11,145

REPORT BY THE

Comptroller General

OF THE UNITED STATES

RESTRICTED — Not to be released outside the General Accounting Office except on the basis of specific approval by the Office of Congressional Relations.

IRS' Audits of Individual Taxpayers And Its Audit Quality Control System Need To Be Better

At the request of the Joint Committee on Taxation, GAO evaluated the quality of IRS audits of individual taxpayers. GAO determined that the quality of these audits needs to be improved. Also, IRS' system for identifying and correcting less than quality audits is not as effective as it should be. The quality control system, in fact, led IRS to believe that quality was not in need of improvement.

Among the results of GAO's analysis of 490 randomly sampled IRS audits that had passed through IRS' quality control system in 5 selected districts were the following:

--These districts assessed incorrect tax estimated to total $1.0 million over a 4-month period because of technical errors, computation errors, or failure to make automatic adjustments.

--IRS personnel too frequently overlooked important audit issues. Since those issues were not audited, GAO lacked the data needed to compute the related tax impact.

IRS has taken some actions to improve audit quality and its quality control system, and is planning others. GAO made a number of recommendations that IRS should include in those plans.

UNITED STATES GENERAL ACCOUNTING OFFICE

110353

GGD-79-59
AUGUST 15, 1979

COMPTROLLER GENERAL OF THE UNITED STATES
WASHINGTON D.C. 20548

B-137762

To the Chairman and Vice Chairman
Joint Committee on Taxation
Congress of the United States

This report, one in a series in response to your Committee's request, points out the need to improve the quality of the Internal Revenue Service's audits of individual taxpayers and the effectiveness of the Service's system for identifying less than quality audits. The Service agreed to most of our recommendations, and effective implementation should improve its audits and the data provided to the Service's management by the audit quality control system.

As arranged with your Committee, unless you publicly announce its contents earlier, we plan no further distribution of the report until 30 days from its date. At that time, we will send copies to interested parties and make copies available to others upon request.

Comptroller General
of the United States

COMPTROLLER GENERAL'S REPORT TO THE JOINT COMMITTEE ON TAXATION

IRS' AUDITS OF INDIVIDUAL TAXPAYERS AND ITS AUDIT QUALITY CONTROL SYSTEM NEED TO BE BETTER

D I G E S T

The quality of IRS audits of taxpayers' individual income tax returns needs to be improved. When IRS does less than quality audits

--some taxpayers may pay more than they owe,

--tax revenues may be lost and

--voluntary compliance may suffer.

Likewise, IRS' system for identifying and correcting less than quality audits is not as effective as it should be. That system, in fact, led IRS to believe that audit quality was not in need of improvement.

It is imperative that audit quality be improved to help assure the continued strength of our voluntary tax assessment system.

IRS' AUDIT QUALITY PROBLEM

GAO analyzed a statistical sample of 490 income tax audits that had passed through IRS' quality control system in 5 selected districts. Of these audit cases, 240 had received a quality review by the districts' technical review staffs (review cases), and the other 250 had been reviewed by IRS' District Conference (conference cases). (See p. 5.)

IRS' definition of audit quality did not provide adequate measurement criteria. Using existing IRS guidelines and input from IRS personnel, GAO identified basic elements which should be present in and are indicative of the quality of any IRS audit. GAO then measured the sampled audits against those elements. (See p. 4.)

Tear Sheet. Upon removal, the report cover date should be noted hereon.

The results: IRS personnel too frequently

--Assessed an incorrect tax because of technical errors, computation errors, or failure to make automatic adjustments. GAO estimated that such errors, totaling $37,000, were made in 34 percent of the review cases and 45 percent of the conference cases in the 5 districts during the 4 months GAO was selecting its sample. Projecting these results to all similar audits completed by the 5 districts during the same period, GAO estimated that these errors totaled $1.0 million. (See pp. 7 to 13.)

--Overlooked significant audit issues. This happened in 58 percent of the review cases and 47 percent of the conference cases. Since IRS had not audited these items, GAO lacked the data necessary to compute the related tax impact. (See pp. 14 to 17.)

--Did not do sufficient work to determine that the item in question had been properly reported, did not record the work they did, or both. Neither GAO nor IRS can really tell. (See pp. 17 to 20.)

--Did not make required probes for unreported income in 62 percent of the review cases and 52 percent of the conference cases. In addition, tax auditors were not required and usually did not probe for unclaimed deductions and credits. (See pp. 21 and 22.)

THE CAUSES OF QUALITY PROBLEMS

Pressures to complete audits quickly

GAO traced the reasons for quality problems to several factors, most of which were connected by the theme of time pressure exerted by management, whether real or perceived, to do audits as quickly as possible. For example:

--Some of the problems were caused by IRS procedures that tended to limit the number of items examined. (See pp. 24 to 26.)

--Opinions differed among IRS personnel about the time it takes to do a quality audit. Better than 70 percent of the tax auditors, group managers, and reviewers in the 5 sampled districts said that tax auditors had less than adequate time to conduct examinations. Conversely, IRS regional and district officials felt that the number of audits IRS planned to do did not put undue pressure on examiners and that examiners had enough time to do quality audits. IRS does not have data available with which to tell whether tax auditors have sufficient time to do a quality audit. (See pp. 26 to 28.)

--Some problems go unnoticed because of a reluctance on the part of IRS personnel to raise questions on cases in which the taxpayer has agreed with the examiner's tax adjustment and because group managers, IRS' first level of review, have been instructed not to conduct in-depth reviews. (See pp. 29 and 30.)

--IRS has not provided its examiners with aids to improve audit depth and documentation. (See p. 31.)

IRS not aware it had a problem

When GAO started its work, IRS management, as a group, described IRS' audit quality as "good" or better. Even the auditors and their supervisors rated IRS quality as high or very high. These perceptions go a long way toward explaining why the problem went unnoticed.

The data produced by IRS' quality control system did not alert management that audit quality needed improvement. Not knowing the need, IRS had no incentive to identify and correct the causes of its quality problem.

QUALITY CONTROL SYSTEM DATA NOT ACCURATE NOR ADEQUATE

The audit quality errors detected by GAO had passed through IRS' quality control system unnoticed. System weaknesses kept IRS from receiving accurate and complete audit quality information.

One major weakness is that reviewers have no standards to follow when reviewing audit cases. To test the effect of the lack of review standards, GAO submitted six of the sample cases from each of the five districts to a cross-district review. The 30 cases were each reviewed 5 times--once by the reviewer in the district that closed the case and once in each of the other 4 districts.

Results showed that management receives an incorrect picture of audit quality and that reviewers are inconsistent in their quality assessments. In one district, for example, the reviewer had not questioned the examiner's work on any of the six cases selected from that district. Reviewers in each of the other districts would have returned at least two of the cases to the examiners. Reviewers in one district would have returned all six cases. (See pp. 46 to 49.

Another weakness was that reviewers did not report all quality problems noted. Reviewers sometimes expedited the review process by resolving questions over the telephone. Other times reviewers merely noted their concerns in the case file. In either event, reviewers did not prepare the documentation needed to bring the problem to management's attention. If a quality control system is to serve management effectively, it must capture information accurately on all quality defects noted. (See pp. 49 and 50.)

Yet another major weakness was the quality control system's inability to provide IRS with needed information. The system now provides little more than error-rate data computed on the basis of the number of tax changes resulting from the review of completed audit cases. This data alone is not sufficient for telling the level of audit quality being attained or identifying problem areas which need attention. Consequently, IRS cannot make informed decisions on necessary corrective actions or consider adequately audit quality when deciding how many audits to do. (See p. 50.)

IRS recently instructed reviewers to start manually accumulating trend information on problems they noted. It has not, however, required the accumulation of this information on a national basis. IRS should take this additional step to facilitate identification of national problems. (See p. 50.)

IRS has not provided sufficient guidance on how to interpret the limited data produced by the quality control system. Thus, managers can react differently given the same information. IRS has established a 5-percent error rate as the alarm level for each category of reviewed cases. Management is to take corrective action when more than 5 percent of the reviewed cases in a category contain an error sufficient to cause a change in the tax auditor's adjustment. "Category," however, is not defined. Thus, depending on each manager's personal definition, it is possible that he may either view audit quality as acceptable or start developing a corrective action plan. IRS should develop uniform criteria to ensure consistent interpretation of each measure of audit quality provided by the quality control system. (See p. 51.)

IRS could improve the effectiveness of the quality control system by redesigning its regional office case review program. This program is supposed to measure and improve audit quality but now produces only limited benefits. By establishing uniform case selection and review standards for this program, IRS could assess periodically the data produced by the quality control system and ensure its continuing reliability and integrity. (See pp. 51 and 52.)

IRS HAS STARTED TO MAKE IMPROVEMENTS

From [illegible] and other sources IRS has been made aware of the need to improve its audit quality and its quality control system. It has established a Quality Control Programs Unit at the national office. This unit is responsible for improving IRS audits and the quality control system through a continuing and detailed study of audit quality. The unit already has begun to make corrections and plans [illegible]. (See pp. 31 and 52.)

GAO RECOMMENDATIONS AND IRS COMMENTS

If IRS is to use the audit function as an effective tool for increasing voluntary compliance, it must audit a sufficient number of returns to make its presence felt. But these audits have to be quality ones. The mix of audit quantity and quality that IRS needs to best accomplish its mission is a decision that properly rests with IRS. However, IRS cannot properly tackle this issue without determining how much quality really needs to be improved and how long it takes to do a quality audit.

The Commissioner of Internal Revenue should:

--Better define audit quality and establish better criteria for measuring it. Criteria that should be considered include the dollar amount of the error and the scope, depth, and documentation of the audit.

--Revise the quality control system so that all errors in terms of the established measurement criteria are recorded and reported, quality problems are identified, and trends are developed.

--Conduct a controlled study to determine how long it takes to do a quality audit and use the study results, along with other factors, in planning the annual number of audits to be done.

These and other recommendations are discussed in detail on pages 32 through 35 and 53 and 54.

In commenting on the report, the Commissioner of Internal Revenue generally agreed with GAO's recommendations and indicated that IRS would take corrective action. The Commissioner's response, however, concerns GAO because it is not clear that the Service is committed to deal with the audit quality problem as a priority matter.

IRS apparently is not convinced of the seriousness of the problem because many of the errors for which GAO was able to quantify the dollar impact did not exceed IRS' dollar tolerances and would not have been corrected under current IRS practice. Thus, under IRS' approach, an error is not an error unless it exceeds a certain dollar amount.

That philosophy is contrary to one of GAO's basic points--an effective quality control system must detect and report all errors, regardless of whether the errors are subsequently corrected. The purpose of an effective quality control system is to measure quality overall and provide the information needed to formulate corrective action plans. A system that overlooks certain errors defeats that purpose.

The Commissioner's response also tended to downplay the significance of GAO's findings by equating audit quality problems to the $1 million in errors GAO was able to quantify. GAO's overall assessment that audit quality should be improved was not based solely on the $1 million. Of equal, if not greater, concern were the findings that (1) significant items were missed in 58 percent of the review cases and 47 percent of the conference cases, and (2) the adequacy of many audits could not be determined from the records.

IRS' comments are included as appendix I to the report. GAO's evaluations of those comments are on pp. 35 and 54.

Contents

ABBREVIATIONS

GAO	General Accounting Office
IRS	Internal Revenue Service

CHAPTER 1

INTRODUCTION

The Joint Committee on Taxation asked us to review Internal Revenue Service (IRS) procedures and controls for insuring audit quality as they relate to individual income tax returns with emphasis on IRS' quality control system. Our objectives were twofold--to evaluate the quality of audits being done and to evaluate the quality control system's effectiveness in identifying and correcting less than quality audits.

WHY IS AUDIT QUALITY IMPORTANT?

IRS' overall mission is to encourage the highest possible degree of voluntary compliance with the tax laws and to maintain the highest possible degree of public confidence in the Service's integrity and efficiency. To do this, IRS communicates the requirements of the law to the public, determines the extent and causes of noncompliance, and enforces the law. Its enforcement activities include auditing returns, collecting delinquent taxes, and recommending prosecution of persons who evade their tax responsibilities. Of all those activities, IRS considers the audit of returns to be the greatest stimulus to voluntary compliance.

The following statistics, relating to audits of individual income tax returns by IRS' 58 district offices, further indicate the significance of IRS' audit effort.

Fiscal year	Number of individual income tax returns audited	Amount of additional tax and penalties recommended as a result of audits (in billions) (note a)
1976	2,022,028	$1.6
1977	1,742,056	1.6
1978	1,845,242	1.8
Total	5,609,326	$5.0

a/These figures represent net audit adjustments proposed by IRS' Examination Division. They do not necessarily represent the net amount that IRS billed the taxpayer or the net amount that IRS will eventually collect.

Faced with limited audit resources, IRS must decide how best to use those resources to achieve its objectives. A basic question in that respect is whether audits of more taxpayers would enhance its mission or whether fewer but better audits would be a better approach--that is, how many audits at what level of quality? It is a difficult question to answer.

The Commissioner of Internal Revenue addressed that question during the appropriation hearings for fiscal year 1975 as follows:

> "* * * the overriding priority of the Internal Revenue Service is to carry out its responsibilities in such a way as to assure the taxpayers that the tax laws are firmly and equitably enforced. Clearly, the Service, by itself, cannot make certain that every taxpayer pays precisely the amount in taxes that he or she owes. We can give the taxpayers a reasonable assurance that the tax system does work: that others are paying their fair shares and those who attempt to pay less than their fair shares are likely to be called to task.
>
> * * *
>
> "In audit one can play a numbers game at the sacrifice of quality, at the sacrifice of doing a comprehensive job, and a comprehensive job that is fair both to the taxpayer and to the Government * * * We propose to maintain the quality of our audits. * * * we propose at the same time to increase the numbers of audits. This increase * * * is necessary to give assurance to the people that do pay that others will."

As the Commissioner acknowledged, audit quality is vital. If IRS performs less than quality audits, tax revenues may be lost, some taxpayers may pay more or less than they really owe, and future voluntary compliance may suffer. In that same respect, taxpayers have a right to expect that IRS examiners understand the tax laws and apply them consistently.

IRS recognizes the importance of audit quality and has a quality control system that produces limited data for monitoring the level of quality being attained. That system includes IRS' Internal Audit Division and various review components at the national, regional, and district office levels. The nucleus of that system is the postaudit review process. This process encompasses audit case review by examiners' supervisors (known as group managers), the district office technical review staffs, and District Conference--until recently, the first step in IRS' administrative appeals process. 1/

The accuracy and completeness of data provided by the quality control system is vital if IRS is to properly consider audit quality when planning the number of audits it wants to do.

1/In the interest of efficiency and economy, IRS revised the appeals process, effective October 2, 1978, by transferring the District Conference function from the Examination Division to the Appellate Division. This reorganization consolidated the appeals function in the Appellate Division and eliminated one level of appeal formerly available to the taxpayer.

CHAPTER 2

IRS NEEDS TO IMPROVE AUDIT QUALITY

IRS' audit quality needs to be improved. Too frequently, IRS tax auditors (1) make technical or computation errors, (2) overlook significant audit issues, (3) either do insufficient work to properly decide whether the item under examination is allowable or fail to document their audit effort, and (4) fail to inquire about unreported income.

These problems can be traced to several factors. The lack of adequate measurement criteria in IRS' definition of audit quality is one. Other factors we identified all seemed to stem from pressure exerted by management, whether real or perceived, to complete audits as quickly as possible.

IRS was not aware that it had audit quality problems because its quality control system failed to bring the problems to management's attention. Responsible IRS officials characterized IRS audit quality as good or better. Their perceptions go a long way toward explaining why action was not taken sooner to improve quality. IRS officials have become increasingly aware that audit quality is not as good as they thought and have begun to improve it.

IRS' CRITERIA FOR MEASURING AUDIT QUALITY IS INADEQUATE

IRS' definition of audit quality does not provide adequate measurement criteria. Faced with that, we developed our own criteria with input from IRS.

The IRS manual for tax auditors states that,

> "A quality audit reflects the development of all significant issues, both factually and technically, a proper determination of the tax liability, the preparation of a complete and an accurate report, and the promotion of improved relations in all taxpayer contacts."

This definition lacks usable audit quality measurement criteria.

Although it is difficult to define quality in quantitative terms, there are certain basic elements which should be present in and are indicative of the quality of any audit.

The extent to which an audit contained these elements generally can be measured. That is how we evaluated IRS' audit quality.

Using existing IRS guidelines with input from various IRS personnel, we developed a list of basic elements that an examiner must satisfy during each audit if he expects to do a quality job. Specifically, the examiner must

(1) examine all significant items,

(2) make technically correct conclusions,

(3) explore all significant items to an adequate depth,

(4) document the examination adequately, and

(5) probe for additional income and expenses.

IRS examination personnel from the national, regional, and district offices generally agreed that these elements were important in performing a quality audit. Some, according to district personnel, are more important than others. The elements as listed above are ranked according to the degree of importance placed on them by district personnel--from most important to least important.

We used these elements to evaluate IRS audit quality by reviewing a statistical sample of tax returns filed in five selected districts by individuals claiming itemized deductions but not business or farm income (hereinafter referred to as nonbusiness returns). All of the sampled returns had been audited by tax auditors through IRS' office audit program--that is, the audit was done by correspondence or through an interview conducted in an IRS office. Office audits of nonbusiness returns accounted for 65 percent of all individual returns audited by IRS during fiscal year 1978.

Our total sample of 490 tax returns consisted of 250 cases selected from 711 in which the taxpayer had disagreed with the auditor's findings and 240 cases selected from 824 in which the taxpayer had agreed. All the sampled cases had passed through IRS' postaudit review process. The unagreed cases had been reviewed by District Conference; the agreed cases had been reviewed by the districts' technical review staffs.

In measuring the quality of a given audit, we attempted to determine whether each person in the examination process had done his job. That is:

--Had the classifier 1/, in selecting that return for audit, identified all significant audit issues on the return?

--Had the examiner done sufficient work to be reasonably assured that the taxpayer correctly reported his income, claimed those exemptions, deductions, and credits to which he was entitled, and was entitled to those exemptions, deductions, and credits he claimed?

--Had group managers, reviewers, or district conferees detected and appropriately documented the problems we noted? If they had, we did not count such items as errors.

In essence then, our results measure the extent and, where possible, the dollar magnitude of those errors missed by IRS' postaudit review process.

When an error was detected, we computed the dollar impact whenever possible no matter how small it was. IRS does not require examiners to correct small dollar errors detected during the postaudit review process if the errors are in the taxpayer's favor. We still included small dollar errors in our statistics because an accurate measure of quality requires that errors be considered, regardless of whether they are corrected.

We discussed with district officials each error for which we were able to quantify the tax impact and each unaudited item which we considered significant. When they disagreed, we considered their reasons in deciding whether to adjust our statistics. Their opinions on our findings varied widely--from almost complete agreement to mostly disagreement--depending on the type of error and the type of case being discussed. The results of these discussions are presented in more detail in subsequent sections of this chapter.

1/Classifiers are experienced examiners who are temporarily assigned to screen returns to determine whether or not they warrant audit.

IRS THOUGHT ITS AUDIT QUALITY WAS GOOD

An important step in assessing quality is to ask responsible IRS personnel how good they think their audits are. Their perceptions can go a long way toward explaining why problem areas seem to go unnoticed. We sought such an assessment from IRS personnel at all levels of responsibility. Their general reaction: audit quality is good.

IRS management as a group described audit quality as "good" or better. The Director, Examination Division, said audit quality was within reasonable bounds, and overall he did not feel that quality was a significant problem. Regional commissioners and assistant regional commmissioners for examination in the regions we visited rated quality as adequate, good, or very good. District directors and examination division officials generally rated audit quality as above average or high.

Tax auditors and group managers also rated audit quality as good. From questionnaire responses, we estimate that 68 percent of IRS group managers and 69 percent of its tax auditors in the five sampled districts rated audit quality as high or very high. 1/

District reviewers, those responsible for assessing audit quality, were not as confident about that quality as were other IRS personnel. Their comments, however, indicated that it was acceptable. Twenty-seven percent of IRS reviewers in the five sampled districts rated audit quality as high or very high, while 60 percent rated it as average.

AUDIT QUALITY IS NOT AS GOOD AS IRS THOUGHT

Our evaluation of the case files associated with 490 audits using the measurement criteria we had developed indicated that audit quality was not as good as IRS thought. Too frequently, IRS tax auditors determined an incorrect tax because of technical errors, computation errors, or failure to make automatic adjustments; overlooked significant audit issues; either did insufficient work in examining the items they did audit or inadequately documented their audit effort; and failed to adequately probe for unreported

---

1/For reporting purposes, questionnaire results are expressed in terms of our best estimate of what all tax auditors and group managers in the five districts would say. See appendix II.

income. In addition, tax auditors usually did not probe for unclaimed deductions and credits but IRS does not require them to do so.

None of the specific errors we identified were identified by IRS' postaudit review process even though each case had been reviewed at least once during that process.

Tax auditors are not always determining the correct tax liability

IRS guidelines state that tax auditors are responsible for determining the correct tax liability. This responsibility would certainly require examiners to arrive at technically accurate and computationally correct conclusions. This is too frequently not the case, however. In 34 percent of the review cases and 45 percent of the conference cases in the 5 districts, the examiner made one or more technical or computation errors or failed to make automatic adjustments totaling $37,000. 1/ This includes errors in the Government's favor of about $11,000 and errors in the taxpayer's favor of about $26,000.

During the period we selected cases, those five districts examined many other returns containing the same characteristics as our sampled cases except that they were not reviewed by District Conference or the technical review staff. Projecting our observed dollar errors to these unreviewed cases, we estimate that the errors totaled $1.0 million.

Technical errors

Technical errors generally involved examiner failure to allow or disallow items clearly defined by law, such as the nondeductible portion of a casualty loss ($100) or medical expenses (up to 1 percent and 3 percent of adjusted gross income).

1/Throughout the report, sample results from the 240 review cases and 250 conference cases were projected to the 824 review cases and 711 conference cases processed by the 5 districts during the period we selected cases unless otherwise specified. Because of the sampling plan and weighting system used, it would be inaccurate to add the percentages for the two types of cases and compute an average for all cases. Therefore, percentages are shown for each type of case throughout the report. Appendix II contains a description of how we projected sample results.

Other technical errors involved situations where the examiner or conferee considered the item but improperly applied the law. Examples include the examiner computing the audit adjustment using the wrong tax rates, or allowing a full year's depreciation on capital assets owned for only part of the year.

On the basis of information on the tax returns and in the audit case files, and after considering district officials' views on each of our findings, we determined that examiners made one or more technical errors in 21 percent of the review and in 16 percent of the conference cases in the five selected districts. The errors, none of which were detected by the postaudit review process, ranged from a $770 error in the taxpayer's favor to a $728 error in the Government's favor. The tax impact of the errors was distributed as follows.

	Technical errors			
	Review cases		Conference cases	
Tax impact	Number of cases	Percent	Number of cases	Percent
Government's favor				
$251 or more	4	.5	2	.3
$101 to $250	7	.8	3	.4
$51 to $100	5	.6	0	0
$1 to $50	9	1.1	15	2.1
	25	3.0	20	2.8
Taxpayer's favor				
$251 or more	3	.4	4	.6
$101 to $250	48	5.8	2	.3
$51 to $100	33	4.0	48	6.8
$1 to $50	65	7.9	37	5.2
	149	18.1	91	12.9
No technical errors found	650	78.9	600	84.4
Total	824	100.0	711	100.1

District officials agreed with 84 percent of the technical errors we identified in the review cases and 75 percent of those we identified in the conference cases. In some cases the officials disagreed because they considered the tax impact insignificant; in other cases, they disagreed because the item in question had not been identified for audit by the classifier; and in still other cases, they disagreed because they assumed the examiner had done something even though the case file contained no evidence to support that assumption.

In deciding to categorize something as an error despite IRS' disagreement, we took the position that (1) all errors should be considered in assessing audit quality no matter what their dollar amount, (2) the fact that an error was due to classifier oversight made it no less an error, and (3) an accurate assessment of audit quality cannot be based on assumption.

We analyzed the technical errors to determine if there were any patterns or relationships between errors and issues. We saw none; that is, the errors were not being made on the same issue. (See app. III for a list of the various issues involved.)

Computation errors

Computation errors include situations where the examiner or conferee (1) did not do such things as add or subtract correctly, or (2) used an incorrect amount when computing the tax impact of proposed adjustments.

One or more computation errors were made in 11 percent of the review cases and 27 percent of the conference cases. As with technical errors, IRS' postaudit review process did not detect them, and most errors favored the taxpayer. Although computation errors ranged from $452 in the taxpayer's favor to $454 in the Government's favor, most had a smaller tax impact--$50 or less.

Tax impact	Computation errors			
	Review cases		Conference cases	
	Number of cases	Percent	Number of cases	Percent
Government's favor				
$101 or more	3	.4	15	2.1
$51 to $100	5	.6	13	1.8
$1 to $50	23	2.8	72	10.1
	31	3.8	100	14.0
Taxpayer's favor				
$101 or more	14	1.7	4	.6
$51 to $100	1	.1	12	1.7
$1 to $50	45	5.5	75	10.5
	60	7.3	91	12.8
No computation errors found	733	89.0	520	73.1
Total	824	100.1	711	99.9

Examples of computation errors include:

--The examiner calculated that the taxpayer's rental expenses were overstated by $594, but failed to consider this adjustment in computing the additional tax owed. The tax impact was $218 in the taxpayer's favor.

--The examiner incorrectly added the totals from the documents provided by the taxpayer to support the amount claimed for interest expense, and gave the taxpayer credit for $110 more than had been paid. The tax impact was $43 in the taxpayer's favor.

--The conferee incorrectly used a 25-percent factor instead of 30 percent in calculating the business portion of the taxpayer's total mileage. The tax impact was $169 in the Government's favor.

IRS officials agreed with all of the computation errors that we identified on review cases and with 92 percent on conference cases. They disagreed with the remaining 8 percent primarily because they considered the amounts involved insignificant. As discussed previously, we did not consider "insignificant amount" as a valid reason for deciding whether an error had been made. Therefore, we retained these errors in our statistics.

Automatic adjustments not made

During an audit, the examiner may make an adjustment that affects other items on the return. To arrive at the correct tax liability, the examiner should automatically adjust those related items. For example, if the examiner has revised the taxpayer's adjusted gross income, he should automatically revise the amount of any deduction, such as medical expense, that is based on the amount of adjusted gross income. Examiners failed to make such adjustments in 9 percent of the review cases and 13 percent of the conference cases examined in the 5 districts. None of these errors were detected by IRS reviewers.

Tax impact	Automatic adjustments not made — Review cases: Number of cases	Review cases: Percent	Conference cases: Number of cases	Conference cases: Percent
Government's favor				
$51 or more	0		1	.1
$ 1 to $50	29	[illegible]	38	5.3
	29	3.[illegible]	39	5.4
Taxpayer's favor				
$ 51 or more	1	.1	8	1.1
$ 1 to $50	44	[illegible]	42	5.9
	45	[illegible].4	50	7.0
No missed automatic adjustments found	750	9[illegible].0	622	87.5
Total	824	99.[illegible]	711	99.9

Following are examples of instances in which IRS personnel failed to make automatic adjustments.

--The examiner made several adjustments that caused the amount of a taxpayer's itemized deductions to fall below the allowable standard deduction. But the examiner failed to use the standard deduction in recomputing the tax liability. The tax impact was $204 in the Government's favor.

--The examiner failed to adjust medical expenses after a change to adjusted gross income. The tax impact was $16 in the taxpayer's favor.

We found no discernible pattern of tax issues where examiners were most likely to not make automatic adjustments.

Tax auditors are not pursuing all significant audit issues

IRS district personnel generally cited the examination of significant items as the most important element of a quality audit. Despite that recognized importance, IRS agreed that at least one significant item was overlooked in several of the audits we reviewed.

The basis for the criteria we used to decide whether all significant items were examined was information obtained from IRS personnel and IRS guidelines. According to IRS personnel, a "significant item" is one that seems questionable in relation to the return, may result in a substantial tax change, or has historically shown a high degree of noncompliance. IRS classification guidelines provide that "questionable" be considered in terms of the dollar amount in relation to other items on the return, the item's description on the return, and the item's very presence on or absence from the return.

In applying the criteria, we determined that IRS did not identify and examine all significant items in 58 percent of the review cases and 47 percent of the conference cases.

Number of items missed	Review cases	Conference cases
	-----(percent)-----	
None identified	42.1	53.0
1	40.2	23.2
2	8.8	14.4
3	4.9	5.6
4	2.6	2.7
5	.1	.5
6	0	0
7	1.3	.7
Total	100.0	100.1

As shown, IRS could have made better audits by extending audit coverage to just one more item.

In terms of items, district officials agreed with only 24 percent of the items we categorized as significant in the review cases and 34 percent of those we categorized as significant in the conference cases. In terms of cases, they also agreed that at least one significant item had been overlooked in 16 percent of the review cases and 20 percent of the conference cases we reviewed. When they agreed, they usually attributed the error to the fact that the classifier did not identify the item for audit.

District officials disagreed with our determinations primarily because they considered the dollar amount of the questionable deduction and/or its related tax impact insignificant. They disagreed other times because the classifier had not identified the item as needing audit (the same reason others gave for agreeing) or because they assumed something that was unsupported by the audit case file.

We discussed earlier our position on IRS arguments that center around the classifier or are based on assumption. (See p. 10.) We disagree with IRS' main argument that the amount involved or its tax impact was insignificant for two reasons. First, other factors besides dollar amount, such as compliance, have to be considered in assessing significance. Second, the significance of the tax impact should not be determined in isolation, but rather along with the total adjustment made. We agree that IRS should not make a proposed adjustment unless the total is significant. However, IRS had already proposed an adjustment to the tax liability in 53 and 50 percent of the review and conference cases, respectively, that contained a missed significant item.

Examples of missed items we considered significant despite IRS' disagreement follow.

1. IRS disagreed with us that a $540 casualty loss claimed by the taxpayer was significant and should have been audited. Considering factors in addition to dollar amount, such as noncompliance and IRS' objective of fairly and equitably enforcing the tax laws, we decided to continue categorizing this item as significant. Casualty loss has historically been an area of high noncompliance. In that respect, classifiers usually select a casualty loss for audit whenever it appears on the return (69 percent of the time in our sample). Also, IRS examined casualty losses of $122, $125, and $150 on other returns we reviewed.

2. The classifier identified only interest expense amounting to $1,220 on the farm rental schedule as warranting examination. IRS disagreed that repairs of $734, automobile expenses of $706,

and insurance of $530 also shown on the farm rental schedule were significant enough to warrant examination. We did not change our position because of other information contained on the return and developed during IRS' audit. The case file showed, for example, that the taxpayer and his spouse worked for wages as opposed to being self-employed, resided on the rental property, and claimed as farm equipment only one-half of a fully-depreciated automobile. Given this, we thought the claimed repairs and automobile expenses were significant enough to warrant examination. We considered the claimed insurance expense to be significant because the case file showed that the taxpayer had no insurance on a farm building which was destroyed by fire during the tax year.

3. A salaried employee claimed a $154 entertainment expense. IRS instructions to taxpayers state that a salaried employee ordinarily may not deduct entertainment expenses. IRS disagreed, however, that a significant item had been overlooked in this instance because the tax impact ($42 if totally disallowed) was insignificant. Our decision to categorize this item as significant, despite IRS' disagreement, was based on the fact that the item was obviously questionable--one aspect of our criteria for assessing significance. The taxpayer paid an additional $602 in taxes and interest as a result of the audit.

Admittedly, any attempt to assess significance requires some judgment and can be expected to result in some disagreement, not only between us and IRS but also among IRS personnel. However, even if we were to limit our statistics to those instances in which district officials agreed with our determination, they would still demonstrate that quality needs to be improved--IRS is too often overlooking significant audit issues.

We analyzed the significant items missed for relationships or discernible trends. We found none. (See app. IV for a list of the affected issues.)

Tax auditors either are not auditing issues in sufficient depth or are not adequately documenting their efforts

IRS personnel generally agree that depth and documentation are both important elements of a quality audit. An examiner must do sufficient audit work to satisfy himself that the item in question has been properly reported and needs to document what he did so that his conclusions can be properly evaluated if reviewed by others. Our review of examiners' efforts, as documented in the audit case files, showed that IRS has a problem--many of its examiners are either not doing sufficient work, or not adequately documenting the work they do, or both.

Audit depth

In assessing audit depth, we considered two questions: did the examiner apply all appropriate techniques in auditing the item in question and, for those techniques he did apply, did he go far enough?

IRS audit guidelines cite specific steps an examiner should take or questions he should ask in examining an issue. These steps and questions are referred to as audit techniques. For example, the guidelines state that when auditing interest expense, the examiner should verify the amount claimed to instruments of indebtedness, assure that the expense is being claimed for the proper year, establish that payments of principal have not been included as interest, and ascertain that the obligation is the taxpayer's. The following table summarizes what the examiners' workpapers showed was done to verify interest expense for those review cases in the five districts in which interest was an audit issue.

Technique	Technique applied	No documentation that technique was applied
	(percent)	(percent)
Verify amount claimed	100	0
Taken in proper year	19	81
Payments of principal included	14	86
Valid debt of taxpayer	6	94

Techniques other than verifying the amount claimed are equally important in determining whether the interest deduction is allowable. If the taxpayer fails to meet any of the other conditions, the deduction is not allowable in the tax year under audit even if the taxpayer can prove that he paid the claimed amount. However, IRS instructs its auditors to not make adjustments between tax years if the overall revenue effect is not significant.

Overall, however, most of IRS' audit effort in the five districts, as documented in the case files, was spent verifying dollar amounts--50 percent of the audit effort expended on review cases and 49 percent of the effort expended on conference cases. Some districts, however, made more in-depth audits than other districts. In the district making the most in-depth audits, only 26 percent of the audit effort on review cases was spent verifying dollar amounts as compared with 57 percent in the district making the least in-depth audits We also noted significant differences among districts on conference cases

Whenever the case file indicated that the examiner had applied a specific technique, we determined whether he had

--obtained documentation that supported all of the item under audit,

--obtained documentation that supported part, but not all, of the item under audit, or

--accepted the taxpayer's oral testimony without additional documentation.

The following table shows that, generally, when examiners did apply a technique they completely verified the matter in question.

Depth of verification	Review cases	Conference cases
	(percent)	
Verified 100 percent	59	64
Verified a portion	5	6
Accepted oral testimony	26	26
Unknown	10	4
Total	100	100

However, a large portion of the overall verification effort involved oral testimony and some districts accepted this evidence more than others.

Some oral testimony resulted from examiners questioning taxpayers about such matters as dependents, marital status, or unreported income. We considered oral testimony sufficient in these situations unless the case file contained information indicating that further audit work was warranted.

We did not, however, consider oral testimony sufficient when the issue warranted documentary support. Of particular concern was examiner acceptance of oral testimony in support of claimed deductions. Auditing is a process whereby the accuracy of a reported figure is usually determined by checking the supporting documentation. Accepting oral testimony when documentation is in order provides little assurance of accuracy since the taxpayer, in effect, said the same thing when he filed his return. An examiner's acceptance of oral testimony when documentation is required can also lead to inconsistent taxpayer treatment. For example, in some of our sampled cases, examiners accepted oral testimony as support for claimed employee business expenses. In other cases, examiners rejected oral testimony and disallowed the employee business expenses claimed.

Audit documentation

Our evaluation of audit depth indicated that many tax auditors were doing insufficient audit work. That assessment could be misleading, however, because it is based on available case file documentation. In reality, examiners might be doing in-depth audits but not recording their efforts.

IRS guidelines require that examiners include in the case file those papers containing evidence of what was done during the audit while cautioning that, in the absence of sufficient documentation, the quality of the audit may justifiably be subject to criticism. The guidelines also require that examiners obtain adequate supporting data to afford a reasonable basis for considering the taxpayer's position on each issue. Examiners have not always followed these guidelines.

Using the IRS manual, we identified the specific audit techniques examiners could have used in each of our sampled audits. We then evaluated the examiner's documentation in each of the case files associated with our sample to determine if we could follow what the examiner had done. We were able to follow what the examiner did on only 28 percent and 32 percent of the applicable audit techniques for review and conference cases, respectively. The techniques that were documented, however, were generally documented well. When we limited our evaluation to only those techniques that had been documented, we could follow what was done in 90 percent and 97 percent of the review and conference cases, respectively.

We also gave 30 of our sample cases to reviewers in districts other than the district in which the audit had been done to obtain their views on case file documentation. The following table shows how they rated that documentation.

Rating	Percent of total
Very Good	3
Good	49
Average	12
Fair	19
Poor	17
Total	100

Our review also showed that the extent to which examiners documented the taxpayer's position depended on whether the taxpayer agreed or disagreed with the audit findings. In that regard, examiners documented the taxpayer's position in only 10 percent of the review cases in which the taxpayer agreed to all adjustments, as opposed to 23 percent of the conference cases in which the taxpayer disagreed with at least one of the examiner's adjustments.

IRS audit personnel generally contend that it is unnecessary to document the taxpayer's position if he agrees to an audit adjustment because, in such a case, the taxpayer's position is the same as IRS'. We disagree with this contention. Illustrative of the value of such documentation is the fact that in 14 percent of the conference cases in the five districts, the conferee, after considering the taxpayer's and examiner's positions, adjusted the audit results because, in the conferee's opinion, the examiner was not technically correct.

Whether or not the taxpayer agrees with the examiner's findings, it is essential that the examiner document his audit efforts and the taxpayer's position as to why he claimed more than was allowed. Without such evidence, IRS cannot fully evaluate audit quality.

Tax auditors are not probing sufficiently for additional income

IRS believes that unreported income amounts to billions of dollars annually, perhaps more than $24 billion. Accordingly, IRS procedures require examiners to probe for unreported income when interviewing taxpayers. Examiners often made no such probes or inadequately documented the probes they did make.

We estimate that examiners did not probe for unreported income in 62 percent of the review cases and 52 percent of the conference cases, even though certain returns contained indications of unreported income. For example:

--A taxpayer claimed and was allowed substantial costs associated with owning and operating an airplane used as a demonstrator but claimed no related income. The file showed no evidence that the examiner asked the taxpayer about income received from demonstrating the airplane.

--A taxpayer claimed expenses incurred in attending Marine Corps Reserve meetings but reported no income from being a Reservist. The file showed no evidence that the examiner inquired about income the taxpayer may have received for being a Reserve member.

--A taxpayer in the armed services claimed and was allowed moving expense but reported no related reimbursement. Military members are usually reimbursed for this expense; however, the file showed no evidence that the examiner had made inquiries in this regard.

Even when the audit case file showed that a probe was made for unreported income, the file did not contain sufficient information for management to assess the nature or adequacy of that probe. This is because IRS does not require examiners to document the nature of the probe, only that a probe was made. Accordingly, in 61 percent and 71 percent of the probes made on review and conference cases respectively, examiners only initialed or marked the "probe for income" line item on a procedure reminder checklist. Since this documentation is consistent with IRS guidelines, IRS should revise the guidelines to require examiners to better document the nature and extent of their verification effort.

Tax auditors are not required to probe for unclaimed deductions

The Internal Revenue Manual states that examiners have a responsibility to both the taxpayer and the Government, and that the examiners should ascertain that taxpayers have not overassessed their tax liabilities. However, IRS does not specifically require the tax auditor to probe for unclaimed deductions or credits to which the taxpayer is entitled.

Lacking specific instructions, examiners did not always probe for unclaimed deductions even though information on the return indicated that a probe should have been made. In our review of case files, we noted 23 returns containing possible unclaimed deductions warranting at least a probe. But none was made. For example, in 9 cases, the taxpayer claimed home mortgage interest expense but did not claim real estate taxes; in 10 cases, the taxpayer claimed medical expenses but no expense for medical mileage; and in 4 cases, the taxpayer's withholding statement showed State income tax withheld but the taxpayer had not claimed the related deduction.

In other cases, examiners failed to correct situations in which taxpayers overlooked allowable credits. For example, four taxpayers failed to claim credits for excess Federal Insurance Contributions Act tax withheld.

Given IRS' responsibility to determine the correct tax liability, IRS should require examiners to probe for unclaimed deductions and credits.

DIFFERENCES IN QUALITY AMONG DISTRICTS SHOW THAT BETTER AUDITS CAN AND SHOULD BE DONE

The level of audit quality being attained among the five selected districts varied significantly for some of the audit elements we used as measurement criteria. These differences show that IRS can improve its audit quality and make that quality more consistent among the districts.

We made statistical comparisons, where possible, of the quality being attained among the five districts for each of the basic audit elements we used as measurement criteria. We also compared each district's average dollar error per case for those cases where we could compute the tax impact. The purpose of the tests was to determine whether significant differences in quality existed among the districts. The statistically significant test results are shown in the following table.

Audit element	Review cases: Lowest district	Review cases: Highest district	Conference cases: Lowest district	Conference cases: Highest district
	(percent)			
Technical error	6	30	(a)	(a)
Computation error	(a)	(a)	12	36
Cases containing at least one error	12	38	(a)	(a)
Significant item missed	36	70	22	72
Unreported income probes not made	8	84	18	68
Audit depth (note b)	26	57	25	58

a/Test results did not show differences to be statistically significant.

b/Verification of amount claimed as a percent of all audit techniques applied. (See pp. 17 to 19.)

No one district attained consistently higher quality than the other districts on all of the tests showing significant results. For example:

--One district consistently made fewer technical errors than the other districts but had the lowest quality in terms of computation errors.

--Another district consistently missed more significant items than the other districts but did not have the lowest quality in terms of probing for unreported income.

--Another district consistently missed fewer significant items than the other districts but had neither the highest nor lowest quality on the other tests.

The highest quality shown for each of the tests in the table above may not be the best attainable, but IRS audit quality and the consistency of that quality would be significantly improved if all districts reached or exceeded those levels.

PRESSURE TO COMPLETE AUDITS AS QUICKLY AS POSSIBLE--A REASON FOR QUALITY PROBLEMS

Why is IRS' audit quality in need of improvement? There is no one answer. But one theme that ran through most of the causes we identified was time pressures exerted by management, whether real or perceived, to do audits as quickly as possible. These causes include:

--IRS procedures that tend to limit the number of items audited.

--A reluctance to raise questions on cases in which the taxpayer has already agreed with the examiner's findings.

--Group managers not instructed to perform in-depth reviews.

--Examiners not provided audit aids which would facilitate the audit and clearly set forth audit depth and documentation goals.

Classifiers tend to limit the number of items they identify for audit

In many instances, IRS district officials disagreed with our findings that certain significant items had not been audited by stating that the item had not been classified (selected for audit). Although that argument does not change the fact that the item was missed or in any way affect our evaluation of IRS audit quality, it is a valid observation when considering cause.

For the type of returns covered by our review, IRS uses classifiers to screen the returns manually and identify the specific items needing examination. The classifier indicates on a classification checklist those items he thinks warrant audit. The checklist generally sets the scope of the examination, because IRS uses it to notify the taxpayer about the audit and to tell him what items on the return he needs to support. The examiner generally does not see the return or the classification checklist until after the taxpayer has been notified. Thus, should the examiner note a significant item during the audit that had not been identified by the classifier, he must either disregard it or inconvenience the taxpayer by requesting additional records.

Classifiers in the five districts we visited identified an average of about three items per return, yet about half of the returns contained at least one significant item that had not been examined.

One reason classifiers do not always identify all significant items on a return may be because they incorrectly perceive limitations in local IRS guidelines. Four of the five districts where we did our work had given their classifiers written instructions to limit the number of identified items to three or four per return for the type of returns in our sample. The instructions did state, however, that more items could be classified if warranted, but in these instances, the cases were to be analyzed by the examiner before IRS contacted the taxpayer. The other district did not have any written instructions which mentioned specific numbers. Classifiers in that district said they consider such things as the time schedules that tax auditors work under and the significance of the items in determining the number of items to be identified.

Tax auditors generally do not audit unclassified items

IRS guidelines do not restrict tax auditors from expanding an office audit beyond those items identified by the classifier. The guidelines provide that the classification checklist is a guide and not intended to preclude examiner judgment. The guidelines further provide that, whenever possible, the tax auditor should consult with his supervisor before extending the scope of an examination beyond those items indicated on the classification checklist. This requirement gives management a degree of assurance that tax auditors will not expand the scope of an audit without good reason. The thrust of the requirement, however, could discourage auditors from expanding the audit to unclassified items even when they should.

The Examination Division Director said that the scope of an office audit is basically set by the classifier and that absent new information developed during the audit examiners would rarely go beyond that scope. District officials generally said that examiners would audit items that had not been identified by the classifier if they detected an unclassified significant item or if the taxpayer provided additional information during the audit which raised other questions.

Our review showed that tax auditors generally do not examine unclassified expense items. In only 18 percent of the review cases and 16 percent of the conference cases did the tax auditor examine such an item. This was not due to lack of opportunity. Our review showed that many returns contained at least one missed significant item.

Examiners may ignore unclassified items because they perceive a lack of authority to expand the audit's scope. We asked examiners what latitude they had to audit items other than those identified by the classifier. Most of them (52 percent) said they could expand the scope to include obvious unallowable items, and 39 percent said they could expand the scope to make an automatic adjustment. However, 41 percent of these examiners said they "never or almost never" audit issues that have not been classified while 59 percent said they "sometimes" do. These responses indicate that examiners either believe they have limited authority to expand the audit scope or are uncertain about the latitude they do have. Thus, IRS needs to clarify the guidelines.

Tax auditors feel pressured to complete audits

IRS annually develops an audit plan that includes expected examination rates. These rates are compiled from historical data on the number of audits done and the time it took to do them. When divided into the annual staff hours available, these rates show the average time IRS expects an audit to take. For those audits in which the nonbusiness taxpayer brings his records to an IRS office, the auditor is expected, on the average, to complete the audit and related paperwork in 1.8 to 4.2 hours depending on the complexity of the issues.

To measure the effect of time on audit quality, we asked examiners, group managers, and reviewers a series of questions dealing with the time allowed to do audits. As previously discussed, they rated audit quality as high or very high. Notwithstanding this rating, most tax auditors, some group managers, and most reviewers felt that audit quantity was overstressed to the detriment of quality. Their opinions in this regard are summarized in the following table.

	Group managers	Reviewers	Tax auditors
	(percent)		
Quantity overstressed to detriment of quality	39	58	70
Quantity stressed but quality usually does not suffer	30	21	19
Adequate balance between quantity and quality	20	17	9
Quality stressed but quantity usually does not suffer	11	4	2
Quality overstressed to detriment of quantity	0	0	0

Tax auditors (85 percent), group managers (71 percent), and reviewers (72 percent) said that tax auditors had less than or significantly less than adequate time to conduct examinations. When asked if a reduction in workload would increase audit quality, most tax auditors (86 percent), reviewers (81 percent), and group managers (72 percent) answered yes.

We gave those auditors who said that time was inadequate several alternatives and asked them to choose from those alternatives what they would do if they had more time. The following table shows the alternatives and their responses.

What auditor would do if he had more time	Percent
Probe more thoroughly for unreported income	10
Go into additional significant items	18
Expand the tests for items under audit (greater depth)	24
Research audit issues more thoroughly	29
Better document the case	12
Probe for additional expenses and credits	0
Other	7

The distribution of their answers shows where the auditors think improvement is needed. As discussed earlier, each of these areas could use more attention.

Contrary to the views of examiners, group managers, and reviewers, IRS regional and district officials generally said that examiners had enough time to do a quality audit. Four of five assistant regional commissioners for examination said that the annual plan did not put undue pressure on examiners. District officials generally said that the audit plan did pressure examiners to complete cases but that examiners still had enough time to do quality audits.

In our earlier report on how IRS selects individual returns for audit 1/, we expressed our concern about whether the examination rates used in developing the annual examination plan provided enough time for examiners to do a quality audit. We recommended that IRS conduct a controlled study to evaluate the reasonableness of those rates. IRS subsequently made a number of changes to its procedures for determining examination rates but did not conduct a controlled study. We still think that a controlled study is necessary, especially since most tax auditors believe that they do not have adequate time to conduct examinations.

1/"How the Internal Revenue Service Selects Individual Income Tax Returns For Audit" (Nov. 5, 1976, GGD-76-55).

IRS is reluctant to upset agreed cases

IRS is reluctant to raise questions about an audit after the taxpayer has agreed to the examiner's findings. This may result in erroneous tax adjustments going unquestioned. The reluctance may stem from IRS manual instructions which state that an agreed issue should not be reopened or a new issue raised by the reviewer unless the grounds for such action are substantial and the potential effect upon the tax liability is material.

We asked regional and district officials whether they believed group managers and reviewers were as likely to raise questions about an audit when the taxpayer agreed as they were when he disagreed. About 42 percent of the regional officials and 38 percent of the district officials answered no.

We asked those responding negatively why group managers and reviewers were less likely to raise questions on agreed cases. Generally, their answers are summarized as follows:

--Reopening a closed case might not be worth the effort.

--Any recontacts with the taxpayer might be considered harassment.

--The taxpayer might decide to change his mind about agreeing if the case is reopened.

--Agreed cases do not have questionable issues.

Most reviewers and group managers also expressed a reluctance to question agreed cases. When asked how often they would not return cases with errors because the audit or issue was agreed, 47 percent of the reviewers and 38 percent of the group managers said "sometimes," 10 percent of the reviewers and group managers said "generally," and 1 percent of the reviewers and group managers said "always." Only 17 percent of the reviewers and 15 percent of the group managers said case agreement was never a consideration.

Group manager reviews need to be more in-depth

A group manager is responsible for the performance and conduct of the 10 to 16 examiners in his group. District officials, reviewers, and group managers said that group managers are in a better position than reviewers to evaluate audit depth, documentation, probes for additional income and expenses, and overall audit quality. Also, most audit personnel (68 percent) who responded to our questionnaires said that the purpose of having group managers review an audit case is to ensure audit quality.

These reviews, however, did not always detect examiner errors. We attribute this, in part, to the guidance given group managers on reviewing cases. IRS guidelines require group managers to take a "quick look" through examination reports and case files to keep informed about the content of completed cases and determine the suitability of releasing those cases from the group. The guidelines also require group managers, when reviewing completed cases, to direct their attention to such areas as scope of examination, examination techniques, significance of issues, development of issues, and clarity and content of reports. However, the guidelines caution group managers to avoid intensive reviews which serve only to duplicate the work of other reviewers. In essence, IRS does not expect group managers to perform in-depth reviews.

No specific requirement exists in terms of the number of cases group managers should review, but they are required, when feasible, to be involved in every case in which the taxpayer disagreed with at least a portion of the examiner's finding. The case files associated with our sample audits showed that group managers actually reviewed 79 percent of the unagreed cases and 13 percent of the agreed cases. These reviews might have been sufficient in number to ensure audit quality if they had been performed in greater depth.

The need to improve group manager reviews is demonstrated by the high percentage of cases that still contained errors after their review. IRS conferees detected technical errors in 2 percent of the cases that group managers had reviewed. Our work showed an even higher percentage. For example, we found technical errors in 11 percent of the review cases and 18 percent of the conference cases that had been reviewed by a group manager.

Group managers caused tax impact changes in 17 percent and 5 percent of the review and conference cases, respectively, that they reviewed. Thus, audit quality would be enhanced if group manager reviews were performed in greater depth.

IRS does not provide tax auditors with necessary audit aids

IRS could help its examiners perform more in-depth audits, better document their audit efforts, and more adequately probe for unreported income by more clearly defining what it expects and providing necessary audit aids. IRS now leaves such matters to examiner judgment.

Pro forma schedules are a type of audit aid that can help examiners understand what is expected of them and provide them with effective vehicles for documenting their audit efforts. IRS has recognized the merits of pro forma schedules and uses them when auditing certain tax issues through correspondence with the taxpayer.

IRS, however, has done little to develop such schedules for audits done through interview even though these audits often address tax issues that lend themselves to a pro forma approach. For example, as discussed earlier, a casualty loss is usually selected for audit whenever it appears on a return (69 percent of the time for our sample returns). The IRS manual lists specific audit techniques for the examiner's use in auditing this issue--so specific that they could readily be incorporated into a pro forma schedule.

IRS IS STARTING TO IMPROVE AUDIT QUALITY

As a result of our work and feedback from other sources, IRS established a Quality Control Program Unit in September 1978. The unit's primary objective is to recommend changes to improve IRS audits through a continuing and detailed study of audit quality.

Among the actions already taken to improve audit quality, this unit has

--emphasized to all audit personnel, through examination program guidelines, the importance of audit quality;

--visited several regions to further emphasize quality; and

--written information notices concerning audit scope, inspection of returns related to the one under audit, and workpaper development.

In addition to establishing the Quality Control Program Unit, IRS has taken other steps to improve quality. These include:

--Mechanizing the tax computations resulting from audit adjustments to improve arithmetical accuracy.

--Requiring the examiner to document the reasons for not extending an audit to subsequent tax years.

--Revising the Tax Auditor Training Program to make it more effective.

These actions are a good start toward improving audit quality, but IRS could do more.

CONCLUSIONS

IRS needs to improve the quality of its individual income tax audits.

--We estimate that during our sample period IRS auditors in our selected districts assessed incorrect tax totaling $1.0 million because of technical errors, computation errors, or failure to make automatic adjustments.

--IRS personnel did not consider at least one significant item in several of the audits we evaluated.

--For those items that were audited, the auditors too frequently either did insufficient work or inadequately documented their audit effort.

--Examiner probes for additional income were not made in a significant number of cases.

Given these results, we concluded that IRS audit quality needs improvement. But how much improvement is needed? There is no simple answer.

If IRS performs less than quality audits, tax revenues may be lost, some taxpayers may pay more than they really owe, and future voluntary compliance may suffer. In that same vein, taxpayers have the right to expect that IRS auditors understand the tax laws and apply them consistently. On the other hand, IRS has limited audit resources. Thus, if it is to effectively use the audit function as a tool for increasing voluntary compliance, IRS must audit a sufficient number of returns to make its presence felt.

The mix of audit quantity and quality that IRS needs to best accomplish its mission is a decision that properly rests with IRS, not with us. IRS cannot properly address the issue, however, without dealing with the question of audit time. The theme that ran through most of the underlying causes for quality problems that we identified was time pressures, whether real or perceived. In this regard, there is a difference of opinion between management and other audit personnel regarding the time needed to do a quality audit. Audit personnel told us that the number of audits called for by the annual plan did not permit sufficient time for quality work. They also indicated that management emphasized quantity to the detriment of quality. Management, however, does not share this opinion. To our knowledge, IRS has no empirical data available with which to tell whether tax auditors have sufficient time to do a quality audit. Thus, such data should be developed.

As we see it, IRS should first define audit quality in measurable terms and then conduct a controlled study to determine how long it takes to do a quality audit. Once this is accomplished, IRS will know the relationship of audit quality to audit time, and, thus, will be better able to consider quality when planning audit quantity. Such a study will be difficult and cannot be accomplished overnight. There are, however, some other actions that IRS should take immediately to improve its audit quality.

Some of the quality problems we identified were caused by IRS procedures and practices for classifying and examining the type returns in our sample. Classifiers generally limit the number of items they identify for audit, and examiners generally limit their audits to those items identified by the classifier because they perceive limitations in the guidance received. Removing these limitations should reduce the number of significant items that are not being examined, provided that classifiers and examiners have sufficient time to do quality work.

Another problem is the lack of assurance that audits are done consistently and comprehensively. IRS could help its auditors perform more in-depth audits, better document their audit efforts, and more adequately probe for unreported income and unclaimed deductions and credits by more clearly defining what is expected and providing audit aids such as pro forma schedules. Examiner use of such aids would also facilitate the reviews of completed audits by both the group manager and the technical review staff.

IRS could also improve audit quality and enhance its assurance of audit consistency and comprehensiveness through more effective group manager reviews. Group managers can have a significant impact on audit quality because of their closeness to the auditors and the audit work. This impact could be increased if IRS provided group managers with standards to follow in evaluating audit quality and in documenting their work. Such standards would also provide IRS with a basis for assessing the adequacy of group manager reviews.

IRS also needs to treat agreed and unagreed audit cases consistently. Quality problems presently can go unnoticed because of a reluctance on the part of IRS personnel to raise questions on cases in which the taxpayer has agreed to the examiner's findings. Unless this reluctance is overcome, quality problems will go unnoticed and IRS management will not receive a true measure of the level of quality being attained. Thus, examiners and reviewers should be required to apply the same standards for both agreed and unagreed cases.

From our work and other sources, IRS has become aware of the need to improve its audit quality, has taken some corrective actions, and is planning others. IRS should incorporate the following recommendations in those plans.

RECOMMENDATIONS TO THE COMMISSIONER OF INTERNAL REVENUE

We recommend that IRS:

--Define audit quality in measurable terms so that examiners know what is expected of them.

--Do a controlled study to determine how long it takes to do a quality audit and use the study results, along with other factors, in planning the annual number of audits to be done.

--Review district classification instructions and require the districts to rescind procedures that may give classifiers the impression that they are to limit the number of items that can be identified for audit.

--Clarify that tax auditors have authority to expand the audit to significant items not identified by the classifier.

--Require examiners to probe for unclaimed deductions and credits.

--Devise pro forma audit aids for those tax issues most often examined to assist examiners in doing consistent and comprehensive audits and in documenting their work so that reviewers can better evaluate audit quality. Pro forma aids should also be devised and used in probing for unreported income, deductions, and credits.

--Establish standards for group managers to follow in evaluating audit quality and in documenting their review efforts.

--Require tax auditors to document and reviewers to include in their review the taxpayer's reasons for claiming more than the tax auditor allowed, even though the taxpayer agreed to the adjustment of the issue.

IRS COMMENTS AND OUR EVALUATION

In a June 20, 1979, letter, the Commissioner of Internal Revenue generally agreed with our recommendations. IRS did not agree, however, that it should document the taxpayer's position on agreed cases. But at the same time, it said it would require that case files be adequately documented so that reviewers can determine whether or not the tax auditors findings were proper. We do not understand how a reviewer can make such an assessment unless he knows the taxpayer's side of the story.

We made this recommendation because conferees found that taxpayers, rather than tax auditors, were technically correct in 14 percent of the unagreed cases and because we do not believe that the fact that a taxpayer agreed with an auditor always means that the auditor was right. We are concerned that many taxpayers may not be sufficiently familiar with the tax laws to dispute an auditor's findings and end up agreeing to those findings even if they are wrong. Thus, we continue to see a need for the taxpayer's position to be developed during the audit and made a part of the record.

IRS took exception to our concern that tax auditors accepted oral testimony when the issue warranted documentary support. IRS pointed out that the acceptability of oral testimony was an exercise of sound judgment--a decision that must be considered in context of the entire audit--and that courts consider oral testimony in arriving at their decision.

We do not take issue with the concept of oral testimony. We recognize that tax auditor judgment must come into play and that reliance on oral testimony might sometimes be acceptable. We are concerned, however, about tax auditors relying on oral testimony as support for 26 percent of the audit techniques applied in both review and conference cases, the general absence of evidence that auditors' judgments in this regard were questioned by group managers or reviewers, and the fact that some districts rely on oral testimony more than others.

IRS also justified acceptance of oral testimony on the basis that such evidence was accepted by the courts. But in court proceedings, unlike audit proceedings, witnesses are under oath and subject to a perjury charge, subject to cross examination, and oral testimony is usually corroborated by documentary evidence or third party testimony.

Although agreeing to implement all our recommendations but one, IRS made several other comments that cause us to question its commitment to deal with the audit quality problem as a priority matter. IRS said that

--"* * * it is necessary to exercise management discretion in our operational planning in assessing the significance of corrections;"

--the perfect quality standard used by GAO would always reflect a degree of error in IRS' operations;

--GAO's method of measuring quality would not be a "practicable operational measure to achieve an appropriate balance between adequate coverage and effective use of resources, or between quantity and quality;" and

--GAO's projection of $1 million in tax-change errors represents only about 4 percent of the combined total of IRS changes, and GAO-identified changes include an undetermined amount of small tax change errors which IRS would not have corrected as a matter of practice.

Apparently IRS is not convinced of the seriousness of the problem because many of the errors we found did not exceed IRS' dollar tolerances and would not have been corrected under current IRS practice. Thus, under IRS' approach, an error is not an error unless it exceeds a certain dollar amount. IRS cannot allow tolerance levels established for management purposes to be the governing criteria when trying to answer the question--how good are its audits?

We are not suggesting that IRS alter its procedures for correcting small dollar errors. Our point is that an effective evaluation of audit quality requires that <u>all</u> errors, regardless of dollar amount, be detected and reported by the quality control system whether or not those errors are subsequently corrected.

We quantified the dollar impact of errors when we could because that information provides some indication of the level of quality attained and the cost to the Treasury and the taxpayer of reduced quality. IRS must know the number and type of errors being made and their related dollar impact to determine if corrective action is needed to improve audit quality overall and, if so, to devise the best strategy for applying that action. The information would also help IRS evaluate periodically the reasonableness of the tolerances it has established.

In making the 4-percent computation, IRS infers that our findings are insignificant. That is not the case. Our overall assessment that audit quality should be improved was not based solely on the $1 million in errors which we were able to quantify. Of equal, if not greater, concern were our findings that (1) significant items were missed in 58 percent of the review cases and 47 percent of the conference cases, and (2) the records show that tax auditors often did not do sufficient work even when they examined an issue. Neither we nor IRS know the dollar impact of those errors.

As we said earlier, the mix of audit quantity and quality that IRS needs to best accomplish its mission is a key decision that properly rests with IRS, not with us. Given our findings, however, IRS is not in a position to make an informed decision in this regard.

CHAPTER 3

IRS MUST IMPROVE ITS QUALITY CONTROL SYSTEM BEFORE IT CAN BEST DETERMINE HOW TO IMPROVE AUDIT QUALITY

As discussed in chapter 2, IRS has begun to improve audit quality. But why did IRS not move sooner? Because, in reality, management cannot be expected to correct a problem unless it knows that a problem exists. In that respect, IRS' quality control system failed to bring audit quality problems to management's attention and led management to believe that quality was not in need of improvement.

The quality control system contained design and procedural weaknesses that led to an inaccurate and inadequate measure or picture of the quality level being attained. We did not have sufficient resources to develop all the data the system should have provided. Thus, it is possible that there are additional causes to the ones we identified.

IRS has started to make changes to improve the quality measuring components of the system.

AN EFFECTIVE QUALITY CONTROL SYSTEM IS VITAL FOR IMPROVED AUDIT QUALITY

A properly designed and functioning quality control system provides management with the data it needs to identify problems, plan corrective action, and improve future performance. Without valid feedback from the system regarding the extent and types of quality problems, management may not know that it has a problem needing correction, and those making the errors may not know that they need to improve their performance.

IRS' quality control system includes internal audit and various review components at the national, regional, and district office levels. Although these various components all have a role in assessing audit quality, the focal point for measuring quality is the technical review function carried out at the district level.

A description of the various components and their involvement in measuring, assessing, or monitoring audit quality follows.

Internal Audit

IRS' Internal Audit Division is responsible for reviewing all IRS activities and preparing evaluative reports for management's use in correcting identified weaknesses. Internal Audit also provides national office program officials with periodic summaries highlighting its most significant audit findings.

Internal Audit's reviews of the examination function generally have been limited in scope and geographic coverage. 1/ It has looked at the managerial and procedural aspects of certain examination processes incrementally--classification, audit, or technical review--but not at the total spectrum of quality from case selection through postaudit review. Also, it has usually evaluated and reported on operations in only one district at a time. From January 1976 through January 1978, for example, Internal Audit issued 41 reports dealing with some aspect of audit quality in 46 districts. Eight of these reports pertained to the five districts we reviewed.

Some of the audit quality problems discussed by Internal Audit in the 41 reports included

--examiners inadequately probing for additional income,

--examiners not making automatic adjustments when needed,

--examiners insufficiently documenting case findings, and

--examiners making inconsistent adjustments.

Internal Audit also criticized reviewers for failing to issue feedback reports to examiners on underdeveloped cases.

In responding to Internal Audit's reports, management officials said that corrective action would be taken. However, we found similar problems in the five districts we reviewed. Management's corrective action generally

---

1/Our overall assessment of IRS' internal audit function is contained in another GAO report, "IRS Inspection Service Functions: Management Can Further Enhance Their Usefulness" (GGD-78-91, Jan. 30, 1979).

consisted of attempting to make responsible employees aware of the problems through increased management emphasis and discussions. Given our findings, this action was not sufficient to satisfactorily resolve the problems on a continuing basis throughout IRS.

National office review components

The national office's quality control efforts consist primarily of conducting regional performance evaluations under the National Office Review Program and reviewing statistical reports generated by IRS' management information system.

Under the National Office Review Program, national office analysts visit each region for 2 to 3 days every 21 months to evaluate how effectively regions are carrying out their primary role of supervising districts. The analysts focus on major program efforts and accomplishments and significant needs and problems. Between January 1976 and September 1978, the analysts prepared 11 reports under the program; 9 of the reports were on the 5 regions in our review.

The 11 reports contained little specific mention of audit quality. Essentially, the reports presented statistics on plan accomplishments (such as the number of audits performed and accomplishments as a percent of the annual examination plan) and reasons for accomplishments falling short of expectations. Some reports, however, did address the problem of examiners failing to document the depth of their audits and their probes for additional income and deductions. In another report, shortcomings in plan accomplishments were attributed to the "emphasis placed on quality audits."

Besides making site visits under the National Office Review Program, national office analysts make special purpose visits to assist in correcting audit quality problems. They also review audit quality error rate reports and other statistical reports on program results and plan accomplishments. These reports, which are prepared monthly, quarterly, or annually, provide statistical data on a national, regional, or district basis. National office analysts review this data, compare the data with previous or associated reports, and contact regional representatives to clarify any questionable items. This effort has some value from an audit quality standpoint since it permits an overview of error rates--IRS' only statistical measure of audit quality. Comparisons of these rates can provide an indication of the direction in which quality is moving. However, the rates do not identify reasons for quality problems.

Local versions of these reports are reviewed in a similar manner by regional officials and analysts and by district officials.

Regional office review components

Regional offices review district operations through two evaluative efforts--the Regional Evaluation Program and the Regional Case Management Program.

Under the Regional Evaluation Program, regional analysts review program accomplishments, evaluate current district systems for performing and processing examinations, and assess the performance of district management. The region provides district management with a written report summarizing the analysts' observations and in some cases containing recommendations for corrective action. We reviewed the reports issued from June 1975 through May 1978 and found little specific mention of audit quality.

Examples of findings from the reports include:

--The overall office audit program is effectively managed as evidenced by program accomplishments during fiscal year 1975.

--Although the district has excellent controls for their office audit program, improvement is needed in the area of controlling business examinations. More managerial involvement is needed in conducting workload reviews and assisting auditors in business examinations.

--The most significant finding is that the Audit Division management team is involved at all levels with its programs and personnel, has sufficient management controls for the identification of problem areas, and is giving proper direction to the resolution of those problem areas.

One reason these reports rarely mention audit quality directly may be because the Regional Evaluation Program guidelines, while being sufficiently broad to include an evaluation of management performance relative to audit quality, do not specifically require such an evaluation.

Under the Regional Case Management Program, regional analysts review completed audit case files to identify managerial, technical, and procedural areas requiring management attention. IRS guidelines require that as a minimum the analysts are to evaluate

--supervisory involvement,

--classification,

--audit planning,

--audit workpapers, and

--district review sampling procedures.

District office review components

District offices monitor, assess, and measure their audit quality primarily through the work of three groups--group managers, the technical review staff, and district conferees. Group managers review examiner performance by observing audits in process, holding discussions with the examiner concerning issues on specific cases, and reviewing certain completed cases. District conferees review and attempt to resolve cases where the taxpayer disagrees with the tax adjustment proposed by the examiner. The heart of the quality control system, however, is the technical review staff.

That staff, which is manned by experienced tax auditors and revenue agents, is responsible for reviewing certain audit case files and identifying quality problems. Information gathered from this process forms the basis for computing error rates--IRS' only statistical measure of the audit quality being attained.

The technical review staff in each district selects audit cases for review on the basis of national criteria which is sometimes supplemented by district management. All cases meeting certain criteria must be reviewed. These cases, called mandatory review cases, include those involving fraud or bankruptcy. All cases not meeting the mandatory review criteria are selected for review through random sampling and called sample review cases. Before October 1978, office audit cases in which the taxpayer disagreed with the examiner's findings and requested a district conference were sent directly to a district conferee and were not reviewed by the technical review staff unless the cases met mandatory review criteria. When IRS eliminated District Conference as part of its reorganization, effective October 1978, it started requiring mandatory review of office audit unagreed cases when the taxpayer asks for an appeal hearing.

The technical review staff's functions are to (1) determine that the tax laws were administered fairly and impartially during the audit process, (2) advise management of areas requiring their attention, and (3) make corrections wherever necessary. To accomplish these objectives, IRS guidelines provide that the review staff must continually strive for

--just, uniform, and impartial treatment of taxpayers while protecting the Government's interest;

--high quality standards in examination reports;

--constant awareness of unfavorable patterns or trends; and

--timely dissemination of information for use by others in evaluating the return classification, assignment, and examination activities.

Procedures require reviewers to consider the managerial, technical, and procedural aspects of a case and to prepare feedback reports whenever warranted--if the case was poorly developed, for example, or contained questionable tax adjustments. Reviewers use three types of feedback reports--observation, advisory, and inquiry--and generally issue only one report per case. Observation reports are used to record significant aspects of a case, such as exceptional examiner performance, which need to be brought to the attention of the examiner's branch chief. Advisory reports are used to inform the examiner and his group manager about errors that the reviewer found but considered not significant enough to correct. This information is provided so that the examiner will not continue to make the same errors. Inquiry reports are used to request additional information when the reviewer has questions about substantial technical, procedural, or mathematical matters.

The inquiry report, noting the reviewer's concerns, and the audit case file are sent to the examiner through his group manager. Should the inquiry report cause the examiner to adjust his original audit findings, the inquiry report becomes a correction report. Only correction reports are counted as errors when IRS computes error rates.

IRS collects information on each review case for management reporting purposes. This information includes

--case selection criteria (mandatory or sample);

--type of return (such as individual non-business with itemized deductions);

--type examiner (tax auditor or revenue agent);

--whether the tax impact of the correction favored the taxpayer or the Government;

--review time; and

--the type of review report (observation, advisory, inquiry or correction) issued, if any.

From this data, IRS prepares a monthly management report summarizing review results by each district, region, and nationally. The report contains separate tables that show technical review results by type of return for tax auditor and revenue agent sample cases and mandatory cases.

Statistics are presented for

--number of returns reviewed,

--total review time used,

--average review time per case,

--the number of reviewer reports issued in total and segregated by type,

--the number of correction reports causing increases and decreases in taxes, and

--error rates.

The error-rate data is supposed to alert management when quality has fallen below an acceptable level so that corrective action can be initiated. To assist management in interpreting error rates, IRS has established an error rate of 5 percent as the alarm level. This equates to 5 out of every 100 cases reviewed that required a change in the examiner's proposed assessment.

We reviewed IRS' reported error rate statistics for the eight consecutive quarters ending September 1978 for the type of cases in our sample. While the reports show both current month and fiscal year-to-date statistics, we concentrated on the fiscal year figures because many IRS personnel stated that the monthly error rates can be misleading due to the small number of cases reviewed in so short a period. In other words, one error in five cases reviewed would produce a 20-percent error rate.

As shown in the following table, audit quality in three out of the five selected districts and nationally was generally acceptable in terms of IRS' 5-percent alarm rate.

	Fiscal year-to-date error rate statistics					
	District					Total
Quarter ended	1	2	3	4	5	U.S.
December 1976	2.8	0.5	0	0	2.0	1.5
March 1977	3.7	0.2	0.8	0	3.9	1.4
June 1977	2.8	0.3	2.2	0	7.8	1.6
September 1977	3.5	0.5	2.7	0.6	7.5	1.8
December 1977	0.9	0.8	24.4	0	4.7	2.3
March 1978	2.4	0.4	19.5	1.3	5.3	2.3
June 1978	2.0	0.4	15.5	4.9	3.0	2.6
September 1978	2.0	0.8	15.4	4.4	3.2	2.7

Most regional and district officials generally viewed error-rate statistics as an acceptable tool for measuring audit quality. Many district officials, however, viewed the error rates as being somewhat inflated because of improper data entries into the system.

We did not attempt to measure the accuracy of the reported error rates in terms of what reviewers found and IRS' reporting criteria. Because the system does not record and report all errors, however, and because our case reviews showed that reviewers were not detecting all conditions that would count as errors, management is not receiving an accurate picture of audit quality.

Because of the importance of the technical review function in measuring and improving audit quality, we reviewed this function to ascertain whether it was providing an accurate measure of the level of quality being attained, and information to assist management in taking corrective action, when necessary.

We determined that the technical review function provided neither. We also evaluated the Regional Case Management Program in terms of these same objectives. It too fell short.

THE QUALITY CONTROL SYSTEM DOES NOT ACCURATELY OR ADEQUATELY MEASURE AUDIT QUALITY

The quality control system is capable of measuring quality and alerting management to quality problems. The system has not been effective, however, because

--reviewers have no standards for evaluating cases,

--reviewers do not report all quality problems noted,

--the system as designed does not provide sufficient data for managing audit quality, and

--IRS has not provided guidelines for its personnel to use in interpreting error rates.

As a result, management receives incorrect error rate data on the aspects of quality that are being measured and receives no statistical data on such important aspects of quality as the adequacy of audit depth and the extent to which all significant items were classified and examined.

Many quality defects go unnoted because reviewers have no standards for evaluating cases

The parameters of technical review are quite broad. The reviewer's role is to ascertain whether each IRS employee involved in the examination process managerially, technically, and procedurally did what he was supposed to do. But IRS has neither developed standards nor provided aids to lead reviewers through this evaluation process. Thus, each reviewer proceeds in whatever fashion he deems appropriate. As a result, examiner errors are often overlooked.

To test the effect of the lack of review standards, we submitted six of our sample cases from each of the five districts to a cross-district review. The 30 cases were each reviewed 5 times--once by the reviewer in the district that closed the case and once in each of the other 4 districts. We asked the reviewers to evaluate these cases just as they normally evaluate their own cases.

The results of this test indicated that management receives an incorrect picture of audit quality and that reviewers are inconsistent in their quality assessments. For example, the reviewer in one district had returned to the examiner none of the six cases selected from that district. Reviewers in the other districts, however, did not share this opinion. Each of them would have returned at least two of the cases to the examiners, and reviewers in one district would have returned all six cases.

The following table shows the number of cases returned by the original reviewer and the number of cases that subsequent reviewers would have returned.

	Number of cases by district (note a)				
	1	2	3	4	5
District 1 cases	$\boxed{1}$	6	2	1	3
District 2 cases	4	$\boxed{3}$	4	1	4
District 3 cases	3	4	$\boxed{3}$	2	2
District 4 cases	4	2	6	$\boxed{0}$	3
District 5 cases	3	2	5	0	$\boxed{0}$

a/Each case was reviewed only once by each of the other districts. The boxed numbers show the number of cases that were returned by the reviewer in the original district.

This data also illustrates inconsistency among districts. For example, districts 3 and 4 reviewed the same 30 cases, but their conclusions differed vastly. District 4 would have returned only 13 percent of the 30 cases while district 3 would have returned 67 percent.

Examples of unanswered questions that subsequent reviewers raised on the sample cases and considered sufficiently significant to warrant returning the case to the examiner include:

--What was the taxpayer's support for employee business expenses? Did the taxpayer furnish a statement from the employer regarding reimbursement policy and job requirements?

--Is the taxpayer claiming depreciation on assets which have already outlasted their useful life?

--Did the taxpayer meet the requirements for deducting the bad debt? What efforts were made to collect it?

--What are the actual years involved for the claimed education expense? Had the taxpayer met the minimum requirements? Was he qualifying for a new trade or business? Were any amounts and/or courses verified?

--Did the taxpayer receive a double deduction for alimony?

--What audit work was done on employee business expenses other than mileage? How was the percentage of mileage for business use determined?

The original reviewer for these cases had raised none of these questions.

Guidelines provided reviewers do not contain specific standards for evaluating the adequacy of audit scope, audit depth, or case file documentation although these are important elements of a quality audit. Without such standards, different reviewers can reach different conclusions about the quality of a given case. These differing conclusions can affect the accuracy of the tax assessed as well as give management an unreliable measure of audit quality. Although differences of opinion are understandable when persons are asked to evaluate quality, those differences could be narrowed if IRS provided adequate review standards.

Reviewers could also use standards or work aids to help them document their review efforts. Without adequate documentation, management cannot always determine what issues or items the reviewer considered or whether reviewers are evaluating cases consistently. Reviewer documentation is primarily limited to the comments they might make on any feedback report they issue, and even then the documentation is limited to the items questioned. Most reviewed cases, however, do not result in a reviewer's report. Consequently, management has no way of knowing what the reviewer looked at in concluding that the case's quality was acceptable.

Reviewers do not report all quality problems noted

Guidelines provide that reviewers will use feedback reports to communicate with managers and examiners. In that regard, the key to recording and reporting errors in the quality control system is the preparation of a reviewer's report. If a reviewer does not prepare a report on certain errors he detects, those errors will never enter the information system. We noted two types of situations in which reviewers handled questions or problems in ways other than through a reviewer's report.

Some reviewers telephone examiners to obtain additional information or correct a problem. We asked 22 reviewers in the 5 districts what they did when they had a question about a case. Most said they would first issue a reviewer's report and then discuss the case with the examiner. However, three said they would first discuss the case with the examiner and usually would not prepare a report. Such discussions expedite case reviews and provide examiners with feedback. Unless a reviewer's report is issued, however, the problem is not recorded in the information system and, thus, escapes management's attention.

Other reviewers note their concerns or questions in the case file as opposed to issuing a report. In that regard, 18 percent of the review cases contained such a note in lieu of a report. For example, one reviewer noted in the case file that a potential Individual Retirement Account issue existed since the taxpayer's records did not indicate participation in a qualified pension plan. However, the reviewer noted that because the item had not been identified by the classifier, it was not necessary for the examiner to audit it. Because the reviewer did not prepare a reviewer's report, the examiner and classifier were not informed that they had missed a significant item. In our opinion, if a reviewer is concerned enough about an issue to write a note, he should prepare a reviewer's report.

While telephone inquiries and reviewers' notes may expedite case closures, documenting all problems is essential if the quality control system is to serve management effectively.

The quality control system does not provide management with needed information

IRS' quality control system does not provide management with sufficient information to determine the level of audit quality being attained, identify problem areas which need attention, make informed decisions on corrective actions that should be taken, or adequately consider audit quality when preparing the annual examination plan.

The quality control system presently provides management with little more than the percentage of cases in which the reviewer found errors significant enough to warrant a change in the tax assessment proposed by the examiner. It does not provide such essential indicators of audit quality as the number of audits having depth and documentation problems, the number of audits in which significant items were missed, the dollar impact of errors, or the tax issues on which quality problems most frequently occur.

Recognizing the need for additional information on the quality of its audits, IRS issued instructions, effective September 12, 1978, requiring district review staffs to manually analyze the reviewer reports issued each quarter. The purpose of this analysis was to identify and report on error trends and patterns, such as classifier failure to identify items for audit, for management's use in evaluating classification and examination activities. The instructions did not, however, require that the information be accumulated and analyzed on a national basis. IRS needs to take this additional step to facilitate identification of national problems.

Even so, the analysis of reviewer reports is a positive step towards acquiring some of the information IRS needs to properly manage audit quality. Manual accumulation and analysis, however, may become quite burdensome and impact adversely on the number of reviews the review staff is able to perform. If so, IRS should consider expanding its mechanized system for recording and reporting error rates to include additional information from reviewer reports. In any event, this type of information is essential if IRS is to give proper attention to audit quality and its relationship to audit quantity.

IRS has not provided guidance for interpreting error rates

IRS established a 5-percent error rate as an alarm figure for each category of sample review cases. IRS guidelines provide that management should take necessary corrective action whenever the error rate exceeds the alarm figure, but the guidelines do not define "category." Consequently, some managers consider the alarm rate in terms of the composite error rates by type of examiner and type of case, such as tax auditor sample review cases; others break it down further by type of return, such as individual, or corporate.

In one of our selected districts, for example, the reported error rate for nonbusiness returns examined by tax auditors was 24 percent. The reported error rates for these returns when grouped according to taxpayer adjusted gross income ranged from 0 to 44 percent. Thus, if category was interpreted to be either all nonbusiness returns or one of the adjusted gross income groupings applicable to nonbusiness returns, the reported error rates should have prompted management action since the 5-percent error rate alarm had been clearly exceeded. On the other hand, if category was interpreted to mean all tax auditor-examined returns, including both business and nonbusiness, the reported error rate was 3 percent--well below the alarm level and, thus, no cause for concern.

Neither has IRS developed criteria for interpreting other statistical information contained on the monthly error rate reports. Besides error rate statistics, the monthly reports show totals for advisory, observation, and nontax change inquiry reviewer reports.

Because IRS has decided that management needs this information, it should provide criteria for uniform interpretation. Likewise, if IRS revises its system to provide such other data as the dollar impact of errors and the audit issues involved with errors, it should also provide uniform interpretative criteria. Otherwise, managers may reach varying conclusions when faced with the same statistical information.

The Regional Case Management Program lacks uniform review standards and a common method of case selection

One objective of the Regional Case Management Program is to assess the quality of selected audits. Although this program surfaces some quality problems, it could provide management with a more accurate and more reliable assessment if it were redesigned.

Presently, case selection and review standards can vary from region to region. Each region we visited independently decided how many and what types of cases (office audit and field audit) to review and how those cases were to be selected. For example, one region contacts personnel at the service center where the cases are stored and asks them to randomly pull 50 to 75 office audit cases. Another region allows the regional analysts to select cases any way they want. Neither the national nor the regional offices have prepared standards for analysts to use when reviewing cases. Some regions provide written suggestions, but each analyst is allowed to apply whatever standards he deems appropriate.

Lacking uniform review standards, analysts in certain regions place greater emphasis on procedural and managerial issues than audit quality. Analysts in other regions directed their reviews primarily toward audit quality.

As presently conducted, the Regional Case Management Program is not achieving its potential as a management tool for evaluating audit quality or the work of district technical review staffs.

IRS IS STARTING TO IMPROVE THE QUALITY CONTROL SYSTEM

As previously discussed, IRS established a Quality Control Program Unit in September 1978 to recommend changes for improving IRS audits through a continuing and detailed study of audit quality. To achieve that objective, the unit is studying the quality control system to determine if new objectives, techniques, and standards in that system would better ensure high quality audits. As part of its effort, the unit has sought the opinions of regional and district examination personnel.

At the time we completed our work, the unit had taken some steps to improve the quality control system. For example, it had revised the IRS manual to require that the Regional Case Management Program include certain technical, managerial, and procedural evaluations and that the resulting reports be sent to the national office. The unit had also approved the development of a training course for district review staffs. These actions should improve the quality control system, but more remains to be done.

CONCLUSIONS

IRS' quality control system did not accurately measure the level of quality being achieved and failed to adequately identify the nature and extent of quality problems. Thus, the system did not convey to management the need to improve the quality of IRS audits of individual income tax returns. IRS needs to revise the quality control system so that it records and reports an accurate picture of the quality being attained and the problems being experienced. Without such information, IRS does not have a sound basis for taking action to improve quality and cannot properly address the basic issue of audit quality versus quantity.

Before the quality control system can be relied on as an accurate indicator of the level of quality being achieved, IRS has to establish specific criteria for measuring quality--criteria such as we used to evaluate quality in chapter 2. Such criteria will enable reviewers to evaluate how well examiners did what they were supposed to do and will allow management to assess IRS performance in achieving the desired level of quality.

The quality control system now provides management with little more than error-rate data computed on the basis of the number of tax changes resulting from the review of completed audit cases. We do not argue the usefulness of this type data as one indicator of audit quality as long as the data is consistently interpreted. IRS, however, has no assurance that this is the case. Criteria should be established to assure that this data is interpreted uniformly throughout IRS. Similarly, IRS should develop criteria for interpreting the other indicators of quality produced by the revised system. IRS should also assure that its sampling plan for selecting review cases is adequate to assure statistical projection for each aspect of quality to be measured.

The quality control system also suffers from the lack of consistent and comprehensive quality reviews. IRS could help its reviewers do better work by providing standards that explain what should be reviewed and how the review effort should be documented. Those standards should be compatible with whatever criteria IRS establishes for measuring audit quality as discussed earlier.

The accuracy and completeness of the information produced by any system should be periodically tested. In this regard, the Regional Case Management Program should be revised to assure uniformity among regions and used to periodically assess the data produced by the quality control system.

IRS has taken some actions to improve the quality control system and plans others. It should incorporate the following recommendations in those plans.

RECOMMENDATIONS TO THE COMMISSIONER OF INTERNAL REVENUE

We recommend that IRS:

--Establish criteria for measuring audit quality. Criteria that should be considered include the dollar amount of the error and the scope, depth, and documentation of the audit.

--Revise the quality control system so that all types of errors in terms of the established measurement criteria are recorded and reported, quality problems are identified, and trends are developed.

--Establish standards for reviewers to follow in evaluating audit quality and in documenting their work so that the consistency and adequacy of their evaluations can be determined.

--Establish criteria to assure uniform interpretation of the error-rate data being produced by the quality control system. Such criteria should also be established for each indicator of quality produced by the revised system.

--Establish a uniform sampling plan for selecting audit cases to evaluate under the Regional Case Management Program, establish standards for regional analysts to follow in evaluating the selected cases, and use the Regional Case Management Program to periodically evaluate the data produced by the quality control system.

IRS COMMENTS AND OUR EVALUATION

IRS agreed with each of our recommendations. In commenting on the text of this chapter, however, IRS expressed concern about the potential bias of our cross-district test. IRS said that reviewers, being aware of our study, felt they were being "tested" and that this feeling may have caused changes in reviewer judgments.

This same concern was raised during our discussions with district officials. Although they generally questioned the work of reviewers in other districts, they did not point out instances in which the reviewers' judgments were wrong. Nor did they question how well their own reviewers performed when evaluating other districts' cases.

Whether reviewers felt they were being tested or not, the test results show that lack of review standards cause inconsistent quality assessments. The results also show that reviewers need to take a harder look at cases and more frequently return questionable cases to examiners for clarification or correction. Unless IRS helps its reviewers do better work, and makes certain that they do, quality review results will continue to be conservatively biased.

CHAPTER 4

SCOPE OF REVIEW

To assess IRS audit quality and evaluate the effectiveness of its quality control system, we

--reviewed a statistical sample of tax returns;

--reviewed IRS policy and procedures manuals and internal audit and management reports;

--interviewed national, regional, and district office management officials, tax auditors and group managers, and those performing quality reviews (reviewers and conferees); and

--mailed questionnaires to all group managers and reviewers and to selected tax auditors in the five districts where we did our work.

In addition, we submitted six of our sample cases from each district to an IRS reviewer in each of the other four districts to assess review consistency.

Effective October 2, 1978, near the end of our field work, IRS reorganized for the stated purpose of better meeting its goal of administering the tax code effectively and economically and serving the taxpayer. As part of that reorganization, IRS eliminated the District Conference function and transferred to the technical review staff the responsibility for reviewing unagreed cases when the taxpayer asks for an appeal hearing. Our report discusses District Conference's role in reviewing unagreed cases as it was during our review.

We did our work at IRS' national office in Washington, D.C.; its Atlanta, Chicago, Dallas, Philadelphia, and San Francisco regional offices; and its Denver, Jacksonville, Los Angeles, Richmond, and St. Louis district offices.

An elaboration of how we selected districts, sampled cases, and controlled and analyzed the data collected is contained in appendix II.

COMMISSIONER OF INTERNAL REVENUE
Washington, DC 20224

JUN 20 1979

Mr. Allen R. Voss
Director, General Government Division
U.S. General Accounting Office
Washington, D.C. 20548

Dear Mr. Voss:

We appreciate the opportunity to review your draft report entitled, "Improvement Needed in IRS' Audits of Individual Taxpayers and Its Audit Quality Control System."

We are in general agreement with most of your recommendations. We enclose our specific comments.

Although we generally agree with the recommendations, which should enhance our quality control system, we believe it is necessary to exercise management discretion in our operational planning in assessing the significance of corrections.

GAO measured quality of audits using a standard of zero error or perfect quality. All dollar errors identified were counted, no matter how small. For example, 68% of the 730 errors GAO identified had a tax impact of $1 to $50. We do not correct small dollar errors if the correction would result in the taxpayer owing additional tax. The additional revenue received from such a correction does not warrant the administrative expense incurred and would unduly inconvenience the taxpayer. Accordingly, the standard used by GAO would always reflect a degree of error in our operations.

GAO recognizes the need for IRS to balance use of its examination resources to achieve adequate audit coverage and yet produce audits of acceptable quality. We believe the method of measuring quality used by GAO would not be a practicable operational measure to achieve an appropriate balance between adequate coverage and effective use of resources, or between quantity and quality.

Department of the Treasury Internal Revenue Service

APPENDIX I

Mr. Allen R. Voss

The report also indicates that oral testimony should not be considered sufficient when an issue warrants documentary support. Drawing the line in appropriate cases involves an exercise of sound judgment. The acceptability of oral testimony is a decision that must be considered in context of the entire audit. Courts consider oral testimony in arriving at their decisions. We have guidelines for our examiners concerning circumstances under which they should consider oral testimony.

With kind regards,

Sincerely,

Enclosure
Comments

Page 8, Paragraph 3

. . . Projecting our observed dollar errors to these (GAO) unreviewed cases, we estimate that the errors totaled $1 million.

(See similar passages on page 32, paragraph 3; page ii, paragraph 1, and in the cover summary, paragraph 2.)

Comments

GAO's projection of $1 million in tax change errors covers a period of four months in five districts. The projection includes both errors that would mean refunds to the taxpayers and errors that would mean additional revenue due the Government.

During the same four month period in the five districts, actual tax change recommendations including both refunds and assessments made by IRS for the type of cases included in the GAO study amounted to $26.3 million; added to the additional changes identified by GAO, the total is $27.3 million.

Viewed in this respect, IRS originally made about 96% of the combined total of actual IRS changes and GAO-identified changes. The GAO projection of $1 million is about 4% of this total.

In addition, the GAO projection includes an undetermined amount of small tax change errors which the IRS would not have corrected as a matter of practice.

Page 6, Paragraph 2

When an error was detected, we computed the dollar impact whenever possible no matter how small it was. IRS does not require examiners to correct small dollar errors detected during the postaudit review process if the errors are in the taxpayer's favor. We still included small dollar errors in our statistics because an accurate measure of quality requires that errors be considered, regardless of whether they are corrected.

Comments

In GAO's study, all errors were recorded, no matter how small, in order to accomplish a measurement of quality. Thus, for example, when GAO observed a $20 error, it was included in GAO's statistical analysis. We understand that in using this approach for statistical purposes, GAO is not suggesting we alter our existing procedures for correcting small errors.

Under our procedures, if correction of an error detected during the review process would result in a benefit to the taxpayer, we correct the error regardless of amount. If, however, correction of an error would result in the taxpayer owing a small amount of additional tax, such as $20, we would not recontact the taxpayer and correct the error. The additional revenue received from such a correction would not warrant the administrative expense incurred and would unduly inconvenience the taxpayer.

Page 19, Paragraph 3

We did not, however, consider oral testimony sufficient when the issue warranted documentary support. Of particular concern was examiner acceptance of oral testimony in support of claimed deductions. Auditing is a process whereby the accuracy of a reported figure is usually determined by checking the supporting documentation. Accepting oral testimony when documentation is in order provides little assurance of accuracy since the taxpayer, in effect, said the same thing when he filed his return. An examiner's acceptance of oral testimony when documention is required can also lead to inconsistent taxpayer treatment. . . .

Comments

The text indicates some concern about the use of oral testimony in resolving issues. However, there are instances when value judgements must be applied by tax auditors. It has been the policy of the Service for many years (Policy Statement P-4-39, approved May 2, 1966) that:

> Due consideration will always be given to the reasonableness of the taxpayer's claimed deductions in relation to his/her reported income, to the reliability and accuracy of his/her records in connection with other items more readily lending themselves to detailed recordkeeping, and to the general credibility of his/her statements in the light of the entire record in the case.

Courts consider the taxpayer's oral testimony in arriving at their decisions, and reasonable tax administration requires us to also consider a taxpayer's oral testimony.

Page 26, Paragraph 4 (See GAO note.)

To measure the effect of time on audit quality, we asked examiners, group managers, and reviewers a series of questions dealing with the time allowed to do audits. As previously discussed, they rated audit quality as high or very high. Notwithstanding this rating, most tax auditors, some group managers, and most reviewers felt that audit quantity was overstressed to the detriment of quality. Their opinions in this regard are summarized in the following table. (See table, page 27.)

Comments

While we have agreed to attempt to determine as much as possible about the relationship between the time spent on audits and their quality, we view the results of GAO's questionnaire on this subject as inconclusive. On Page 7 GAO estimated that tax auditors (69%) and group managers (68%) rated audit quality as high or very high. Conversely, the table on Page 27 shows that tax auditors (70%) and group managers (39%) feel that quantity is overstressed to the detriment of quality. Further, GAO reports that the same tax auditors (85%) and group managers (71%) said that tax auditors had less than or significantly less than adequate time to conduct audits.

GAO NOTE: We interpreted the response to mean that quality was acceptable considering the related time constraints.

Page 34, Recommendation 1

We recommend that IRS define audit quality in measurable terms so that examiners know what is expected of them.

Comments

Although audit quality is defined in our manual, we will attempt to develop a more comprehensive definition.

Page 34, Recommendation 2

We recommend that IRS do a controlled study to determine how long it takes to do a quality audit and use the study results, along with other factors, in planning the annual number of audits to be done.

Comments

We will attempt to develop a study to determine as much as possible about the relationship between the amount of time spent on audits and their quality.

Page 34, Recommendation 3

We recommend that IRS review district classification instructions and require the districts to rescind those procedures that may give classifiers the impression that they are to limit the number of items that can be identified for audit.

Comments

We concur. We will also issue guidelines to our classifiers in order to provide greater uniformity and consistency in identifying the significant issues on nonbusiness tax returns assigned to tax auditors.

Proper identification of the significant issues by the classifier will help assure IRS that the scope of the audit is adequate, and will keep taxpayer inconvenience to a minimum by identifying at one time all the records the taxpayer will normally be required to furnish for the audit.

Page 35, Recommendation 4

We recommend that IRS clarify that tax auditors have authority to expand the audit to significant items not identified by the classifier.

Comments

We generally concur. Our instructions will provide that tax auditors will be allowed to expand the audit when significant issues arise as a result of new information brought out during the audit. However, we will continue to require the approval of the tax auditor's group manager before a new issue may be raised during an audit based on information on the tax return. We believe this requirement is necessary to prevent a tax auditor from substituting his/her judgement for that of the classifier's, which could create inconsistent and non-uniform treatment of taxpayers, and could require taxpayers to frequently furnish additional records and return for another interview.

Page 35, Recommendation 5

We recommend that IRS require examiners to probe for unclaimed deductions and credits.

Comments

We concur.

Page 35, Recommendation 6

We recommend that IRS devise proforma audit aids for those tax issues most often examined to assist examiners in doing consistent and comprehensive audits and in documenting their work so that reviewers can better evaluate audit quality. Proforma aids should also be devised and used in probing for unreported income, deductions, and credits.

Comments

We concur.

Page 35, Recommendation 7

We recommend that IRS establish standards for group managers to follow in evaluating audit quality and in documenting their review efforts.

Comments

We concur.

Page 35, Recommendation 8

We recommend that IRS require tax auditors to document and reviewers to include in their review the taxpayer's reasons for claiming more than the tax auditor allowed, even though the taxpayer agreed to the adjustment of the issue.

Comments

In agreed cases, when all relevant facts and applicable provisions of the tax laws have been adequately documented in the workpapers, the adjustments have been explained to the taxpayer, and the taxpayer has agreed to the findings, we do not concur that documention of the taxpayer's position at the time the return was filed is relevant to assessing audit quality.

However, we will require that case files be adequately documented so that reviewers can review whether or not the findings of tax auditors were proper, and that the report to the taxpayer adequately explains why adjustments to the tax return were necessary.

Page 47, Paragraph 1

To test the effect of the lack of review standards, we submitted six of our sample cases from each of the five districts to a cross district review. The 30 cases were each reviewed 5 times -- once by the reviewer in the district that closed the case and once in each of the other 4 districts. We asked the reviewers to evaluate these cases just as they normally evaluate their own cases.

Comments

Several of our field offices have expressed concern with this methodology because of potential bias.

When reviewers were subsequently reviewing cases that had been previously reviewed in other districts, they had been made aware of GAO's study. Despite GAO's instructions, several reviewers felt they were being "tested," and this may have, itself, caused changes in their judgement.

Page 52, Recommendation 1

We recommend that IRS establish criteria for measuring audit quality. Criteria that should be considered include the dollar amount of the error and the scope, depth, and documentation of the audit.

Comments

We will attempt to develop criteria for measuring audit quality.

Page 54, Recommendation 2

We recommend that IRS revise the quality control system so that all types of errors in terms of the established measurement criteria are recorded and reported, quality problems are identified, and trends are developed.

Comments

We concur.

Page 54, Recommendation 3

We recommend that IRS establish standards for reviewers to follow in evaluating audit quality and in documenting their work so that the consistency and adequacy of their evaluations can be determined.

Comments

We concur. One of our recent actions has been to develop a quality control worksheet to be used in all districts. This worksheet will be tested in June 1979 and if found satisfactory will be implemented nationwide.

Page 54, Recommendation 4

We recommend that IRS establish criteria to assure uniform interpretation of the error rate data being produced by the quality control system. Such criteria should also be established for each indicator of quality produced by the revised system.

Comments

We concur.

Page 54, Recommendation 5

We recommend that IRS establish a uniform sampling plan for selecting audit cases to evaluate under the Regional Case Management Program, establish standards for regional analysts to follow in evaluating the selected cases, and use the Regional Case Management Program to periodically evaluate the data produced by the quality control system.

Comments

We concur.

GAO NOTE: The page references contained in IRS' comments have been revised to agree with the page numbers in this report.

SAMPLING AND DATA ANALYSIS METHODOLOGY

This appendix explains how we selected locations, audit cases, and questionnaire recipients; how we maintained quality control over the data collected; and how we projected and analyzed that data.

Selection of locations

In determining which IRS districts to include in our review, we considered three factors. We wanted to (1) obtain coverage across the United States and in different regions, (2) include different size districts in terms of the number of audits closed in a year, and (3) include districts with varying levels of quality as indicated by the error rates reported by IRS' quality control system. The offices selected and their characteristics are shown below.

	IRS region	Size	Error rate
Denver	Southwest	Small	Low
Jacksonville	Southeast	Large	Low
Los Angeles	Western	Very Large	High
Richmond	Mid-Atlantic	Small	Low
St. Louis	Mid-Western	Medium	High

Selection of audit cases

IRS groups individual income tax returns into business and nonbusiness returns. These two groups are further divided into audit classes on the basis of adjusted gross income. Business returns, which are not to be confused with corporate returns, include those that have a Schedule C (Profit and Loss From Business or Profession) or Schedule F (Farm Income and Expense) attached. Nonbusiness returns have neither of these schedules attached.

IRS audits of individual income tax returns are categorized as either office audits or field audits. Office audits generally involve less complex issues and are conducted either by correspondence or by interview at an IRS office. In a correspondence audit the taxpayer is asked to submit relevant material by mail. If IRS wants to audit the return through an interview, it tells the taxpayer what items he needs to support and sets an appointment at the IRS office.

Field audits generally involve the more complex issues and are usually conducted at the taxpayer's place of business.

The sample size required to provide statistically reliable data for all types of individual income tax returns would have been prohibitive in terms of time and GAO resources. Therefore, we decided to evaluate audit quality as it relates to office audits of individual nonbusiness returns, regardless of income class, wherein the taxpayer claimed itemized deductions. These audits account for most of the individual income tax audits done by IRS.

The method of selecting individual cases at each district was as follows. The cases completed by the Review and Conference staffs were made available to us each day. This point in IRS' process was selected because the post-audit review process should have been completed and observed errors corrected. Based on an analysis of previous review and conference workloads, a selection plan for each type of case was provided GAO teams at each district. In districts with small workloads, we selected all returns; at larger districts, we selected returns randomly. Our objective was to select an average of 2 to 2.5 cases per day of each type. Using this approach, we selected 240 of 824 review cases and 250 of 711 conference cases processed during the selection period. This period ranged from April 12, 1978, through September 1, 1978, depending on district workloads.

Because the cases were still in process at the time we selected them for analysis, we copied all case file documents and returned the original documents to IRS for further processing.

Selection of questionnaire recipients

To obtain information about audit quality and IRS' post-audit review process, we mailed questionnaires to examiners, group managers, and reviewers. We sent a questionnaire to each group manager and reviewer in the five districts visited. Because of the large number of examiners in the five districts, however, we randomly selected those to whom we would send questionnaires. Proportionate district samples of tax auditors and revenue agents were drawn according to the ratio they represented to the total examiner universe in the five districts.

To encourage those selected to participate in the review, we mailed up to two followup letters to nonrespondents.

Quality control over data collected

Maintaining quality control over the data to be collected and what was collected was an important consideration throughout the review.

Questionnaire data

We developed our questionnaires from discussions held with IRS employees during the initial phase of this assignment. We then pretested the questionnaires by administering them to IRS employees, asking the employees for their comments and opinions, and discussing their answers with them to see if they understood the questions and what their answers meant. As a result of information gathered during the pretest, we modified the questionnaires.

Each completed questionnaire was reviewed by a GAO staff member for completeness and a determination if the respondent's answers indicated an understanding of the question. Optional written comments were also reviewed to gain a better understanding of the respondent's opinions. We then keypunched the responses to create a computerized data base. We inspected the data base for obvious errors and verified a 10-percent random sample of the data elements back to the questionnaires. This verification process indicated an error rate of .1 percent, substantially under our tolerance level of 1 percent.

Audit case file data

We also maintained strict quality control over the information collected from audit case files. We incorporated IRS guidance on auditing procedures and techniques into a special manual and data collection instrument for recording the results of our case evaluation. We tested the instrument and the manual on actual cases and modified them where appropriate. All staff members attended training sessions on the use of the manual and instrument. Once we started our review, questions relating to either the instrument or the manual were centrally answered and each location was notified by phone and in writing of any further changes. If required, we reevaluated cases already completed in light of the approved modifications. GAO supervisors or another GAO employee compared the information recorded on each instrument with the related case file and signed off on the accuracy of the recorded information.

GAO staff members having overall responsibility for the review visited each location and reviewed cases for conformance to the manual. Finally, we discussed our conclusions with IRS officials at each location and recorded their reaction to each of our findings. We then keypunched the data collection instruments and subjected the resulting data bases to machine and manual edits. We also verified 10 percent of the keypunched data with an observed error rate of only .05 percent, well within a tolerable level.

Projection of sample results

After computerizing the two data bases, we weighted them in order to project sample results to

--all cases processed in a district during our review,

--all cases processed in the five districts during our review, and

--the universe from which questionnaire recipients were selected.

The following example illustrates our weighting methodology. One district's review staff processed 126 cases of which we selected 50. We calculated the weighting factor by dividing the stratum universe size by the stratum sample size (126/50 = 2.52). Therefore, any observed condition about one sampled review case in this one district can be projected to 2.52 review cases in that district. A similar system was used for weighting questionnaire responses.

For report purposes, we used the weighted mean as the best estimate because the projections are based on a statistical sample rather than a complete enumeration. The figures presented are subject to variation except where the weight used was 1, which would happen if all universe cases were selected.

Methods of analysis

We analyzed the data to determine the impact of certain factors on a given variable. We used the analysis of variance, the chi-square test of independence, and regression analysis.

Analysis of variance

Analysis of variance is a statistical technique used to test differences in means between two or more independent groups. In our analysis, we used it to test for differences in means between the five districts covered in our review. For example, were the districts significantly different in the extent to which they accepted oral evidence from the taxpayer?

We determined the significance of the differences in means between groups by using confidence levels which represent the probabilities that the differences were not products of chance. In interpreting the analysis results, we used a confidence level of 95 percent or greater as being significant.

Chi-square

Using the chi-square test of independence we (1) established the association between the variables tested and (2) determined the significance of the identified association.

To illustrate, data collected on one variable produced the following proportions of answers.

IRS district	Yes answers	No answers	Total
1	8	92	100
2	8	92	100
3	12	88	100
4	18	82	100
5	6	94	100

The proportion of yes answers varied from a high of 18 percent in District 4 to a low of 6 in District 5. But is this difference in proportions significant or merely the result of chance? We used the chi-square test of independence to evaluate these possibilities.

We determined the significance of the associations between the variables tested by using confidence levels which represent the probabilities that the associations were not products of chance. In interpreting our results, we used a confidence level of 95 percent or greater as being significant.

We used the chi-square test of independence for categorical variables and continuous variables expressed in terms of ranges. Continuous data were grouped into ranges so that expected cell frequencies would be at least five.

Regression analysis

We used regression analysis to determine what impact certain factors had on a given variable. For example, can the variance in the number of problems found on an audited return after its evaluation by the quality control system be explained in terms of certain characteristics of the case, such as the adjusted gross income reported by the taxpayer? Again, we used a 95 percent-confidence level.

TAX ISSUES WHERE TECHNICAL ERRORS WERE MADE

Tax Issues	Review cases	Conference cases
	(percent)	
State and local taxes	4.7	4.5
Employee business expenses	4.2	.4
Personal insurance premiums	4.0	0
Rents and royalties	1.7	1.8
Contributions	2.7	0
Wages	1.5	.8
Casualty or theft loss	1.3	.3
Business auto expenses	.4	1.5
Personal living expenses	.1	1.5
Unreported income	1.3	0
Duplicate deduction	.6	.7
Moving expenses	0	1.1
Child care	.2	0
Sick pay exclusion	.4	.3
Nonexcludable unearned income	.4	.3
All other	5.0	3.9

TAX ISSUES WHERE SIGNIFICANT ITEMS WERE MISSED

Tax Issues	Review cases	Conference cases
	----------(percent)----------	
Interest expense	11.2	9.3
Deductible employee expense	5.0	11.0
Sales tax	4.9	8.6
State tax refund	10.4	1.8
Rent and royalty	7.2	5.1
Other income	4.2	8.6
Medical (except mileage)	6.6	4.5
Capital gains and losses	5.0	3.4
Contributions	6.1	1.8
Other deductible expenses	3.6	2.8
Gas tax	2.1	4.4
Real estate tax	2.4	2.3
Filing status	4.0	0
Medical mileage	2.3	1.8
Personal property tax	1.9	1.4
Dependents and exemptions	1.8	1.1
All taxes	2.1	.8
Travel	1.6	.8
Education expenses	.7	1.8
Income producing expenses	.4	2.3
IRA or Keogh	1.1	1.1
Casualty and theft losses	1.0	.8
Sick pay exclusion	1.3	.1
Credit for the elderly	1.3	0
Other	5.5	9.8

(268028)

☆ U.S. GOVERNMENT PRINTING OFFICE: 1979 - 620-167/277

Single copies of GAO reports are available free of charge. Requests (except by Members of Congress) for additional quantities should be accompanied by payment of $1.00 per copy.

Requests for single copies (without charge) should be sent to:

U.S. General Accounting Office
Distribution Section, Room 1518
441 G Street, NW.
Washington, DC 20548

Requests for multiple copies should be sent with checks or money orders to:

U.S. General Accounting Office
Distribution Section
P.O. Box 1020
Washington, DC 20013

Checks or money orders should be made payable to the U.S. General Accounting Office. NOTE: Stamps or Superintendent of Documents coupons will not be accepted.

PLEASE DO NOT SEND CASH

To expedite filling your order, use the report number and date in the lower right corner of the front cover.

GAO reports are now available on microfiche. If such copies will meet your needs, be sure to specify that you want microfiche copies.