United States General Accounting Office

GAO

Report to the Chairman, Subcommittee on Oversight, Committee on Ways and Means, House of Representatives

March 1991

TAX ADMINISTRATION

IRS Needs to Improve Certain Measures of Service Center Quality

GAO/GGD-91-66

GAO

United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-242491

March 20, 1991

The Honorable J.J. Pickle
Chairman, Subcommittee on Oversight
Committee on Ways and Means
House of Representatives

Dear Mr. Chairman:

This report responds to your request that we provide information on steps that the Internal Revenue Service (IRS) is taking to monitor and improve the quality of various service center activities. As agreed with the Subcommittee, we focused on three specific activities: (1) the quality of correspondence with taxpayers about adjustments to their accounts, (2) the extent to which IRS errs in processing tax returns, and (3) the accuracy of notices sent to taxpayers informing them about changes made to their returns during processing.

Results in Brief

IRS appears to have improved the quality of correspondence it sends taxpayers to tell them about adjustments to their accounts. Over the last 2 years, IRS statistics show that the percentage of its letters with critical errors decreased from 38 percent to 14 percent. IRS plans to begin monitoring other types of service center correspondence in 1991.

IRS' progress in improving the accuracy with which it processes tax returns is less clear. The data IRS uses to monitor its progress in improving processing quality show a decrease in error rates from 23 percent in 1988 to 18 percent in 1990. This error rate is not a valid indicator of IRS' processing quality, however, because it (1) includes errors made by taxpayers and (2) counts as errors things that are not errors. IRS' progress in improving the quality of its processing performance is further clouded by the fact that service centers have been making the same types of processing errors year after year. IRS has begun studying ways to reduce some of those recurring errors in the hopes of having solutions in place for the 1991 filing season. Reports on those projects had not been finalized as of January 1991.

IRS lacks an effective measure of the accuracy of notices it sends taxpayers. Output review—the service center unit that reviews notices for accuracy—primarily selects notices for review on the basis of the historical probability of that type of notice having an error. While that strategy helps target limited resources toward the most error-prone notices, it does not provide IRS with a valid measure of the accuracy of all notices. Moreover, the opportunity to reduce errors on notices was hampered at two service centers we visited. Output review units at the centers were not following IRS procedures directing them to report identified errors to the managers of units that originated the notices. IRS' Internal Audit Division officials said that the division is completing a review of the notice review process that focuses on some of these issues.

In light of the decrease in correspondence errors reported by IRS and IRS' plans to expand its efforts, we are not making any recommendations in this area. We are recommending, however, that IRS develop better measures for assessing the quality of returns processing and notices. We think that IRS could fund development of those measures with existing resources if it reassessed the nature and extent of its current quality control efforts and the way in which it uses the data it currently collects. We also see opportunities for IRS to make better use of output review results to identify and correct the causes of erroneous notices.

Objective, Scope, and Methodology

Our objective was to provide information on IRS' efforts to measure and improve the quality of service center (1) correspondence, (2) returns processing, and (3) notices. Our correspondence work focused on letters sent out by the service centers' Adjustments/Correspondence Branch to explain adjustments to taxpayer accounts. It did not include correspondence sent out by other service center functions such as collection and examination. Our returns processing work focused on income tax returns going to the service centers' Error Resolution System—a system for correcting errors on tax returns. Our notice work focused on those notices issued as a result of processing tax returns, such as notices to advise taxpayers of changes IRS made to their returns because of math errors or name and address corrections. We did not inquire into other types of notices, such as collection notices or notices of underreported income.

To meet our objective, we

- analyzed descriptive information on the various types of service center quality reviews;
- assessed available IRS data on quality trends and the level of quality achieved in correspondence, returns processing, and notices;
- interviewed managers in each of the three areas and the employees responsible for quality measurement in IRS' National Office and at 2 of its 10 service centers (Andover, Massachusetts, and Cincinnati, Ohio);

- observed quality measurement personnel in Andover and Cincinnati as they analyzed correspondence and notice cases and examined a limited number of those cases to verify the case review process; and
- coordinated with IRS' internal auditors who were doing related work.

We selected the Andover and Cincinnati service centers in an attempt to attain some geographic coverage.

As agreed with the Subcommittee, we prepared this report on the basis of work we had done as of mid-December 1990. We did not verify the accuracy of the IRS data discussed in the report nor did we determine causes or evaluate effects of any noted discrepancies.

We did our work from June to December 1990 in accordance with generally accepted government auditing standards. IRS provided written comments on a draft of this report. Those comments are presented in appendix I.

IRS Reports Significant Improvement in Correspondence Quality

The Adjustments/Correspondence Branch in each of IRS' 10 service centers is responsible for quickly and accurately responding to taxpayer inquiries and notifying taxpayers of adjustments to their accounts. Among other things, tax examiners in the branches may need to explain an action taken by IRS, respond to questions about a taxpayer's account balance, decide whether a penalty should be abated, or adjust the taxpayer's account.

In July 1988, we reported on the results of our random sample of correspondence sent to taxpayers by the Adjustments/Correspondence branches in three service centers (Fresno, California; Kansas City, Missouri; and Philadelphia, Pennsylvania) between May 4 and July 31, 1987.[1] Our review showed that 31 percent of the correspondence cases closed by the three centers during those 3 months had critical problems. We considered a problem critical if the tax examiner provided incorrect information, failed to address all the taxpayer's questions, or acted incorrectly in adjusting or failing to adjust the taxpayer's account.

We identified various factors that contributed to the correspondence problems at those three centers. For example, tax examiners generally responded to taxpayers by selecting prepared letters from a catalog of

[1] Tax Administration: IRS' Service Centers Need to Improve Handling of Taxpayer Correspondence (GAO/GGD-88-101, July 13, 1988).

about 500 letters. These computer-generated form letters contained required and optional paragraphs and various data fields to be filled in by the examiner. Because of the large number of potential letters and paragraphs, and because the examiner did not get to review the completed letter, this system did not facilitate preparation of appropriate responses. We also noted that our 31-percent critical error rate was much higher than that reported by the service centers. We attributed that, in part, to the fact that our rate was based solely on a review of cases involving correspondence, while the service centers' rates were based on a review of all Adjustments/Correspondence Branch activities, some of which do not involve correspondence. We recommended, among other things, that tax examiners be allowed to view a letter after it is composed and that IRS report adjustment case correspondence error rates separately.

After receiving our report, IRS modified its primary quality measurement program at the service centers—the Program Analysis System (PAS)—to enable it to specifically assess the quality of correspondence being sent to taxpayers by the Adjustments/Correspondence Branch. In January 1988, PAS analysts began reviewing sample adjustments/correspondence cases, compiling statistical data, and reporting their findings to the branch and the National Office. The analysts check correspondence for such things as completeness, accuracy, and grammatical correctness. They categorize any errors as either critical (such as those affecting the letter's accuracy or completeness) or noncritical (such as those involving spelling or punctuation).

PAS reports showed that the nationwide critical error rate decreased from 38 percent in February 1989 (the earliest PAS report available) to 14 percent in August 1990 (the most recent PAS report available at the time of our review). We did not assess the PAS methodology and thus have no basis for determining whether those error rates are reliable.

As shown in table 1, the nationwide error rate reported by PAS generally declined from February to August 1990, although the rates at individual centers varied (sometimes significantly) among centers and from month to month.

Table 1: Critical Error Rates for Letters Sent to Taxpayers by Service Center Adjustments/Correspondence Branches (February - August 1990)

	Error rates						
Service center	**February**	**March**	**April**	**May**	**June**	**July**	**August**
Andover	50.0	28.1[a]	36.0[a]	40.5	22.9[a]	20.8[a]	11.8[a]
Atlanta	32.1	34.0	32.7	44.3	20.7	19.6	9.2
Austin	15.8	7.8	13.2[a]	8.8	18.4[a]	13.8	16.0
Brookhaven	27.1	14.3[a]	21.7	18.3	12.9	15.9	18.8
Cincinnati	23.5	23.8	14.0	17.6	18.3	23.1	23.8
Fresno	31.7	13.3	13.0	6.5	6.1	8.1	8.6
Kansas City	10.0	29.6	9.9	18.1	8.0	16.2	20.8
Memphis	30.3	18.4	24.0	7.7	20.0	17.7	8.2
Ogden	5.0	21.0	9.4	13.5	10.7	7.3	5.2
Philadelphia	27.1	25.6	15.6	29.3	13.5[a]	13.2	0.0[a]
National	24.4	19.3	16.8	15.5	13.9	15.1	14.1

[a]Sample too small to meet IRS criteria of 95-percent confidence level and a sampling error, or precision, of $\pm$ 5 percent. IRS changed the PAS sampling methodology for fiscal year 1991, which should better ensure sufficient sample sizes in the future.

Sources: IRS PAS reports, July and August 1990.

The director of IRS' Returns Processing and Accounting Division, who oversees the adjustments/correspondence function, said that the varying service center rates could be due either to differences in performance or differences in the experience levels of PAS analysts. He explained that a more experienced analyst would be more adept at detecting errors, which would be reflected in higher error rates.

IRS officials attributed the improved performance reflected in table 1 to such things as an increased emphasis on improving quality, the initiation of PAS quality reviews, and a stabilizing work force of more experienced tax examiners. Additionally, IRS implemented a new correspondence system—the Expert correspondence system—in all service centers in 1990 to help tax examiners compose letters and allow examiners to review composed letters before they are mailed. We observed the Expert system being used at the Cincinnati Service Center and found it to be an improvement over the process that existed during our review in 1987.

IRS Plans to Expand Its Correspondence Quality Reviews

There are other service center functions, besides the Adjustments/Correspondence Branch, that correspond with taxpayers. Those functions include taxpayer service, collection, examination, and document matching. In August 1990, an internal correspondence task force recommended, among other things, that a quality measurement system like PAS be implemented for all service center functions during fiscal year 1991. IRS' National Office PAS correspondence analyst said that service centers are scheduled to implement PAS correspondence reviews in the document matching and taxpayer service functions in June or July 1991 and in the collection and examination functions in October 1991.

Imprecise Measures of Quality and Recurring Processing Errors Hinder Returns Processing Quality Improvement

Each filing season, IRS receives and processes millions of tax returns at its 10 service centers. That processing involves several functions. The receipt and control function receives, extracts, and sorts the tax returns. The document perfection function edits the returns for completeness, legibility, and accuracy and places codes on the returns to facilitate entry of return data into the computer. The data conversion function transcribes information from each return onto magnetic tape for computer processing. The computer checks the transcribed information for such things as math errors. If the computer finds what appears to be an error, it refers the return to the service center's error resolution function for review and any necessary correction.

IRS measures the quality of its returns processing by tracking the percentage of individual returns that go to error resolution. As shown in table 2, that percentage has declined over the past 3 years.

Table 2: Percentage of Individual Returns Sent to Error Resolution

Year	Percentage[a]
1988	23
1989	20
1990	18

[a]This percentage is calculated by dividing the number of individual returns going to error resolution by the total number of processed individual returns We focused our analysis on individual returns because they account for the bulk of the tax returns processed by service centers

One of IRS' performance priorities for fiscal year 1991 is to improve returns processing quality during the 1991 tax return filing season. It has thus tasked its service center managers with lowering the rate of returns going to error resolution by 2 percentage points. Our work indicated, however, that the percentage of processed returns going to error resolution may not be a valid indicator of returns processing quality, because not all of the errors referred to error resolution result from processing problems and some returns are sent to error resolution for reasons other than the presence of an error.

For example, one thing the computer is programmed to look for in identifying returns that should be sent to error resolution is an error, such as numbers that do not add. Those errors could be due to mistakes by taxpayers in preparing their returns or by service center staff in processing the returns. The computer also refers returns to error resolution not because they contain errors but because they involve conditions that meet certain predetermined criteria that IRS has programmed into the computer, such as a return coded to indicate the existence of fringe benefits but with no wages reported. IRS has decided that it wants such returns reviewed, as a form of consistency check, even if they prove to be error free.

For the reasons cited above, we believe that the inclusion of taxpayer errors and consistency checks in the figures cited in table 2 produces a distorted picture of returns processing quality. A decrease in error rates and an apparent improvement in quality could be achieved, for example, simply by a reduction in the number of consistency checks. IRS has, in fact, changed its consistency check criteria from year to year, thus causing the number of returns sent to error resolution for consistency checks to vary. For example, there were 71 types of consistency checks in 1990 compared to 85 in 1989.

We believe that a more valid indicator of returns processing quality is provided by PAS. In 1985, IRS began using PAS to sample individual returns handled by error resolution—an effort that was expanded in 1988 to include business returns. PAS analysts review the sample returns to identify the type of error involved, if any, and the source of the error. PAS segregates findings into three main source categories—errors made by taxpayers in filling out their returns, errors made by IRS in processing the returns, and consistency checks in which no errors were identified. By segregating the source, PAS provides IRS with an opportunity to focus on IRS-caused errors—the errors that IRS has the best chance to correct.

Using PAS data for 1988, 1989, and 1990, we stratified returns sent to error resolution on the basis of the source of the error. Table 3 shows that returns with taxpayer errors, as a percentage of all processed returns, declined by 2 percentage points, while the percentage of returns with IRS processing errors declined by 1 percentage point since 1988—a different picture than that provided by the data in table 2.

Table 3: Reasons Why Returns Were Sent to Error Resolution in 1988, 1989, and 1990

	Percentage of processed returns		
Reason	**1988**	**1989**	**1990**
Taxpayer error	9	8	7
IRS processing error	10	10	9
Consistency check[a]	4	2	2
Total returns	**23**	**20**	**18**

[a]These are returns on which error resolution identified no error. If an error was identified, the return would be categorized as either a taxpayer or IRS error, whichever applied.

Sources: National Office error analysis reports dated August 24, 1988, October 18, 1989, and August 7, 1990.

The director of the Returns Processing and Accounting Division agreed that a better indicator of returns processing quality is needed. He said, however, that before PAS results could be used for that purpose, the number of returns sampled by PAS would have to be increased substantially. As an alternative, he said the Cincinnati Service Center is testing a procedure whereby error resolution staff are recording the types and sources of all errors they correct. If the test is successful and the procedure can be implemented nationwide, it will provide IRS with a quality indicator for all returns sent to error resolution, not just the sample analyzed by PAS. The director noted, however, that funds might not be available to implement the procedure even if the test proves successful. He expects to have test results later in 1991.

IRS Is Attempting to Correct Chronic Processing Problems

One purpose of PAS is to identify systemic errors so that IRS can isolate and correct the causes of those errors and prevent their recurrence. PAS results show that most IRS processing errors occur when (1) service center staff in the document perfection function fail to add codes that direct the computer to take certain actions or (2) staff in the data conversion function fail to enter information into the computer. Table 4 shows the six most prevalent IRS processing errors identified by PAS during the 1990 filing season. As also shown in the table, the same errors were prevalent in prior years—some for the past 4 years.

ɔle 4: Top Six IRS Processing Errors Individual Returns in 1990 and Their nk in Prior Years

Errors	1990	1989	1988	1987
Document perfection failed to properly enter a code needed by IRS to process the return	1	1	1	1
Document perfection failed to properly code a dependent's identification number	2	2	2	[a]
Data conversion failed to enter the dependency status indicator	3	4	4	[a]
Document perfection failed to code the taxpayer's filing status	4	5	5	5
Document perfection failed to code the taxpayer's zip code	5	6	[a]	[a]
Data conversion failed to enter an estimated tax penalty	6	7	[a]	[a]

[a]Not ranked in the top 10

Sources: IRS PAS reports

In 1990, as part of its overall quality improvement effort, IRS assigned 13 processing concerns to various field offices for detailed review in the hopes of having some solutions in place for the 1991 filing season. These concerns included the problem ranked first in table 4 and other problems involving (1) taxpayer errors in claiming earned income tax credits, (2) IRS errors in processing business returns, and (3) IRS errors in processing taxpayer correspondence. IRS expects the field offices to assess the problems, develop solutions, and recommend corrective actions. Reports on those projects had not been finalized by late January 1991. We do not know whether the projects will have an effect on the 1991 filing season. As discussed with the Subcommittee, we will monitor and update the status of those projects as part of our 1991 filing season review.

Problems Identified in IRS' Notice Review Process

As noted earlier, many of the returns that go to error resolution involve mistakes made by taxpayers in preparing their returns. Those mistakes involve such things as math errors and omitted entries. When IRS corrects a taxpayer's mistake, it informs the taxpayer of the correction via a processing notice.[2] Tax examiners in error resolution assign a code to each error they correct and input the code to the computer, which automatically generates a notice. The code tells the computer which explanation to include on the notice. In 1989, the 10 service centers generated about 47 million processing notices for individuals and businesses.

Output review units in the service centers review samples of the processing notices before they are mailed to ensure their accuracy and appropriateness. When unit examiners receive notices to review, they also receive such related documents as account transcripts and tax returns, if needed to aid their review. There are various things an examiner might check in reviewing a notice. If the notice refers to a missing tax payment, for example, the examiner is supposed to check the most current data to determine whether payment was received after the notice was generated. If payment was received, IRS would void the notice.

Output review's goal is to stop erroneous notices from being issued. To do that, it attempts to focus on notices that are most likely to contain errors. The primary vehicle through which IRS identifies such notices is an automated system known as the Notice Review Processing System (NRPS). The system uses selection criteria established by the National Office to identify notices for review at each of the 10 service centers. The NRPS selection criteria are based on error rate information provided by output review unit officials. Each center can adjust the number of notices it reviews to reflect local conditions. If a center believes, for example, that it needs to review more of a particular type of notice than would result from application of the NRPS criteria, it can take steps to have more of those notices identified for review.

PAS Reviews Indicate That Output Review Is Performing Effectively

IRS keeps statistics on the extent to which output review finds errors on the notices it reviews. As shown in table 5, IRS statistics for the first 5 months of 1988, 1989, and 1990 show that notice error rates varied from service center to service center and from year to year. We could

[2]We refer to these as processing notices to distinguish them from other notices that service centers send taxpayers, such as notices related to underreported income identified through the matching of information returns with tax returns.

not determine whether the varying rates were due, for example, to a variance in the quality of the notices being generated by those centers, adjustments made to the notice selection criteria, or a variance in performance of the output review units in those centers.

Table 5: Error Rates on Notices Reviewed by Output Review for the First Months of 1988, 1989, and 1990

Service center	Notices to individuals 1988	1989	1990	Notices to businesses 1988	1989	1990
Andover	15.0	14.2	11.9	54.6	51.3	47.8
Atlanta	7.8	13.8	10.3	28.0	48.3	29.9
Austin	15.7	15.9	15.5	45.6	50.3	44.5
Brookhaven	14.5	12.4	10.9	30.4	30.0	35.4
Cincinnati	10.9	11.5	13.6	32.2	44.0	48.5
Fresno	8.5	16.5	10.2	13.1	48.3	42.7
Kansas City	8.0	17.4	15.8	13.6	47.4	63.0
Memphis	7.8	15.2	10.4	31.5	52.0	45.9
Ogden	11.0	12.2	7.9	45.4	39.4	42.2
Philadelphia	9.6	8.1	16.8	41.7	18.8	38.1
National	11.0	13.6	12.4	33.5	44.9	43.4

Note: We used data for the first 5 months because that was the only period for which comparable data were available from IRS. For the year through December 7, 1990, the national error rate for notices to individuals was 11.8 percent, ranging from 7.7 percent in Ogden to 15.7 percent in Philadelphia, and the national error rate for notices to businesses was 38.4 percent, ranging from 22.3 percent in Fresno to 56.7 percent in Kansas City.

Sources: 1988 and 1989 figures came from a June 1989 IRS report. 1990 figures came from comparable data in IRS' Management Information System for Top Level Executives.

A measure of the relative effectiveness of each center's output review unit is provided by PAS, which reviews samples of notices examined by the units. PAS measures effectiveness by comparing the accuracy of notices received by output review with the accuracy of the notices when they leave output review. Table 6 shows this information for March through September 1990, the period IRS used in its 1990 PAS report.

Table 6: Output Review Effectiveness
(March - September 1990)

Service center	Received accurate[a]	Sent accurate[a]
Andover	84 4	90.7
Atlanta	78.3	90.4
Austin	93.5	97.7
Brookhaven	74.9	92.8
Cincinnati	82.5	95.4
Fresno	87.5	94.2
Kansas City	78.9	95.9
Memphis	86.7	95.9
Ogden	73.0	91.8
Philadelphia	81.8	92.4
National	81.2	93.6

[a]The "received accurate" figures indicate the percentage of notices that were accurate when received by output review and thus needed no change. The "sent accurate" figures indicate the percentage of notices that were accurate when they left output review. The difference between the two columns reflects output review's effectiveness in improving the quality of the notices it reviews

Source: September 1990 PAS report on output review.

As shown in table 6, the accuracy of notices received by output review units ranged from 73 percent to 93.5 percent while those released to taxpayers had less variability—ranging from 90.4 percent to 97.7 percent accurate.

IRS Does Not Measure Overall Notice Quality

Processing notices are the vehicle through which service centers advise taxpayers of errors they made in filling out their returns. Considering the potentially adverse impact of incorrect notices on taxpayer relations, we believe that the quality of those notices is an important indicator of a service center's performance. IRS does not have such an indicator.

As discussed earlier, IRS has error rate information generated by output review and data on output review's effectiveness generated by PAS. Neither of those measures, however, is an indicator of notice quality. The error rate information generated by output review units relates only to the notices they review. Because NRPS and output review seek to select notices with the highest potential for error, some types of notices do not get reviewed. IRS, therefore, has no measure of the quality of those unreviewed notice types. Also, because the sampling rates can vary for those notice types that are reviewed, the error rate information generated by output review may not even be a valid indicator of the quality of those types of notices.

Because NRPS is intended to identify those notices most likely to have errors, IRS uses the error rate data as an indicator of the effectiveness of that computerized selection system, not as a measure of overall notice quality. If output review finds a high percentage of errors on the notices that NRPS sends it, IRS considers NRPS effective in identifying erroneous notices. A low rate, on the other hand, might be grounds for adjusting the sampling rate, either nationally or locally, in order to focus on error-prone notices. Such adjustments further reduce the usefulness of NRPS error rates as a measure of notice quality.

Because PAS generates its data by reviewing samples of output review's work, its results, as discussed earlier, are an indicator of output review's effectiveness, not of notice quality.

Output Review Information Not Being Used to Identify Causes of Errors or to Update NRPS Selection Criteria

IRS procedures specify that output review units are to identify causes of erroneous notices and provide information to the responsible service center area so they can eliminate causes. It is important that IRS correct systemic problems so that it can prevent erroneous notices and thus limit the costs incurred in reviewing and correcting those notices.

At the time of our visits, output review examiners in the Andover and Cincinnati service centers were recording the disposition of each of their reviews on checksheets. In so doing, they indicated whether the reviewed notice was correct or whether it had to be revised or voided. However, the examiners did not record the type of error found on the notice nor the source of the error—information that would be useful in identifying systemic errors and their causes. Furthermore, neither of the two centers' output review units had procedures for routinely recording the causes of erroneous notices and providing feedback to the service center units that originated the notices. Instead, officials at both centers said informal feedback was provided if a reviewer happened to observe recurring errors. Without formal feedback, IRS cannot be assured that the causes of erroneous notices are identified and corrected.

Better feedback might, for example, be useful for reducing the high error rates for business notices shown in table 5. The Cincinnati Service Center output review manager said that error rates for business notices are much higher than those experienced for individual notices because many of the errors involve situations in which payments were received after the notice was generated. IRS did not have specific data, however, showing the extent to which erroneous business notices were due to late-received payments versus other reasons involving mistakes by IRS.

In our opinion, IRS needs that kind of causal information to have a sound basis for deciding what type of action is needed to improve its process for generating notices.

As noted earlier, output review examiners are required to record the disposition of each notice reviewed. The Cincinnati Service Center output review manager said that service centers send weekly disposition information to the National Office for use in determining whether the NRPS selection criteria need to be adjusted. According to a National Office NRPS official, however, the data are not being used to adjust the NRPS criteria. The official said that criteria adjustments are based on input from service center output review officials during end-of-year meetings and informal discussions throughout the year. In our opinion, use of the weekly disposition information provided by service centers would enable the National Office to make more timely, and perhaps more precise, adjustments to NRPS selection criteria. Nonuse of these data might mean that service centers are incurring unnecessary costs in compiling and forwarding the information.

Internal Audit Has Been Evaluating the Notice Review Process

In March 1990, IRS' internal auditors began a review of NRPS to evaluate (1) NRPS selection criteria; (2) the adequacy of NRPS operating procedures; (3) the adequacy of management information for measuring service center performance and monitoring the effectiveness and efficiency of NRPS; and (4) the National Office's role in planning, directing, controlling, and monitoring the design and implementation of NRPS.

Internal Audit did work at the Austin and Cincinnati service centers and, among other things, reviewed about 4,500 notice cases. As the review progressed, Internal Audit developed additional areas of interest and expanded its scope. In January 1991, Internal Audit representatives said that their report would focus on NRPS and the selection of notices for review, the effectiveness and efficiency of output review in examining selected notices, and the reporting of activities and results. They expect to address, as part of that focus, error rate variances among service centers and the extent to which output review does or does not identify the causes of erroneous notices and communicate such information to the responsible area. As such, that report should address some of the areas that our work indicated needed attention, including the use of output review dispositions to adjust NRPS selection criteria. Internal Audit plans to issue its report to IRS management in March 1991.

IRS May Have to Reassess How It Funds Quality Control

As discussed previously, IRS does not have reliable indicators of the quality of its returns processing function or its processing notices. Without these indicators, IRS does not have valid, consistent measures of the errors it makes when processing tax returns and sending notices to taxpayers. The development of such indicators could entail additional costs.

IRS might be able to minimize the costs associated with developing improved quality indicators by reassessing the way it now spends its quality assurance dollars. The service centers now do various types of quality reviews directed at different aspects of performance ranging from in-depth evaluations of complete processes to evaluations of a particular employee's performance. We believe that IRS needs to reassess all of those quality assurance efforts to determine which are most important in helping IRS satisfy its mission of providing quality service to taxpayers and what would be the most appropriate mix of those efforts. That assessment should include consideration of the need for quality indicators, such as those discussed in this report, that IRS can rely on to measure its success in meeting that mission.

IRS might also be able to minimize the additional costs associated with developing improved indicators by assessing how it uses the quality data it already collects. For example, IRS might be able to use statistical techniques to combine the results from its reviews of notices identified by NRPS with random samples of other types of notices to produce a valid indicator of overall notice quality. It might also be able to use the samples it takes of individual employees' work to help develop measures of overall returns processing performance rather than building entirely separate samples.

Conclusions

In each of the three service center activities discussed in this report, IRS has taken steps directed at enhancing quality. It has (1) improved the process by which the Adjustments/Correspondence branches compose letters to taxpayers and has begun measuring the quality of those letters, (2) assigned certain recurring processing problems for in-depth analysis in the hope of identifying causes and solutions, and (3) established a system to review certain returns processing notices before they are sent to taxpayers.

We believe that IRS' quality efforts could be further enhanced if IRS used performance indicators that more precisely measured the level of quality being achieved. The indicator IRS uses to measure returns processing quality and to hold managers accountable for improving quality is flawed by the fact that it measures more than the quality of the returns processing function. It includes taxpayer errors, which are outside the control of the returns processing function, and it includes consistency checks that are not errors. IRS also has no true measure of the quality of its processing notices. The indicator it now has only measures the quality of notices reviewed by the output review units—reviews that are targeted at notices deemed most likely to be in error. If IRS wants to know how good a job it is doing in preparing notices and whether and to what extent its performance is improving, it needs to develop another indicator—not in place of but in addition to the indicator produced by output review.

We recognize that developing such indicators may entail additional costs. We believe, however, that IRS might be able to minimize or absorb that cost within existing funds if it reassessed the way it now spends its quality assurance dollars and the way it uses the quality data it now collects.

Another important step IRS could take to enhance its quality improvement efforts is to ensure that output review units document the kinds of errors being made and the sources of those errors and that the information is used to identify systemic problems. Such a process is vital if IRS hopes to eventually correct those problems and thus prevent errors from occurring in the first place.

Recommendations to the Commissioner of Internal Revenue

To better assess its returns processing performance, IRS should (1) use an indicator to measure returns processing quality that identifies the extent to which returns are being sent to error resolution specifically because of errors made by service center staff in processing returns and (2) measure the overall quality of returns processing notices rather than just those that are referred to output review. To help fund the development of such indicators, IRS should reconsider how it spends the money currently available for service center quality control efforts and assess how the quality data it now collects might be used to help build the needed indicators.

IRS should also (1) compile data on output review results in a way that enables management to identify specific problems that need to be addressed and (2) ensure that the results are so used.

Agency Comments and Our Evaluation

The Commissioner of Internal Revenue commented on a draft of this report by letter dated March 11, 1991 (see app. I). The Commissioner said that improving quality has been a major IRS objective as reflected in IRS' Strategic Business Plan, which includes actions that he said parallel several of our recommendations.

The Commissioner said that (1) the returns processing function is developing a system that will enable IRS to separate processing errors caused by the service centers from taxpayer errors, (2) notices have been identified as one of the key returns processing product areas for which quality standards are to be developed during fiscal year 1991 for application in fiscal year 1992, and (3) IRS would consider approaches to reviewing all notices, not just those in output review.

While agreeing that improvements can be made in how IRS spends its quality assurance dollars, the Commissioner pointed out that measuring employees' performance commands most of IRS' quality resources because the frequency of reviews and sample size have been negotiated with the National Treasury Employees Union. We understand the need to measure individual employees' performance and are not suggesting that IRS stop doing so (although IRS might want to consider how frequency and sample sizes might be adjusted in future negotiations with the union). As we noted earlier, however, IRS should consider how it might use the samples it takes of individual work to help build a measure of overall returns processing performance. In commenting on that, the Commissioner said that IRS is taking steps to develop quality standards for key returns processing products and services and is assessing available quality data for use in developing those standards.

The Commissioner acknowledged that (1) output review examiners were not recording the types or sources of errors found on notices and (2) output review units did not have procedures for routinely recording the causes of erroneous notices and providing feedback to the units that originated the notices. He noted, however, that the PAS operation in output review is designed to accumulate that type of information. Because PAS only reviews a sample of output review's work, that information would not be as useful, in our opinion, as similar information recorded by output review examiners for each of the notices they review. Finally, the Commissioner said that weekly notice disposition reports were not being used to adjust the NRPS criteria because of a programming deficiency that caused incorrect data. He said that the notice disposition reporting system has been redesigned and is to be tested in July 1991.

As arranged with the Subcommittee, we are sending copies of the report to other congressional committees, the Secretary of the Treasury, the Commissioner of Internal Revenue, and other interested parties.

Major contributors to this report are listed in appendix II. Please contact me on (202) 275-6407 if you or your staff have any questions concerning the report or our continuing work for the Subcommittee.

Sincerely yours,

Jennie S. Stathis

Jennie S. Stathis
Director, Tax Policy and
Administration Issues

Comments From the Internal Revenue Service

Note: A GAO comment supplementing those in the report text appears at the end of this appendix.

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, D.C. 20224

COMMISSIONER

MAR 11 1991

Mr. Richard L. Fogel
Assistant Comptroller General
United States General Accounting Office
Washington, DC 20548

Dear Mr. Fogel:

We have reviewed your recent draft report entitled, "Tax Administration: Quality of Various IRS Service Center Activities".

We are pleased that the report notes significant improvement in the quality of Service Center correspondence. Improving the quality has been a major IRS objective as reflected in the Strategic Business Plan which includes actions that parallel several of the report's recommendations.

Detailed comments regarding the report recommendations are enclosed. We have also enclosed general comments on additional actions the IRS has undertaken to improve the quality of our taxpayer correspondence.

Best regards.

Sincerely,

Fred T. Goldberg, Jr.

Enclosure

IRS COMMENTS ON RECOMMENDATIONS
CONTAINED IN GAO DRAFT REPORT ENTITLED
"TAX ADMINISTRATION: QUALITY OF VARIOUS IRS
SERVICE CENTER ACTIVITIES"

Recommendation:

IRS should use an indicator to measure returns processing quality that identifies the extent to which returns are being sent to error resolution specifically because of errors made by service center staff in processing returns.

Comment:

We agree that the quality of returns processing should not be measured solely by tracking the number of individual returns that go to error resolution. As a result, Returns Processing is developing a system that will enable us to separate processing errors caused by the service centers from taxpayer errors.

Recommendation:

Measure the overall quality of returns processing notices rather than just those that are referred to output review.

Comment:

In conjunction with Corporate Critical Success Factor #6 in the Service's current Strategic Business Plan, notices have been identified as one of the key Returns Processing product areas for which quality standards are to be developed during FY 1991 for application in FY 1992.

As stated in the draft report, Output Review generally reviews only those selected notices which are directed to them from the Notice Review Processing System (NRPS). However, service centers can and do conduct special reviews of Non-NRPS notices.

Recommendation:

To help fund the development of such indicators, IRS should reconsider how it spends the money currently available for service center quality control efforts.

Comment:

We agree that improvements can be made in how we spend our quality assurance dollars. Measuring employees' performance commands a majority of our quality resources since the frequency of reviews and sample size have been negotiated with our National Treasury Employees Union. We must allocate our quality resources in a manner that allows us to satisfy this obligation.

-2-

Recommendation:

IRS should assess how the quality data it now collects might be used to help build the needed indicators.

Comment:

As the report indicates, we are testing a system which would provide the types and sources of errors on all returns. However, as also stated in the report, funding for implementation may not be available even if the test proves successful.

As stated earlier, in conjunction with Corporate Critical Success Factor #6 in the Service's current Strategic Business Plan, we are taking steps to develop quality standards for key Returns Processing products and services, and are including assessments of currently available quality data for use in developing these standards.

Recommendation:

IRS should also (1) compile data on output review results in a way that enables management to identify specific problems that need to be addressed and (2) ensure that the results are so used.

Comment:

The report states that Output Review examiners did not record the type of error found on the notice nor the source of the error, and that Output Review did not have procedures for routinely recording the causes of erroneous notices and providing feedback to the service center units that originated the notice.

While this is true for Output Review examiners, our Program Analysis System (PAS) operation in Output Review is designed to review for the specific location of errors on the notices and identify who, what, where and why the error was made; either service center area or taxpayer. The PAS Output Review manual instructs the analysts to code for Pipeline and Non-pipeline functions that may be causing bad notices. This information is used to identify systemic errors and their causes.

While our PAS operation in Output Review is limited to those items which are actually processed by Output Review, we will consider approaches to reviewing not only the functional area of Output Review but all notices. It would mean the expansion of PAS Output Review and would require additional sampling and staff years. This could result in a duplication of review (i.e., sampling of notices at one point in the notice stream, and again if those notices end up in Output Review).

-3-

The report correctly observes that the weekly notice disposition reports submitted by the service centers to the National Office are not being used to adjust the Notice Review Processing System criteria. However, National Office has not used these reports because we discovered a programming deficiency which resulted in incorrect cumulative data. Returns Processing and Information Systems Management have redesigned the entire notice disposition reporting system and have scheduled a test to take place at Andover Service Center in July 1991.

ee comment 1

General Comments:

The following items, although not cited by GAO, also contributed to the improvement in the quality of our correspondence:

Correspondex Letter File

Rewritten and reorganized files of IDRS "C" letters were implemented in early 1990. This file now provides a more manageable and usable system. New letters allow for input of tax examiner name, telephone number, and tour of duty hours. The system provides:

- Improved indices
- Ability to address multiple issues
- Pre-composed paragraphs (Review aid)
- Delayed printing to allow for final review
- Facilitated letter correction process

Professional Letter System (PLS)

The PLS was tested in two service centers in 1990, and implemented nationwide in February 1991. The PLS utilizes personal computers and word processing programs to create letters. Tax examiners prepare correspondence action sheet; typist prepares and prints letter.

This decentralized printing of letter allows:

- Proofreading by tax examiners
- Tax examiner signatures on letters
- Proper attachment of enclosures
- Attachment of letter copy to the case file
- Printing professional looking letter

The following is GAO's comment on the Internal Revenue Service's letter dated March 11, 1991.

GAO Comment

1. We did not discuss these two items in our report because (1) we were not trying to discuss everything that might have had an impact on improved correspondence quality; rather, we wanted to highlight some of those items that were cited by officials we talked with during our study and (2) implementation of the Professional Letter System was outside the February to August 1990 period that we use in the report to show the improvement in correspondence quality.

Major Contributors to This Report

General Government Division, Washington, D.C.

David J. Attianese, Assistant Director, Tax Policy and Administration Issues

Cincinnati Regional Office

Homer N. Carrington, Evaluator-in-Charge
Robert I. Lidman, Assignment Manager
Daniel J. Meadows, Evaluator