT

FORM ONE S

Name_____ Course Section Number_____

Address_____ Exam Number_____

Phone Number_____ Exam Code_____

Student ID#_____ Exam Score_____
(From Line 8 below)

Year in School_____

Major_____

1. Section one, points earned _____
2. Section two, points earned _____
3. Section three, points earned _____
4. Section four, problem one, points earned _____
5. Section Four, problem two, points earned _____
6. Section five, problem one, points earned _____
7. Section five, problem two, points earned _____
8. EXAM Score, Add lines 1 through 7 _____

I have examined this report form and to the best of my knowledge and belief the information reported herein is true, correct, and complete. I understand that any variance between the exam score reported and an audited exam score will result in a loss of exam points, based on the penalty structure detailed in the course syllabus. Additionally, if there is evidence that the cause of the variance is due to academic dishonesty or misconduct, then penalties allowed under university guidelines will be invoked.

_____ _____
(Your Signature) (Date)

REPORTING FORM AND RELATED SCHEDULE
FOR COMPLEX COMPUTATIONAL TREATMENT

FORM ONE

Name_____ Course Section Number_____

Address_____ Exam Number_____

Phone Number_____ Exam Code_____

Student ID#_____ Exam Score_____
(From Line 12 below)

Year in School_____

Major_____

1.	Percent Correct, Multiple Choice Questions (attach Schedule MC)	____.____%
2.	Multiply the percentage on Line 1 by .375 and enter here	____.____%
3.	Percent Correct, True/False Questions (attach Schedule TF)	____.____%
4.	Multiply the percentage on Line 3 by .0625 and enter here	____.____%
5.	Percent Correct, Fill-in-the-Blank Questions (attach Schedule FB)	____.____%
6.	Multiply the percentage on Line 5 by .1875 and enter here	____.____%
7.	Percent Correct, Journal Entries (attach Schedule JE)	____.____%
8.	Multiply the percentage on Line 7 by .125 and enter here	____.____%
9.	Percent Correct, Problems (attach Schedule PB)	____.____%
10.	Multiply the percentage on Line 9 by .25 and enter here	____.____%
11.	Total Percent of Correct Answers (Add Lines 2, 4, 6, 8, and 10)	____.____%
12.	EXAM Score (Multiply Line 11 by 160)	_____

I have examined this report form and to the best of my knowledge and belief the information reported herein is true, correct, and complete. I understand that any variance between the exam score reported and an audited exam score will result in a loss of exam points, based on the penalty structure detailed in the course syllabus. Additionally, if there is evidence that the cause of the variance is due to academic dishonesty or misconduct, then penalties allowed under university guidelines will be invoked.

_____ _____
(Your Signature) (Date)

SCHEDULE MC — Multiple Choice Section

Name_____ Student Number_____

1. Multiple Choice, number of correct answers _____

2. Multiply the number on Line 1 by 3 _____

3. Multiple Choice, number of incorrect answers _____

4. Subtract Line 3 from Line 2. If less than zero, enter zero. _____

5. Divide number on Line 4 by 60. Write the percentage here and on Form One, Line 1 _____.___%

SCHEDULE TF — True/False Section

Name_____ Student Number_____

1. True/False, Number of correct answers _____

2. True/False, Number of incorrect answers _____

3. Multiply number on Line 2 by .5 _____

4. Subtrace Line 3 from Line 1. If less than zero, enter zero. _____

5. Divide number on Line 4 by 10. Write the percentage here and on Form One, Line 3 _____.___%

SCHEDULE FB — Fill-in-the-Blank Section

Name_____ Student Number_____

1. Fill-in-the-Blank Questions, number of correct answers _____

2. Multiply number on Line 1 by 3 _____

3. Divide number on Line 2 by 30. Write the percentage here and on Form One, Line 5 _____.___%

SCHEDULE JE — Journal Entry Section

Name_____ Student Number_____

1. Journal entries, problem one, number of points earned _____

2. Journal entries, problem two, number of pionts earned _____

3. Journal entries, total points earned (add lines 1 and 2) _____

4. Divide the number on Line 3 by 20. Write the percentage here and on Form One, Line 7 _____.___%

SCHEDULE PB — Accounting Number Problems Section

Name_____ Student Number_____

1. Accounting Problems, problem one, number of points earned _____

2. Accounting Problems, problem two, number of points earned _____

3. Accounting Problems, Total Points earned, (Add Lines 1 and 2) _____

4. Divide the number on Line 3 by 40. Write the percentage here and on Form One, Line 9. _____.___%

APPENDIX D

SYLLABUS SUPPLEMENT--PENALTY STRUCTURE AND AUDIT NOTIFICATION LETTER

SYLLABUS SUPPLEMENT - EXAMINATION PROCEDURES

For purposes of grading the exams in this course, a self-assessing (self-grading) procedure will be utilized. The following is intended to provide a clear description of how this system will function.

Exam day will follow typical procedures--exams will be closed-book and given during class time. Upon completion of the exam, or at the end of class time, whichever comes first, the exam is to be handed in to the instructor.

At the next class meeting, the students will have the exams returned, along with answer sheet, grading instructions and reporting forms. The student is to complete the reporting forms by utilizing the answer sheet and grading instructions. The reporting form will communicate to the instructor the points earned on the exam. Reporting forms will be due at the beginning of the third class period following the exam (1 week after the exam is returned). The student need turn in only the required form(s). The exam should be kept by the student as support for grading computations! Additionally, the student may want to keep a copy of reporting forms.

There are different exam versions and, accordingly, different answer sheets, grading instructions, and reporting forms. You must use the answer sheets, grading instructions, and reporting forms provided to you. Use of materials provided to other students could cause assessment of penalty points for use of improper forms.

Note that 70% of your grade in this course will be based on the exam scores you report. Therefore, it is in your best self-interest to follow the instructions to avoid loss of exam points due to penalty assessments.

Enforcement Procedures

In order to assure compliance with grading instructions, reported exam scores will be subject to audit verification with penalties assessed for noncompliance.

Exams will be selected for audit verification using stratified random sampling. If your exam is selected for audit, it is essential that you are able to produce evidence to support the reported score. Accordingly, you must retain your exam.

Audit Procedures

Those selected for audit will be informed of such shortly after filing the reporting forms. Audit conference time and date will be set. At the audit conference the student will be asked to support their exam score calculations, either in whole or in part. Towards this end it is critical that the student keep their exams as it will be the primary source of backup for reported score. The auditor will have the report forms the student has filed. Additionally, random pages of random

exams will be xeroxed by the auditor and may be used for audit purposes. If there is a discrepency between the reported score and the audited score a penalty will be assessed.

GENERAL PENALTIES

Penalty for Late Filing

Students will be assessed a penalty for late filing of an exam score on the required forms as follows. If forms are filed after the due date, but before the next class meeting, the penalty will be 5% of reported exam score. For filing one or two class meetings late the penalty will be 10% of reported exam score. If filing occurs following the second class meeting after the due date, the penalty will be 25% of reported score. Additionally, there will be a failure to file penalty--a zero on the exam.

Penalty for Filure to Sign Reporting Form

If the exam score reporting form lacks a proper signature, then a penalty of 1% of reported score will be assessed.

Penalty for Use of Improper Forms

If the exam score reporting form is not the same as the form required as per the original exam and grading instructions, then a penalty of 10% of reported score will be assessed.

Additionally, failure to attach required schedules to exam score reporting form will result in assessment of a penalty of 1% of reported exam score, per schedule.

Math Error Corrections

If math errors are discovered in the exam score reporting form, then the reported score will be adjusted. This adjustment will be made without an audit conference. Students will be informed of such adjustment.

Audit Penalty Structure

1) There will be an automatic "Interest" assessment of 1% of the reported exam score if the reported exam score is in excess of the audited score.

2) An additional penalty will be assessed based on the following sliding scale:

Excess of Reported Score Over Audited Score (in points)	Penalty Assessment in % of Reported Score
1-5	0%
6-10	5%
11-15	10%
16-20	15%
21-25	20%
over 25	25%

3) If fraud (cheating) is evident, there will be a fraud penalty assessed, ranging from loss of one letter grade to an F in the course.

4) No penalties will be assessed if the audited score is greater than the reported score.

5) If the student cannot produce evidence to support reported exam score, then, for those section(s) of the exam that are audited, the score will be zero and the scope of the audit may be expanded.

This examination procedure should enhance the learning process. It also places a high degree of responsibility on you. Time deadlines will be enforced; therefore, if you miss class, you should make other arrangements with me to pick up your exam and related materials or to turn in report forms. Finally, I cannot emphasize strongly enough how critical it is that you retain your exam copy--without it you will be unable to justify your reported score should you be called on to do so.

AUDIT NOTIFICATION LETTER

Doug Laufer
Acctg. 2103
An Husb 005E
624-7619

Sally Smith
433 Drummond
Stillwater, OK
SS#aaa-aaa-aaaa

Your Acctg. 2103 exam (exam #3) has been selected for audit verification. The sections of your exam to be reviewed are noted with checkmarks:

_____ Section One

_____ Section Two

_____ Section Three

_____ Section Four

_____ Section Five

Audits of exams will be conducted October 9, 10 and 11. It is important you contact me to set a specific appointment time. The space below is provided to record your appointment.

APPOINTMENT INFORMATION

Place: Animal Husbandry Building Date:

Room No.: 005E Time:

For the sections indicated above, the examiner will review your exam score reporting form and related information. You will be called upon to substantiate the score you have reported. Accordingly, you will need to bring the appropriate section(s) of your exam and any other relevant material. The audit determination will be on the basis of available information only. You will be informed of any proposed exam score changes during the audit conference.

If you have any questions, please contact me at the number shown in the heading of this letter (624-7619), of if I cannot be reached at that number, call 624-6377 and leave a message.

Thank you for your cooperation.

Sincerely,

Doug Laufer

APPENDIX E

POSTEXPERIMENT QUESTIONNAIRE AND RESULTS

Accounting 2103
Experiment Questionnaire
Laufer

There is concern and interest about how students feel about the experience of using a self-grading system. Could you please answer the following questions? If you would like to commend on any facet of the experience, do so on the class evaluations. Thank you for your cooperation.

Please use the following scale to answer the questions:

(A) Definitely yes
(B) Yes
(C) Not applicable
(D) No
(E) Definitely no

Record your answers on the computer sheet using #2 lead. Do not put your name or any information, other than responses to the questions, on the sheet.

1. I enjoyed grading my own exam.

2. By grading my own exam, I learned more than I would have otherwise.

3. Grading my own exam caused me to study less for this course.

4. I feel the course workload was substantially increased due to self-grading of the exams.

5. My effort level in complying with the grading instructions was consistent from one exam to the next.

6. I felt that the grading instructions on some exams were more difficult to comply with than other exams.

7. If I was unsure of how to grade a particular question, I gave myself the benefit of the doubt.

8. If the instructor had graded my exam I think I would have received the same score.

9. If the instructor had graded my exam I would have received a lower exam score.

10. If the instructor had graded my exam I would have received a higher exam score.

11. The possibility of an audit caused me to make a stronger effort to follow the grading instructions than I would have otherwise.

12. The possibility of penalties caused me to make a stronger effort to follow the grading instructions than I would have otherwise.

13. At the beginning of the semester I expected to have all of my exams audited.

14. At the beginning of the semester I expected to have at least one of my exams audited.

15. I found complying with the grading instructions created stress for me.

16. I feel that the self-grading system was unfair.

TABLE IV

POSTEXPERIMENT QUESTIONNAIRE--
SUMMARY RESULTS

Question Number	Mean	Standard Deviation
1	3.618	1.100
2	3.472	1.090
3	1.791	0.779
4	2.264	0.945
5	4.227	0.725
6	3.518	1.106
7	2.455	1.209
8	3.782	0.942
9	1.991	0.697
10	2.609	1.050
11	3.945	1.056
12	3.930	1.059
13	2.300	0.863
14	4.091	0.808
15	2.192	1.014
16	2.189	0.957

VITA 2

Douglas M. Laufer

Candidate for the Degree of

Doctor of Philosphy

Thesis: AN EXPERIMENTAL STUDY OF THE IMPACT OF RULE COMPLEXITY ON RULE COMPLIANCE

Major Field: Business Administration

Biographical:

Personal Data: Born in New York City, New York, December 28, 1952, the son of Mr. and Mrs. A. P. Laufer, Jr.

Education: Graduated from Alameda High School, Lakewood, Colorado, in June 1971; received Bachelor of Science degree in Accounting from the University of Denver, in 1974; received Masters of Taxation degree from Colorado State University, in 1980; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1985.

Professional Experience: Staff Accountant, Stone, Gray and Company, CPAs, August, 1974 to August, 1976; Partner, McMahon, Sipp and Company, CPAs, August, 1976 to December, 1978; Graduate Teaching Assistant, Colorado State University, January, 1979 to May, 1980; Instructor, Northern Arizona University, August, 1980 to May, 1982; Lecturer, Oklahoma State University, August, 1982 to May, 1985.

Professional Memberships: American Accounting Association; American Institute of Certified Public Accountants; American Taxation Association; Colorado Society of Certified Public Accountants; The Athletic Congress.