July 2005

TAX ADMINISTRATION

IRS Needs Better Strategic Planning and Evaluation of Taxpayer Assistance Training

GAO

Accountability ★ Integrity ★ Reliability

TAX ADMINISTRATION

IRS Needs Better Strategic Planning and Evaluation of Taxpayer Assistance Training

Why GAO Did This Study

Millions of taxpayers ask IRS questions about tax law each year. While the accuracy of IRS's answers has improved in some cases, it is still not always what taxpayers or Congress expect. Concerns about accuracy have raised questions about the adequacy of the training IRS provides to its taxpayer assistance staff. Because of these questions, GAO was asked to assess the extent to which IRS's planning and evaluation of its taxpayer assistor training conformed to guidance published by GAO and others. Planning and evaluation are part of a feedback loop whereby lessons from one year can be applied to making improvements in future years.

What GAO Recommends

GAO is making several recommendations to IRS. Regarding planning, our recommendations include:

- establishing long-term goals;
- determining the relative importance of the factors that affect accuracy, including training; and
- benchmarking training against other organizations.

Regarding evaluation, we recommend that IRS continue to pursue evaluations of the impact of its taxpayer assistance training efforts on accuracy.

IRS agreed with five of our eight recommendations, including those summarized above, and partially responded to the remaining three.

www.gao.gov/cgi-bin/getrpt?GAO-05-782.

To view the full product, including the scope and methodology, click on the link above. For more information, contact James R. White at (202) 512-9110 or whitej@gao.gov.

What GAO Found

IRS devotes significant resources to training its tax law assistors to answer questions, by telephone and at walk-in sites, and prepare tax returns. Although IRS cannot separate the costs of training tax law assistors from other assistance staff, the thousands of staff devoted to providing tax law assistance receive training each year. The training includes classroom and computer-based training on such subjects as tax law and communication. While IRS has some data on travel and course development costs associated with training, it does not have data on what is likely the largest cost component, the value of staff time devoted to tax law training.

Responsibility for training IRS taxpayer assistance staff is decentralized. IRS's Human Capital Office provides guidance and sets policy. The two divisions responsible for tax law assistance each have a human resources office with technical staff that are assigned to the various taxpayer assistance programs. Generally speaking, the taxpayer assistance programs share responsibility for planning training with human resource staff. The human resource staff are responsible for training evaluations.

IRS's planning for taxpayer assistor training could be enhanced by a more strategic approach. To their credit, some taxpayer assistance programs clearly communicated the importance of training to staff and had knowledge and skills inventories. All the units had analyses of some of the individual factors that affect accuracy, such as the quality and use of their taxpayer assistance guidance. However, none of the programs had long-term goals for either accuracy or training or had benchmarked their training efforts against those of other organizations. Nor had they done a combined analysis of the major factors that affect accuracy in order to determine their relative importance. Setting long-term goals and analyzing training needs and relative impacts are important steps in strategic planning. Goals can provide a yardstick for measuring progress and benchmarking can help identify best practices. Analyses of relative impacts can help IRS make informed decisions about a strategy for improving accuracy, including the importance of training compared to other factors in that strategy.

The taxpayer assistance programs routinely conducted evaluations of their training efforts. However, with one exception, the evaluations did not include analyses of the impact of training on the accuracy of assistance. Instead, the units conducted less sophisticated analyses of more immediate impacts, such as trainees' satisfaction. Given the importance of accurate answers to taxpayers' questions and the resources spent on training, the four assistance programs would benefit from more sophisticated evaluations of the effectiveness of training. One program had recognized the potential value of a more sophisticated evaluation of training and pilot tested an analysis in 2004. The value of evaluation is that it provides feedback about the effectiveness of one year's training that can be used to plan improvements to future training.

United States Government Accountability Office

Contents

Tables

Figures

GAO

Accountability * Integrity * Reliability

United States Government Accountability Office
Washington, DC 20548

July 11, 2005

The Honorable Max Baucus
Ranking Minority Member
Committee on Finance
United States Senate

The Honorable Byron L. Dorgan
United States Senate

Taxpayers expect timely and accurate assistance from the Internal
Revenue Service (IRS) when they have tax law questions or have tax
returns prepared. The quality of IRS's assistance can reduce the time and
aggravation of preparing tax returns and increase taxpayers' compliance
with the tax laws. In 2004, IRS answered almost 9 million tax law
questions by telephone and prepared almost half a million tax returns for
low income taxpayers.

IRS's past performance has shown that taxpayers cannot always rely on it
to provide accurate information. While the accuracy of IRS's taxpayer
assistance has improved in some cases, it has been inconsistent or below
expectations in others. After 2 years of decline, in the first weeks of the
2005 filing season IRS telephone assistance accuracy was estimated at 87,
percent compared to 76 percent for the same time period in 2004.
Although data on the accuracy of assistance at IRS's walk-in sites are
limited by not being representative, Treasury Inspector General for Tax
Administration (TIGTA) reports have raised concerns about the accuracy
of the returns IRS prepares and that IRS had not met its annual tax law
accuracy goal.[1]

This performance has raised questions about the adequacy of the training
IRS provides its taxpayer assistance staff. Although a number of factors
can affect the accuracy of assistance IRS provides taxpayers, effective
training and development programs can enhance an organization's ability

---

[1]Treasury Inspector General for Tax Administration, *Improvements Are Needed to Ensure
Tax Returns Are Correctly Prepared at Taxpayer Assistance Centers*, Reference No. 2004-
40-025 (Washington, D.C.: 2003) and Treasury Inspector General for Tax Administration,
*Customer Service at the Taxpayer Assistance Centers Is Improving but Is Still Not
Meeting Expectations*, Reference No. 2005-40-021 (Washington, D.C.: 2004).

GAO-05-782 Tax Administration

to achieve its mission and goals, such as improving accuracy. At the same time training and developing staff is costly, making it important that such investments are targeted strategically and not wasted on efforts that are irrelevant, duplicative, or ineffective.

Because of your interest in ensuring that taxpayers receive accurate information when they contact IRS for assistance, and the contribution that training makes to that end, you asked that we review how IRS plans and evaluates its primary taxpayer assistance training and development efforts. We focused on planning and evaluation from a strategic perspective, that is, how planning and evaluation of training can help improve accuracy. As discussed with your offices, our objectives were to assess whether IRS's processes for planning and evaluating the training and development of taxpayer assistance staff conform to published guidance.

In conducting our work we developed detailed criteria to assess how IRS plans and evaluates its training and development of taxpayer assistance staff based on guidance we published for assessing strategic training and development efforts in the federal government.[2] We developed 27 separate criteria for planning and evaluation, such as establishing goals, conducting knowledge and skills needs analyses, benchmarking against other organizations, systematically collecting data, and comparing benefits and costs. We shared these criteria with officials in IRS's Human Capital Office and they said the criteria are appropriate and consistent with their policy guidance. We assessed training and development for four types of assistance: less complex tax law questions answered by phone, more complex tax law questions answered by phone, tax law questions answered at IRS walk-in sites, and return preparation at IRS walk-in sites. We collected documents describing IRS's planning and evaluation process for the four taxpayer assistance programs, interviewed officials in IRS's Wage and Investment (W&I) and Small Business/Self-employed (SB/SE) divisions to get more detail, and compared IRS's practices to our criteria. The scope and methodology section provides more details.

We conducted our work from May 2004 through May 2005 in accordance with generally accepted government auditing standards.

---

[2]GAO, *Human Capital: A Guide for Assessing Strategic Training and Development Efforts in the Federal Government*, GAO-04-546G (Washington, D.C.: March 2004).

Results in Brief

IRS's planning for taxpayer assistance training could be enhanced by long-term goals and analyses of the relative importance of the factors that affect accuracy, other organizations' experiences, and gaps between long-term needs and existing skills. To their credit, all four taxpayer assistance programs had efforts that conformed, at least in part, to many of our planning criteria. They had efforts to involve stakeholders in key annual planning decisions, communicate the importance of training to staff, and plan changes to their training to address annual tax law changes. The programs also had analyses of some of the individual factors that affect accuracy, such as the quality and use of their taxpayer assistance guidance. Two programs had annual goals for accuracy. However, the programs did not conform to other criteria. None of the programs had long-term accuracy goals, training goals, or measures suitable for an assessment of the impact of training on accuracy. Nor had they determined the relative importance of the various factors that impact accuracy, benchmarked the training practices of other organizations, or conducted assessments of long-term skill needs. Setting goals and conducting these analyses could set a direction and provide a more informed basis for planning improvements to training in order to improve accuracy.

The four taxpayer assistance programs routinely conducted evaluations of their training efforts, but these efforts did not fully comply with criteria for strategic evaluations. IRS's Human Capital Office has policy guidance urging such evaluations. One program met some of the criteria for evaluating the impact of training on accuracy. Given the importance of accurate answers to taxpayers' questions and the resources spent on training, the four assistance programs would benefit from more sophisticated evaluations of the effectiveness of training. One program had recognized the potential value of a more sophisticated evaluation of training and pilot tested such an analysis in 2004 and plans to conduct a similar analysis in 2005.

We are making recommendations to the Commissioner of Internal Revenue to improve IRS's planning and evaluations of taxpayer assistance training and development efforts. The recommendations include establishing long-term goals, determining the importance of the various factors that affect accuracy, and conducting long-term skills and knowledge gap analyses.

In commenting on a draft of this report (see app. XII), the Commissioner of Internal Revenue agreed with five of our eight recommendations and partially responded to the remaining three recommendations.

Background

IRS provides tax law assistance to taxpayers through IRS's toll-free telephone lines and tax law and return preparation assistance in person at IRS taxpayer assistance centers (formerly known as "walk-in sites") nationwide, in addition to other means. The four taxpayer assistance programs we reviewed were:

- **W&I Customer Accounts Services (CAS) general tax law assistance by telephone**. Many taxpayers contact IRS by calling its toll-free telephone number which is operated and staffed by CAS.[3] According to IRS fiscal year 2004 data, 3,420 CAS staff[4] handled more than 8.7 million telephone calls from taxpayers with general, less complex tax law questions. We did not verify these IRS data or other IRS data on taxpayer assistance programs' workload and staffing.
- **SB/SE Taxpayer Education Communication (TEC) complex tax law assistance by phone**. According to IRS officials, during the 2005 filing season and in previous years, TEC staff supported telephone service by answering taxpayers' questions on selected, more complex tax topics. When taxpayers called IRS's toll-free number about these topics, W&I CAS staff recorded the taxpayers' contact information and questions so that TEC staff could call the taxpayers back later to provide answers. According to IRS, in the 2004 filing season, TEC handled about 320,000 telephone calls for complex tax law assistance. According to IRS, TEC trained approximately 400 staff and used 272,212 staff hours to provide telephone assistance in 2004. In 2006 these calls will be handled by W&I CAS staff.
- **W&I Customer Assistance, Relationships, and Education (CARE) walk-in tax law assistance**. According to IRS, about 1.4 million staff hours were devoted to providing walk-in assistance in fiscal year 2004. Also according to IRS, 1,654 staff were trained to provide tax law assistance in 2004. In 2004, there were about 7.7 million contacts with taxpayers at IRS's approximately 400 taxpayer assistance centers. IRS did not have information on how many of these contacts were for tax law assistance. Many of the contacts were for other services, such as tax forms, publications, or accounts issues.
- **W&I Customer Assistance, Relationships, and Education (CARE) walk-in return preparation assistance**. According to IRS data, IRS staff prepared 476,813 tax returns in fiscal year 2004. IRS did not have data on

---

[3]Millions of taxpayers use IRS's Web site for taxpayer assistance.

[4]This is the total number of staff that were assigned to provide tax law assistance by telephone. The staff also had other duties. IRS could not provide data on the time staff spent providing telephone assistance.

the number of staff that prepared tax returns at taxpayer assistance centers.

IRS does not have data on the amount of time its assistors spend annually being trained on tax law.[5] However, all staff providing tax law assistance receive some training each year. IRS has continuing professional education and refresher training requirements for all its taxpayer assistance staff.

Tax law assistance staff receive a variety of training. According to an IRS official, the staff get primarily classroom training but also receive computer-based training. On-the-job training, managerial coaching, and workshops are also part of training. Taxpayer assistance staff receive training on technical tax law topics, how to use IRS systems and guidance to answer questions, and communication.

Responsibility for training and developing IRS's tax law assistance staff is decentralized. IRS's Human Capital Office provides guidance and sets policy and standards on training and development for the agency. W&I and SB/SE each have a human resources office with Learning and Education (L&E) staff who are assigned to the taxpayer assistance programs. L&E staff provide program staff advice and analysis on related policies and issues and formulate strategies, procedures, and practices to address the programs' human capital needs, including training. Generally speaking, the taxpayer assistance program offices identify training needs, L&E staff work with program staff to develop and fund annual training plans, and the program offices administer training. According to IRS policy, L&E staff are responsible for evaluating training.

According to fiscal year 2004 IRS data, IRS invested about $7 million in training W&I CAS and CARE staff to provide taxpayer assistance, including such expenses as travel, supplies, contractor fees, and development costs. IRS could not separate these costs into amounts spent on tax law training and other topics. According to SB/SE TEC fiscal year 2004 data, training costs were $325,072 in student and instructor travel. However, the cost data IRS provided did not include the costs of assistors' time associated with training. As with the staffing and workload data, we did not verify the accuracy of IRS's training cost data.

---

[5]In commenting on a draft of this report, an IRS official noted that IRS tracks training time, but does not have time data specific to tax law training.

In March 2004, we issued an assessment guide that introduced a framework for evaluating a federal agency's training and development efforts.[6] This assessment guide consists of a set of principles and key questions that federal agencies can use to ensure that their training and development investments are targeted strategically and are not wasted on efforts that are irrelevant, duplicative, or ineffective. As detailed in our assessment guide, the training and development process can loosely be segmented into four broad, interrelated components: (1) planning/front-end analysis, (2) design/development, (3) implementation, and (4) evaluation. Figure 1 depicts the general relationships between the four components, including the feedback loop between evaluation and planning. Planning and evaluation are highlighted because they are the focus of this report. Although these components can be discussed separately, they are not mutually exclusive and encompass subcomponents that may blend with one another. For instance, evaluation is part of the planning as organizations should reach agreement up front on how the success of training and development efforts will be assessed.

---

[6]GAO-04-546G.

Figure 1: Planning and Evaluation as Part of Four Components of the Training and Development Process

Planning/
Front-end Analysis

• Develop a strategic approach that establishes priorities and leverages investments in training and development to achieve agency results.

Design/
Development

• Identify specific training and development initiatives that, in conjunction with other strategies, improve individual and agency performance.

Implementation

• Ensure effective and efficient delivery of training and development opportunities in an environment that supports learning and change.

Evaluation

• Demonstrate how training and development efforts contribute to improved performance and results.

Source: GAO.

TIGTA is conducting a review that covers some aspects of the design and implementation of training.

Scope and Methodology

Our work examined the training and development of employees who provide tax law and return preparation assistance to taxpayers over the telephone and at walk-in centers, and covered both seasonal and full-time employees in IRS's W&I and SB/SE divisions. Our assessment of IRS's training and development program for taxpayer assistance employees was

based on analyses of IRS data and interviews with IRS officials. We also obtained background information from sources outside IRS such as TIGTA, the IRS Oversight Board, and the IRS National Taxpayer Advocate.

To assess how IRS plans for and evaluates the training and development of IRS employees who provide tax law and return preparation assistance to taxpayers over the telephone and at walk-in centers, we used our guide for assessing training and development programs in the federal government[7] as a framework.

We used the parts of the GAO guide dealing with the planning and evaluation components, along with supplemental guidance from the Office of Personnel Management, to identify the detailed strategic training and development criteria applicable to IRS and organized them into chronological phases. We focused our criteria on strategic planning and evaluation, that is, planning and evaluation intended to help achieve IRS's accuracy goals. We developed 27 criteria, listed in appendices I and II. We also developed examples of evidence that would demonstrate conformance with each of the criteria. For example, for the criterion "conduct a knowledge and skills inventory to identify employees' current skills and knowledge," we looked for evidence such as a completed knowledge and skills surveys and proficiency tests (see app. I for the planning phase criteria and further evidence examples and app. II for the same information on evaluation.) We shared the criteria with officials in IRS's Human Capital Office and the taxpayer assistance programs and assigned L&E staff. The officials from the Human Capital Office said the criteria are appropriate and consistent with their policy guidance. Some program and L&E officials expressed concerns about their need and ability to satisfy the criteria. Their concerns are discussed in the planning and evaluation sections of this report.

In applying the criteria to IRS, we collected documents describing IRS's planning and evaluation processes for the four types of assistance. We also interviewed officials responsible for the taxpayer assistance programs in W&I CAS and CARE units and SB/SE TEC unit and associated L&E staff responsible for training planning and evaluation, to get more details where needed. We reviewed all the evidence and made a judgment about the extent to which IRS's practices conformed to the criteria. We then discussed our initial assessments with IRS officials responsible for

---

[7]GAO-04-546G.

planning and evaluating the taxpayer assistance training and development programs who, for some of our preliminary assessments, provided additional written evidence for us to consider in making our final assessments. We then revised our initial assessments based on the evidence they provided. We made our assessments in two steps. First, an analyst reviewed all the evidence and made a judgment about the extent to which it conformed to the criteria. Then a second analyst independently reviewed the assessments. The evidence supporting our assessments is provided in appendices III through X.

Much of the information we relied on was descriptive. We determined that the information was reliable for our purposes. To the extent possible we corroborated interview evidence with documentation. Where not possible, the description is attributable to IRS officials. Where relevant we corroborated that policy guidance, such as *Internal Revenue Manual* guidance, was being implemented by collecting documentation and reports showing implementation. With respect to controls over databases, we reviewed documentation of the controls but did not assess their adequacy or test data in the databases.

We conducted our work at the Wage and Investment Division headquarters in Atlanta, Ga.; the Small Business/Self-Employed Division offices in Philadelphia, Pa.; and IRS headquarters Human Capital Office in Washington, D.C. from May 2004 through May 2005 in accordance with generally accepted government auditing standards.

Long-term Goals and Analyses of Needs and Impacts Could Improve IRS's Planning for Training and Development

Summarizing table 1, IRS's planning for taxpayer assistance training primarily focuses on meeting short-term needs, such as the challenge of training staff about tax law changes in preparation for the next year's filing season. Planning could be enhanced by long-term goals and analyses of the relative importance of the factors that affect accuracy, other organizations' experiences, and gaps between long-term needs and existing skills. Table 1 shows our assessments of planning for training by IRS's four primary taxpayer assistance programs for each phase of the planning process: goals and priorities, information gathering and assessments, skills and knowledge analysis, and strategy development and selection. The evidence supporting our assessments is shown in appendices III through VI.

Table 1: Assessments of IRS's Planning Practices in Training and Developing Staff to Provide Accurate Taxpayer Assistance

Planning phase and criteria	Extent to which IRS's practices conform to strategic guidance by type of service			
	Telephone tax law		Walk-in assistance	
	General (CAS)	More complex (TEC)	Tax law (CARE)	Return preparation (CARE)
Phase: goals and priorities				
Establish quantitative long-term accuracy goals that link to IRS's strategic goals	◐	○	◐	○
Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term planning decisions	◐	◐	◐	◐
Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	○	○	○	○
Affirm, through upper-level management communication, the importance of training and development to improve accuracy	◐	◐	●	●
Phase: information gathering and assessments				
Analytically determine and track the strategic and operational factors, including training, that affect accuracy	◐	◐	◐	◐
Incorporate evaluation results into training and development planning	See table 2 for our assessments			
Benchmark the training and development program against high-performing organizations	○	○	○	○
Assess whether human resource information systems provide needed data in a timely manner	It is too early for an assessment on the human resource information. For details see appendices III through VI.			
Phase: skills and knowledge analysis				
Perform a needs assessment to determine the knowledge and skills needed now and in the future	◐	◐	◐	◐
Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	●	○	●	●
Perform a gap analysis to determine the differences between current and needed skills and knowledge	◐	○	◐	◐
Phase: strategy development and selection				
Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	○	○	○	○
Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	◐	◐	◐	◐

Planning phase and criteria	Extent to which IRS's practices conform to strategic guidance by type of service			
	Telephone tax law		Walk-in assistance	
	General (CAS)	More complex (TEC)	Tax law (CARE)	Return preparation (CARE)
Consider the anticipated costs and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development	○	○	○	○
Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses would be used	See table 2 for our assessments			
Establish process to ensure that strategic and tactical changes can be incorporated into T&D efforts	◑	◑	◑	◑

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Goals and Priorities

In the goals and priorities phase, the four units were relatively strong in involving key stakeholders and communicating the importance of training. The programs involved key stakeholders in annual training planning decisions, shown by half-circles in table 1. However, they did not have long-term planning processes in which to involve stakeholders. All four programs had some upper-level management efforts to communicate to staff the importance of training and development in improving accuracy. CARE's was a model of communicating to all levels of staff the importance of training. For example, CARE's management guidance included managers responsible, communication vehicles, key dates, and the message to be conveyed. The other programs communicated with managers, but did not communicate more widely through the organization.

With respect to goals, IRS does not have long-term goals, as opposed to annual goals, for accuracy; nor do the four programs have goals for training and development or measures of the impact of training on accuracy. This observation about the lack of goals for accuracy is consistent with other recent reports where we cited a lack of IRS

long-term goals, as well as reports and other assessments by OMB and TIGTA.[8] Two of the programs had annual goals, shown by half-circles.

Some taxpayer assistance officials that we talked with questioned the need for long-term accuracy goals or training and development goals and measures. For example, W&I CAS telephone assistance officials said setting long-term accuracy goals is not necessary because their annual accuracy goals would move IRS toward improved performance. Program officials also said that since staff training is considered a part of doing business and is not managed as a program, training goals are unnecessary. Further, some said that, given the number of factors that affect accuracy, developing measures of training effectiveness would not be possible.

However, without long-term accuracy goals, IRS, Congress, and others are hampered in evaluating whether IRS is making satisfactory progress improving taxpayers' service. As we said in a prior report on IRS's telephone assistance program, a long-term, results-oriented goal would provide a meaningful sense of direction as well as a yardstick for measuring the results of operations and evaluating the extent of improvements resulting from changes in resources, new technology, or management of human capital.[9] Similarly, goals and performance measures for training would provide direction and help measure progress. We recognize that collecting performance data can sometimes be challenging, as discussed in the evaluation section. One IRS taxpayer assistance program has been collecting performance data suitable for determining the effectiveness of training.

Information Gathering and Assessments

In this phase, the four programs determined and tracked selected factors that affect accuracy, but would benefit from a more complete analysis as well as benchmarking. IRS has identified, and tracked on an individual basis, selected factors that affect accuracy. One nontraining factor that IRS has tracked and analyzed is the use and quality of the *Probe and Response Guide*, a manual provided to taxpayer assistance staff intended

---

[8]For example, see GAO, *Internal Revenue Service: Assessment of Fiscal Year 2006 Budget Request and Interim Results of the 2005 Filing Season*, GAO-05-416T (Washington D.C.: Apr. 14, 2005); and Treasury Inspector General For Tax Administration, *The Accounts Management Program Has Annual Performance Goals But Should Develop Long-term Performance Goals*, Reference No. 2005-40-079 (Washington D.C: May 6, 2005).

[9]U.S. General Accounting Office, *IRS Telephone Assistance: Opportunities to Improve Human Capital Management*, GAO-01-144 (Washington, D.C.: Jan. 30, 2001).

to guide them in responding to taxpayers' questions. IRS blamed decreases in accuracy in 2003 and 2004 on problems with the *Probe and Response Guide*. According to IRS officials, they used their analysis of the *Probe and Response Guide* to make the guidance more usable by taxpayer assistance staff. Officials attribute improvements in accuracy in 2005, at least in part, to this effort.

However, IRS has not conducted an analysis of the factors that affect accuracy, including training—shown by half-circles in Table 1. Specifically, IRS has not done an analysis to determine the relative importance of the various factors, including training, that affect accuracy. Determining the relative importance of the factors that affect accuracy is important because there are multiple factors. Some factors that affect accuracy are strategic and outside of IRS's direct control such as tax law changes. Other factors are operational and subject to IRS control such as manuals, information systems, and training. Without an understanding of the relative importance each factor has on accuracy, it is difficult for IRS to make informed decisions about a strategy for improving accuracy, including training's role in that strategy. Because of the purpose of the analysis and the number of factors involved, the analysis might not give a precise measure of each factor's impact. Furthermore, such analyses may be conducted on an occasional basis.

None of the taxpayer assistance programs collected another type of information—best practices of other organizations learned through benchmarking. Some officials told us it would not be possible to benchmark training programs because IRS has a unique mission. However, the telephone taxpayer assistance programs have benchmarked other aspects of their operations such as performance measures. Also, as we have stated in another report, many processes that seem unique to the government actually have counterparts in the private sector.[10] Looking at processes in dissimilar organizations can lead to rewarding improvements because it stimulates new thinking about traditional approaches to doing work. For example, in the report cited above, we noted that Xerox benchmarked L.L. Bean to improve order fulfillment.

Still another type of information useful for planning can be obtained from evaluations of training efforts. As we discuss in the background,

---

[10]GAO, *Tax Administration: Planning for IRS's Enforcement Process Changes Included Many Key Steps But Can Be Improved*, GAO-04-287 (Washington, D.C.: Jan. 20, 2004).

evaluation can provide useful feedback about the impact of existing training to help plan improvements to training. Evaluation of IRS's training efforts is the subject of our second section.

Skills and Knowledge Analysis

All four assistance taxpayer assistance programs conducted annual needs assessments. These assessments identified the knowledge and skills assistors needed for the following year. None of them had longer term analyses to project future needs. Determining skills and knowledge needed is challenging for IRS, especially when IRS must react quickly to tax law changes. For example, Congress passed a law in early January 2005 allowing taxpayers to deduct on either their 2004 or 2005 tax returns contributions made during January 2005 for relief of the victims of the December 2004 Indian Ocean Tsunami.[11] According to an IRS official, IRS had to react to this change, which took place after the filing season had already begun, by quickly alerting staff and providing them the information necessary to assist taxpayers who had questions about the deduction.

However, none of the four programs had analyses of long-term needs, such as improved proficiency in doing research, which is the basis for the half-circles in table 1. Longer-term analyses of needs could help the four programs better plan strategies to meet future needs. Such planning could help ensure that the programs acquire, train, and retain the needed staff.

Three of the four programs had knowledge and skills inventories based on actual testing to determine employees' proficiency and knowledge levels. TEC was the exception. TEC assumes the employees have an underlying technical foundation because of their background in IRS. However, in a recent TEC survey of 133 employees, the percentage of employees who self-reported having pre-existing skills or being fully proficient in several technical categories, such as depreciation, sale of property, and trusts and fiduciaries, ranged from 31 percent to 62 percent.

Three of the four programs performed gap analyses to determine the difference between current and needed skills and knowledge. Because of the above limitations in needs analyses, the gap analysis was necessarily annual not long term. Again, TEC was the exception. Because TEC had not done a needs analysis, it could not do a gap analysis.

---

[11]Pub. L. No. 109-1 (2005).

Strategy Development and Selection

Weaknesses in analyses, described in the previous sections, hampered the four programs' ability to develop criteria for when to use training, as opposed to other strategies, for filling knowledge and skills gaps, such as hiring and retention efforts. In addition, the four programs lacked information about the benefits and costs of their training efforts. As noted in the background section, IRS trains several thousand tax law assistance staff annually but does not have data on the cost of this significant effort. Without an understanding of the usefulness of other strategies, and without an understanding of the benefits and costs of training, IRS lacked information that would be useful for making resource allocation decisions. This matters because the resources devoted to training are significant. With better information to select and develop a strategy, IRS might be able to improve accuracy and save resources.

Evaluating the Impact of Training on Accuracy Could Provide a Basis for Future Improvements

Given the importance of accurate answers to taxpayers' questions and the resources spent on training, the four assistance programs would benefit from more sophisticated evaluations of the effectiveness of training. The four programs all conduct some evaluations of their training efforts. However, only TEC attempted an evaluation of the impact of training on accuracy. In table 2, the lack of evaluation plans in the other three programs had a cascading effect that left the programs generally unable to fully satisfy our assessment criteria in subsequent evaluation phases—data collection, data analysis, and application of evaluation results. Appendices VII through X show the evidence supporting our assessments in table 2.

Table 2: Assessments of IRS Practices in Evaluating Training and Development of Staff to Provide Accurate Taxpayer Assistance

Evaluation phase and criteria	Extent to which IRS's practice conforms to strategic guidance by type of service			
	Telephone tax law		Walk-in assistance	
	General (CAS)	Complex (TEC)	Tax law (CARE)	Return preparation (CARE)
Phase: evaluation plan				
Establish an overall approach for evaluating the impact of training on accuracy	◑	●	◑	◑
Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	○	○	○	○
Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	◑	●	◑	◑
Phase: data collection				
Ensure that data are collected systematically and in a timely manner	◑	●	◑	◑
Ensure that data are accurate, valid, complete, and reliable	◑	◑	◑	◑
Phase: data analysis				
Analyze data collected to assess the impact of training on accuracy	◑	◑	◑	◑
Compare accuracy benefits to training costs	○	○	○	○
Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	○	○	○	○
Analyze internal and external stakeholders' assessments of training to include impact on accuracy	◑	◑	◑	◑
Phase: application of evaluation results				
Evaluate training program as a whole, in addition to individual course evaluations, and document the results	○	◑	○	○
Incorporate evaluation results into the training and development program to improve accuracy	◑	◑	◑	◑

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Evaluation Plan Phase

IRS adopted a four-level model based on the widely accepted Kirkpatrick model[12] for evaluating training. Under this model, the sophistication of analysis increases as the numerical level of analysis increases:

- Level 1 — Reaction. The goal is to measure participants' reaction to training, usually through questionnaires.
- Level 2 – Learning. The goal is to determine what the training participants learned through various kinds of tests administered immediately after the training is completed.
- Level 3 – Behavior. The goal is to determine if the job performance of the training participants changed in the aftermath of the training. The most common means for making this determination is the administration of a survey of trained staff and their supervisors 3 months, on average, after training is complete.
- Level 4 – Results. The goal is to determine if the training led to the desired results, in this case, improved accuracy of taxpayer assistance.

Options for level 4 analyses might include statistical correlations between training and accuracy and controlled experiments where some staff receive new training and others do not. Such ROI analyses are not a part of IRS's evaluation approach. This level is sometimes split into two levels with the fifth level—often referred to as return on investment (ROI)—representing a comparison of costs and benefits quantified in dollars.

Not all training and development programs are suitable for evaluations of their effect on organizational results. The difficulty and costs of analyzing the impact of training on accuracy need to be weighed against the benefits of such an evaluation.

However, the more significant the activity targeted by the training and development program, the greater the need for level 4 analysis. Factors to consider when deciding the appropriate level of evaluation include estimated costs of the training effort, size of the training audience, management interest, program visibility, and anticipated life span of the effort. As noted in the background, IRS devotes significant resources to assisting taxpayers and training assistance staff. Congress has also expressed great interest in improving taxpayer service, such as accuracy.

---

[12]Donald L. Kirkpatrick, *Evaluating Training Programs: The Four Levels* (San Francisco: Berrett-Koehler, 1994).

For example, much of the focus of the Internal Revenue Service Reform and Restructuring Act of 1998[13] was on in improving taxpayer service.

IRS's Human Capital Office recognizes the need to do level 4 evaluations. The office has a policy statement stating that level 4 evaluations should be done for all mission-critical training.

Despite the large investment of resources, significant congressional attention, and Human Capital Office guidance, IRS officials involved in three of the four taxpayer assistance training programs had not agreed on whether to conduct a level 4 evaluation. They conducted analyses only at lower levels, usually levels 1 or 2. By contrast, officials at L&E and TEC agreed to conduct a pilot test of a level 4 analysis in 2004 for more complex tax law questions answered by telephone. Although the pilot test was not successful in 2004—for reasons discussed below in the data analysis phase—L&E and TEC intend to use the data collection and analysis plan created for 2004 to conduct a similar evaluation in 2005.

Data Collection Phase

For three of the four taxpayer assistance programs, as shown in table 2, the lack of level 4 analyses had a cascading effect that left the programs unable to fully satisfy our assessment criteria in the subsequent evaluation phases. All four programs had controls in place to help ensure systematic, timely data collection for level 1 and level 3 evaluations. TEC also had such controls in place for its level 4 evaluation. However, the database used to store level 2 data—the Administrative Corporate Education System (ACES)—is no longer in place, and there is no replacement system planned. With the exception of TEC, the programs did not attempt to collect level 4 data. Database controls and data collection plans for TEC's level 4 pilot evaluation helped ensure the systematic and timely collection of data.

All four units had controls in place to help ensure that level 1 and level 3 data were accurate, valid, reliable, and complete.[14] Because of the lack of a level 2 database, none of the four units had controls in place for that data.

---

[13]Pub. L. No. 105-206 (1998).

[14]Our review was to determine if some basic controls were in place to ensure that data were collected timely and systematically and that data were accurate, valid, reliable, and complete. We did not assess or test the adequacy of the controls, data collection methods, or data quality or completeness.

TEC was concerned that the accuracy data, which were based on test telephone calls, did not reflect the types of calls that taxpayers actually made. TEC and IRS's quality review staff are working together to try to address these concerns.

Data Analysis Phase

All four assistance programs conducted some analyses of their training efforts and analyzed stakeholders' assessments of training to identify potential improvements to individual courses. A level 1 analysis, usually surveys of trainees' opinions about a course, was conducted for all courses. The officials also said level 2 analyses, often testing, are generally done while level 3 analyses are infrequently done.

None of the four assistance programs successfully completed a level 4 evaluation of the impact of training on accuracy, compared benefits to costs, benchmarked training evaluation, or analyzed stakeholders' assessments of the impact of training on accuracy. The 2004 TEC level 4 pilot test was not completed because the level of proficiency of TEC staff in the five technical categories reviewed was insufficient.

Application of Results Phase

All the business units applied the results of their data analysis to individual training courses, not their training efforts as a whole. For example, they used the results of their level 1 and level 3 analyses to make improvements in individual training courses and to identify training needs for the upcoming year. However, because there had been no successful evaluations of the impact of training on accuracy, evaluation results could not be used to plan a strategy to improve accuracy.

Conclusions

Training has the potential to improve the service received from IRS by millions of taxpayers. To its credit, IRS has planning processes in place to address the challenges of training staff for each year's filing season. The challenges include training staff to answer questions about annual tax law changes—changes that are often very complex. Although we do not know how much training has contributed, IRS's taxpayer assistance accuracy has improved in recent years. On the other hand, IRS's current level of accuracy remains a concern, especially the accuracy of walk-in assistance and return preparation. In addition, IRS devotes significant resources to training its tax law assistance staff.

A more strategic approach to planning and evaluation would have several benefits. Strategic planning could help managers better understand the

extent to which training, as opposed to other factors that affect accuracy, could be used to improve accuracy. Evaluations of training's impact on accuracy could help managers better understand which specific training techniques are effective and which are not. While potentially difficult to design and costly to conduct, IRS's Human Capital Office has policy guidance urging more such analyses. In addition, one taxpayer assistance program is pilot testing such an analysis. Gap analyses, evaluations, and cost-benefit comparisons might also contribute to providing training at lower cost by distinguishing between effective and ineffective training.

Recommendations

We recommend that the Commissioner of Internal Revenue take appropriate action to ensure that IRS's planning and evaluations of its taxpayer assistance training and development efforts better conform to guidance for strategically planning and evaluating the training efforts.

Specifically, in the area of planning, we recommend that IRS:

- Establish a long-term goal for the accuracy of taxpayer assistance.
- Establish goals and measures for training and development logically linked to accuracy.
- Determine and track the relative importance of the various factors, including training, that affect accuracy.
- Benchmark training and development programs against high-performing organizations.
- Conduct skills and knowledge gap analyses for all taxpayer assistance programs, to include identifying and comparing current skills to long-term skill needs.
- Consider costs, benefits, ways to mitigate risks, and the appropriate level of investment for training and development efforts.

In the area of evaluation, we recommend that IRS continue to pursue the level 4 pilot in TEC and, if that analysis is shown to be feasible, conduct level 4 evaluations for its other taxpayer assistance training and development programs. The evaluations should include the following:

- an analysis of the feasibility and cost effectiveness of alternative level 4 methodologies and a data collection and analysis plan,
- a comparison of the accuracy benefits to the costs of the training,
- benchmarking of the analytical methods and the results of the data analysis against high-performing organizations, and
- an analysis of stakeholder assessments of the impact of training on accuracy.

We also recommend that IRS replace the defunct ACES database, which had been used to store level 2 data, with another database for this purpose.

Agency Comments and Our Evaluation

The Commissioner of Internal Revenue provided written comments on a draft of this report in a letter dated July 6, 2005 (see app. XII). He said the report offers valuable insight, is timely, and has been shared with the project manager for IRS's recently-initiated effort to reengineer its learning and education processes. Of the eight recommendations we made, the Commissioner agreed with five recommendations: (1) establish long-term accuracy goals, (2) determine and track the relative importance of factors that affect accuracy, (3) benchmark training and development programs, (4) conduct level 4 evaluations, and (5) replace IRS's system for storing level 2 evaluation data.

The Commissioner partially responded to the remaining three recommendations. In commenting on our recommendation to establish training goals and measures linked to accuracy, the Commissioner focused on evaluation. Although related to after-the-fact evaluation, setting clear training goals and measures to ascertain progress toward those goals—consistent with agency goals which, in the case of taxpayer assistance, would include accuracy—is an important up-front planning step by which key stakeholders should agree on what training success is and how it will be assessed.
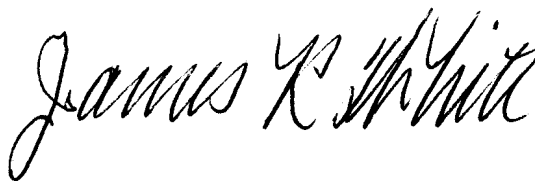
The Commissioner said that IRS recognizes the value of conducting skills and knowledge gap analyses. He summarized the programs' efforts to identify short-term skills needs, which we recognized in this report. However, he did not identify how short-term gap analyses would be conducted for the program for responding to more complex tax questions. Nor did the Commissioner discuss gap analyses to identify long-term skills needed to reach accuracy goals. Failing to conduct gap analyses, including analyses of strategic changes—such as economic, technological, and demographic developments—can hinder performance and the development of strategies that integrate new capabilities and provide flexibility to meet new challenges and improve service.

In commenting on our recommendation that IRS consider the costs and benefits of training efforts, the Commissioner's comments did not specifically mention benefits and costs but did mention unfunded needs. Given the resources dedicated to training staff and providing taxpayer

assistance, and the impact that assistance accuracy can have on taxpayers, the taxpayer assistance programs would benefit from analytically-based assurance that training efforts focus in a cost beneficial way on achieving accuracy goals.

As we agreed with your offices, unless you publicly announce the contents of this report earlier, we plan no further distribution of it until 30 days from the date of this letter. We will then send copies of this report to the Chairman of the Senate Committee on Finance, the Chairman and Ranking Minority Member of the House Committee on Ways and Means, and the Chairman and Ranking Minority Member, Subcommittee on Oversight, House Committee on Ways and Means. We will also send copies to the Secretary of the Treasury; the Commissioner of Internal Revenue; the Director, Office of Management and Budget; and other interested parties. We will also make copies available to others on request. In addition, the report will be available at no charge on the GAO Web site at http://www.gao.gov.

If you or your staff have any questions about this report, please contact me at (202) 512-9110 or whitej@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix XIII.

James R. White
Director, Tax Issues

Appendix I: What GAO Looked For: Examples of Planning Practices That Would Conform to Strategic Guidance for Training

Phase	Criterion	Conforming examples
Goals and priorities	Establish quantitative long-term accuracy goals that link to IRS's strategic goals	Authoritative quantitative targets for the level of accuracy to be achieved at a future point in time across multiple fiscal years
		A logical link between the accuracy goals and higher-level strategic goals e.g., such as to improve customer service
	Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term training planning decisions	Documented management processes to ensure that stakeholders' views are taken into account in key planning milestones and related decisions, such as developing the training and development budget request and what measures will be used to evaluate training and development
		Surveys or interview employees to determine their views and perceptions on training and development in general and more specifically on competencies and skills needed for the future
	Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	Authoritative quantitative targets related to training and development program activities to be achieved at a future point in time that are logically linked to accuracy goals
		A documented set of established measures used to assess the impacts of the training and development programs, such as on accuracy or employee skills and knowledge
		A list of methods, tools, and measures used to assess the impact of training and development, such as control groups, surveys, or a trend analysis
		Targets and goals in strategic and performance plans that establish how training and development strategies are expected to contribute to improved organizational and programmatic results
	Affirm, through upper-level management communication, the importance of training and development to improve accuracy	Demonstrated efforts to communicate throughout the organization the importance that upper-level management attaches to training and development, such as memos, letters, statements found in strategic plans, or other authoritative documents that articulate management emphasis on training and development to improve accuracy
		Comments from midlevel mangers and employees on the extent to which management communicates the importance of training, such as statements captured in an employee survey

Phase	Criterion	Conforming examples
Information gathering and assessments	Analytically determine and track the strategic and operational factors, including training, that affect accuracy	Management processes to systematically collect and analyze data that will allow IRS to comprehensively identify and report on the major factors affecting the level of accuracy achieved such as:
		Strategic factors (factors largely outside IRS's immediate control) that might include:
		• Complexity/ changes of the tax law
		• Levels of funding or budget limitations
		• Economic trends
		• Available pools of staff to select from
		Operational factors (factors within IRS's immediate control) that might may include:
		• Workload placed upon assistors
		• Changes in CSR guidance
		• Available learning tools and job aides
		• Community of practice & communication of lessons learned
		• Manager/employee relations
		• Incentive/rewards & recognition program
		Documentation that outlines the analytical methods used to identify factors affecting accuracy
		Periodic, scheduled, systematic studies to determine the key factors affecting accuracy
		Periodic regular efforts to systemically monitor the impacts strategic and operational factors have had on accuracy, such as studies of factors over an extended amount of time, such as 2-3 years, to gauge whether IRS has experienced increased or decreased levels of accuracy due to these factors
	Incorporate evaluation results into training and development planning	See appendix II
	Benchmark the training and development program against high-performing organizations	Periodic studies conducted of the training and development programs of high-performing customer service organizations providing similar services compared to identify potential improvements, to include such elements as, for example:
		• Employee skill level
		• Resources invested in training and development
		• Training and development curriculum
		• Geographic and demographic trends

Phase	Criterion	Conforming examples
	Assess whether human resource information systems provide needed data in a timely manner	Management processes to identify the user requirements for its human resource information system used in planning training and development, such as: • Identifying the user needs • Surveying managers to determine if the information provided met their needs in a timely manner
Skills and knowledge analysis	Perform a needs assessment to determine the knowledge and skills needed now and in the future	Management processes that include steps that require IRS to identify the critical knowledge and skills its employees need now and in the future to provide accurate tax law assistance, including evidence such as documentation that explicitly outlines the knowledge and skills requirements for taxpayer assistors to reach strategic accuracy goals
	Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	Management processes designed to ensure that IRS has determined the current knowledge and skills of its employees. Methods used to obtain this information may include: • Completed knowledge and skills surveys • Proficiency tests • Manager assessment
	Perform a gap analysis to determine the differences between skills and knowledge needed now and in the future	A documented analysis that compares IRS's current knowledge and skills to the knowledge and skills needed to effectively assist taxpayers now and in the future to reach accuracy goals
Strategy development and selection	Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	As part of its process for determining how to fill skill gaps and address factors that affect accuracy, or as part of its process for determining whether to implement a given training and development program, steps to: • Require that approaches other than training be considered to fill the skill gap or address the factor and • Systematically analyze and consider the relative costs and benefits of the alternative(s) to training • Develop a documented action plan or similar document that outlines when training and development interventions should take place once a knowledge and or skill gap has been identified

Phase	Criterion	Conforming examples
	Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	As part of its process for deciding which training and development programs to implement, steps to ensure efforts will target needed improvements and enhance needed skills, which would yield such items as
		A logical or explicit link between the training and development programs offered and (1) the knowledge and skills identified in the skill needs assessment or (2) the strategic and operational factors identified as affecting accuracy (to the extent that the needs and the identified factors could be addressed through training)
		A documented audit trail showing that training programs offered now and planned for the future resulted from identified needed skills and knowledge
	Consider the anticipated cost and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development	Management processes to establish and apply criteria to select the level of funding invested for training and development, such a process would yield such items as: • A comprehensive systematic effort that considers all costs related to training and development and the associated level of return linked to the investment • Historical data that show IRS reallocates training and development resources based on data derived from a cost-benefit analysis • An analysis of training and development investments from prior years and the outcomes achieved • An analysis documenting various training and development investments scenarios and the respective expected outcomes
		As part of its process for deciding which training and development efforts to implement, steps to require that the relative costs and benefits of alternative efforts be compared to current training and development efforts which would include such components as: • The relative projected impact of the efforts on needed skills, factors that need to be addressed, and accuracy levels • Training and development delivery mechanisms • Staff time involved/FTE's • Logistical options for staff travel • Products used to train and develop employees
		As part of its processes for deciding which training and development efforts to implement, steps to minimize risks by requiring that risks of alternative investments be identified and that alternative ways of mitigating those risks are identified and considered, to yield such items as a documented analysis for each proposed investment describing the details of the risks and how they could be addressed

Phase	Criterion	Conforming examples
	Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses will be used	See appendix II
	Establish a process to ensure that strategic and tactical changes can be incorporated into training and development efforts	As part of IRS's process for planning for training and development, steps to ensure decision makers incorporate a forward looking approach to planning which lay out steps to address potential strategic and operational changes the could impact IRS's training and development programs such as: • A decrease in the training and education budget • Legislative changes affecting the tax law • Unexpected staffing shortages • Technological innovations • Restructuring or a reorganization

Source: GAO.

Phase	Criterion	Conforming examples
Evaluation plan	Establish an overall approach for evaluating the impact of training on accuracy (a level 4 evaluation)	An authoritative document, or set of documents, that describes how IRS will evaluate taxpayer assistance training and development over time, to include evaluations of the impact of training on accuracy
	Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	For each line of effort to evaluate training and development impact, and for each major training and development program, a documented comparison of alternative analytical approaches to include comparisons of their practicality and potential costs and effectiveness
	Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	For each line of effort to evaluate training and development impact, and for each major training and development program, a document that describes in detail the specific data that will be collected and the analyses that will be conducted
Data collection	Ensure that data are collected systematically and in a timely manner	Documentation of protocols and systematic procedures for data collection specified in a data collection plan that includes details of methods to ensure timely collection of data
	Ensure that data are accurate, valid, complete, and reliable	For each line of effort and for each set of data collected to evaluate training and development, spot checks for accuracy and completeness, validity reports, confirmation for each set of data, tests to ensure data reliability, and/or data collection procedures designed to ensure data reliability
Data analysis	Analyze data collected to assess the impact of training on accuracy	Steps to ensure that analyses will help determine the impact of training and development on accuracy performance, such as • Analyses have considered and/or used an array of approaches (both qualitative and quantitative) • Analyses that separate training from other strategic and operational factors that might affect accuracy (to the extent that it is cost-effective and/or feasible) • Analyses enable relating processes to outputs and outcomes • Analyses allow for comparisons before and after training is taken • Analyses contain consistent qualities, allowing for tracking and comparison of processes and results over time

Phase	Criterion	Conforming examples
	Compare accuracy benefits to training costs	Steps to ensure that there are documented comparisons of the costs and benefits for each key training and development program effort to the extent deemed feasible, such as • Comparisons of benefits (including qualitative, estimated, and in some cases monetized benefits) to the costs of the training and development program • Use of forward-looking analytical approaches, such as forecasting and trend analysis, to aid in estimating and comparing future performance with and without the training intervention
	Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	Efforts to compare IRS's approach to that of high-performing organizations in such areas as training for customer assistance, including telephone and live assistance, spending on training, ways to analyze the impact of training, and the effectiveness of training
	Analyze internal and external stakeholders' assessments of training to include impact on accuracy	Efforts to analyze and consider stakeholder feedback, including feedback from internal and external obtained from such methods as: • survey results used as data in various analytical reports on training to improve accuracy • Panel studies, task forces, etc. aimed at accuracy-specific training guidance • Regular employee satisfaction surveys and focus groups focusing directly on training for accuracy
Application of evaluation results	Evaluate training program as a whole, in addition to individual course evaluations, and document the results	Reports detailing the results of analyses performed to evaluate accuracy training and development with assessments of quantitative and qualitative data pulled together in a comprehensive way to integrate conclusions from each set of data
	Incorporate evaluation results into the training and development program to improve accuracy	Authoritative documents stating how evaluation results will be used to inform, modify, and improve planning, such as • Annual performance reports comparing actual to target performance Memos, minutes, etc. from budget and training planning sessions that describe decision making based on evaluations of training initiatives aimed at specific accuracy goals Indications that the agency is making fact-based determinations of the impact of its training and development programs by using these assessments to refine or redesign training and development efforts as needed

Source: GAO.

Planning phase and criteria	Extent	Evidence summary
Phase: goals and priorities		
Establish quantitative long-term accuracy goals that link to IRS's strategic goals	◑	CAS reported an annual accuracy goal linked to IRS's strategic goal of improving taxpayer service. The accuracy goal was not long-term. CAS had accuracy projections through 2010, but according to IRS officials, these projections were only internal and subject to change. In addition, OMB stated that IRS did not have long-term goals.[1]
Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term training planning decisions	◑	CAS involved stakeholders in its annual planning decisions. For example, CAS sought input from employees and managers on training priorities and needed changes to training. However, CAS did not have a long-term training planning process for involving key stakeholders in such strategic issues as establishing training program goals and long-term employee developmental needs.
Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	○	CAS had no such goals or measures.
Affirm, through upper-level management communication, the importance of training and development to improve accuracy	◑	CAS management had efforts that conveyed the importance of training, such as an annual letter sent to field office directors that referred to the need for specific types of training. However, management did not communicate more widely throughout the organization the strategic value of training and development and its importance in achieving long-term accuracy goals.
Phase: information gathering and assessments		
Analytically determine and track the strategic and operational factors, including training, that affect accuracy	◑	CAS collected and analyzed accuracy data by using a quality review system to identify operational factors affecting accuracy, such as the proper use of IRS guidance to assistors, and the use of a pareto analysis to determine the correlation between the causes and effects of key errors. However, CAS did not analyze and track the strategic factors such as tax law changes. As a result, CAS did not have information on the impact of training on accuracy, while holding other factors such as the quality of assistors' guidance or tax law changes, constant.
Incorporate evaluation results into training and development planning	__	For our assessment, see "Application of Evaluation Results" phase in appendix VII.
Benchmark the training and development program against high-performing organizations	○	CAS had not benchmarked its training and development program. IRS had benchmarked other nontraining practices related to customer service such as performance measures.
Assess whether human resource information systems provide needed data in a timely manner	__	It was too early for an assessment because the Enterprise Learning Management System (ELMS), which would be CAS's primary system used for providing human resource information, is being implemented in fiscal year 2005.

---

[1]As reported in GAO-05-416T.

Planning phase and criteria	Extent	Evidence summary
Phase: skills and knowledge analysis		
Perform a needs assessment to determine the knowledge and skills needed now and in the future	◑	CAS identified the knowledge and skills assistors needed for the following year. For example, CAS projected the number of questions expected by tax topic and took into account recent tax law changes. However, CAS had not identified the specific knowledge and skills needed to achieve long-term accuracy goals.
Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	●	CAS had an inventory of assistors' current knowledge and skills based on testing.
Perform a gap analysis to determine the differences between skills and knowledge needed now and in the future	◑	CAS did an annual gap analysis, but because CAS had not performed a needs assessment for a period that extended beyond the next year, CAS was unable to determine the differences between current skills and knowledge and skills and knowledge needed in the future.
Phase: strategy development and selection		
Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	○	CAS had no criteria to guide decisions on using training and development strategies, as opposed to other strategies, such as improving guidance or hiring practices, to address skills and knowledge gaps.
Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	◑	CAS had practices such as employee testing, and studies conducted to identify the most frequent errors made on an annual basis, that allowed CAS to link training efforts to identified short-term skills and knowledge gaps. However, because of limitations in long-term gap analyses, CAS had limited ability to ensure training and development efforts were linked to long-term needed skills and knowledge.
Consider the anticipated cost and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development	○	CAS did not have these analyses.
Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses will be used	__	For our assessment, see "Evaluation Plan" phase in appendix VII.
Establish a process to ensure that strategic and tactical changes can be incorporated into training and development efforts	◑	CAS had efforts to anticipate and react to potential short-term changes that could affect training, such as tax law changes or lack of experienced assistors due to staff turnover. However, CAS did not have a training planning process to identify potential long-term changes such as technological innovations or changes in the economy that might impact training and development.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Appendix IV: Assessments of Planning: More Complex Tax Law Telephone Service by SB/SE[1] TEC

Planning phase and criteria	Extent	Evidence summary
Phase: goals and priorities		
Establish quantitative long-term accuracy goals that link to IRS's strategic goals	○	TEC did not have such goals.
Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term planning decisions	◑	TEC involved stakeholders in its annual training planning. For example, TEC surveyed employees to obtain input on training needs for the upcoming year. However, TEC did not have a long-term training planning process for involving key stakeholders in such strategic issues as establishing training program goals and long-term employee developmental needs.
Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	○	TEC had no such goals or measures.
Affirm, through upper-level management communication, the importance of training and development on accuracy	◑	TEC management had efforts that conveyed the importance of training, such as an annual letter sent to field office directors that referred to the need for specific types of training. However, management did not communicate more widely throughout the organization the strategic value of training and development and its importance in achieving long-term accuracy goals.
Phase: information gathering and assessments		
Analytically determine and track the strategic and operational factors, including training, that affect accuracy	◑	TEC collected and analyzed accuracy data to identify factors, such as an analysis of quality review system data. However, TEC did not analyze and track the strategic factors such as changes to tax law. As a result, TEC did not have information on the impact of training on accuracy, while holding other factors such as the quality of assistors' guidance or tax law changes, constant.
Incorporate evaluation results into training and development planning	___	For our assessment, see "Application of Evaluation Results" phase in appendix VIII.
Benchmark the training and development program against high-performing organizations	○	TEC had not benchmarked its training and development program.
Assess whether human resource information systems provide needed data in a timely manner	__	It was too early for an assessment because the Enterprise Learning Management System (ELMS), which would be TEC's primary system used for providing human resource information, is being implemented in fiscal year 2005.
Phase: skills and knowledge analysis		
Perform a needs assessment to determine the knowledge and skills needed now and in the future	◑	Annually, TEC identified the types of tax law topics it would be handling and identified the knowledge and skills assistors needed. This process included taking account of recent tax law changes. However, TEC had not identified the specific knowledge and skills needed to achieve long-term accuracy goals.

[1]In 2006, W&I will be handling these calls and training their staff to provide answers to these more complex tax law questions.

Planning phase and criteria	Extent	Evidence summary
Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	○	TEC did not conduct an inventory of its employees' current skills and knowledge. Instead, TEC assumed that its employees had an underlying technical foundation because of their prior IRS experience, and the annual training they received was to supplement previously acquired skills.
Perform a gap analysis to determine the differences between current and needed skills and knowledge now and in the future	○	Because TEC had not done a knowledge and skills inventory, a gap analysis could not be done.
Phase: strategy development and selection		
Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	○	TEC had no criteria to guide decisions on using training and development strategies, as opposed to other strategies such as improving guidance or hiring practices, to address skills and knowledge gaps.
Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	◑	In 2004, TEC matched its training courses to the general skills and knowledge assistors needed. However, because of limitations in long-term gap analyses, TEC had limited ability to ensure training and development efforts were linked to long-term needed skills and knowledge.
Consider the anticipated cost and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development	○	TEC did not have these analyses.
Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses will be used	—	For our assessment, see "Evaluation Plan" phase in appendix VIII.
Establish a process to ensure that strategic and tactical changes can be incorporated into training and development efforts	◑	TEC had informal processes for anticipating and reacting to potential short-term changes that could affect training. However, TEC did not have a training planning process to identify potential long-term changes such as technological innovations or changes in the economy that might impact training and development.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Appendix V: Assessments of Planning: Tax Law Questions Walk-In Service by W&I CARE

Planning phase and criteria	Extent	Evidence summary
Phase: goals and priorities		
Establish quantitative long-term accuracy goals that link to IRS's strategic goals	◐	CARE reported an annual accuracy goal linked to IRS's strategic goal of improving taxpayer service. The goal was not long-term. CARE had accuracy projections through 2010, but according to IRS officials, these projections were internal and subject to change. In addition, OMB stated that IRS did not have long-term goals.[1]
Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term planning decisions	◐	CARE involved stakeholders in its annual planning. For example, human capital training staff developed an annual training plan based on needs identified by program staff. However, CARE did not have a long-term training planning process in which to involve stakeholders in such strategic issues as identifying measures to assess training progress or identifying long-term employee developmental needs.
Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	○	CARE had no such goals or measures.
Affirm, through upper-level management communication, the importance of training and development to improve accuracy	●	CARE implemented a plan for communicating to various levels of staff, including front-line employees, management's vision for training. The plan included the vehicle and product to be used to communicate this message, along with key dates activities should occur to deliver the message of how training and development plays a role in providing quality taxpayer service and enhancing employee development to achieve future goals.
Phase: information gathering and assessments		
Analytically determine and track the strategic and operational factors, including training, that affect accuracy	◐	CARE collected and analyzed data, including quality review system data, to identify operational factors affecting accuracy to identify the most frequent errors. CARE also had studies to determine why those errors occurred, such as why staff did not properly use IRS guidance to answer questions. However, CARE did not analyze and track the strategic factors such as changes to the tax law or attrition. As a result, CARE did not have information on the impact these factors had on accuracy holding other factors such as changes in the tax law or change in guidance constant.
Incorporate evaluation results into training and development planning	___	For our assessment see the "Application of evaluation results" phase in appendix IX.
Benchmark the training and development program against high-performing organizations	○	CARE had not benchmarked its training and development program. IRS had benchmarked other nontraining practices related to customer service such as performance measures.
Assess whether human resource information systems provide needed data in a timely manner	___	It was too early for an assessment because the Enterprise Learning Management System (ELMS), which would be CARE's primary system used for providing human resource information, is being implemented in fiscal year 2005.

---

[1]As reported in GAO-05-416T .

Planning phase and criteria	Extent	Evidence summary
Phase: skills and knowledge analysis		
Perform a needs assessment to determine the knowledge and skills needed now and in the future	◐	CARE had identified the knowledge and skills assistors needed for the following year. One example is ensuring that assistors have the skills and knowledge needed to respond to questions about recent tax law changes. However, CARE had not identified the specific knowledge and skills needed to achieve long-term accuracy goals.
Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	●	CARE had an inventory of assistors' current knowledge and skills based on a prescreening testing process.
Perform a gap analysis to determine the differences between current and needed skills and knowledge now and in the future	◐	CARE did an annual gap analysis, but because CARE had not performed a needs assessment for a period that extended beyond the next year, it had not performed a gap analysis to determine the differences between current skills and knowledge and skills and knowledge needed in the future.
Phase: strategy development and selection		
Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	○	CARE had no criteria to guide decisions on using training and development strategies, as opposed to other strategies, such as improving guidance, or hiring practices to address skill gaps.
Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	◐	CARE has practices such as employee testing, and studies conducted to identify the most frequent errors made on an annual basis, that allow CARE to link training efforts to identified short-term skills and knowledge gaps. However because of limitations in long term gap analyses, staff have limited ability to ensure training and development efforts are linked to long term needed skills and knowledge.
Consider the anticipated costs and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development	○	CARE did not have these analyses.
Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses will be used	__	For our assessment, see "Evaluation Plan" phase in appendix IX.
Establish a process to ensure that strategic and tactical changes can be incorporated into training and development efforts	◐	CARE had efforts to anticipate and react to potential short-term changes that could affect training, such as tax law changes or lack of experienced assistors due to staff turnover. However, CARE did not have a training planning process to identify potential long-term changes, such as technological innovations or changes in the economy, that might impact training and development.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◐

Conforms to guidance to little or no extent ○

Appendix VI: Assessments of Planning: Return Preparation Walk-In Service by W&I CARE

Planning phase and criteria	Extent	Evidence summary
Phase: goals and priorities		
Establish quantitative long-term accuracy goals that link to IRS's strategic goals	○	CARE did not have quantitative long-term accuracy goals for return preparation. Fiscal year 2005 will be the baseline year for a new measure of tax preparation accuracy, so therefore there are no annual or long-term goals.
Involve key stakeholders, including human capital professionals, managers, and employees, in key long-term planning decisions	◑	CARE involves stakeholders in annual planning. For example, human capital training staff developed an annual training plan based on needs identified by program staff. However, CARE did not have a long-term training planning process in which to involve stakeholders in such strategic issues as identifying measures to assess training progress or identifying long-term employee developmental needs.
Establish goals for, and measures to assess the effectiveness of, training and development on accuracy	○	CARE had no such goals or measures.
Affirm, through upper-level management communication, the importance of training and development to improve accuracy	●	CARE implemented a plan for communicating to various levels of staff, including front-line employees, management's vision for training. The plan included the vehicle and product to be used to communicate this message, along with key dates activities should occur to deliver the message of how training and development play a role in providing quality taxpayer service and enhancing employee development to achieve future goals
Phase: information gathering and assessments		
Analytically determine and track the strategic and operational factors, including training, that affect accuracy	◑	CARE collected and analyzed accuracy data, including quality review system data, to identify operational factors affecting accuracy to identify the most frequent errors. CARE also had studies to determine why those errors occurred, such as why staff did not properly use IRS guidance to answer questions. However, CARE did not analyze and track the strategic factors such as changes to the tax law or attrition. As a result, CARE did not have information on the impact these factors will have on accuracy holding other factors such as changes in the tax law or change in guidance constant.
Incorporate evaluation results into training and development planning	__	For our assessment see the "Application of evaluation results" phase in appendix X.
Benchmark of the training and development program against high- performing organizations	○	CARE had not benchmarked its training and development program. IRS had benchmarked other nontraining practices related to customer service such as performance measures.
Assess whether human resource information systems provide needed data in a timely manner	__	It was too early for an assessment because the Enterprise Learning Management System (ELMS), which would be CARE's primary system used for providing human resource information, is being implemented in fiscal year 2005.
Phase: skills and knowledge analysis		
Perform a needs assessment to determine the knowledge and skills needed now and in the future	◑	CARE identified the knowledge and skills assistors needed for the following year. One example is ensuring that assistors have the skills and knowledge needed to respond to questions about recent tax law changes. However, CARE had not identified the specific knowledge and skills enhancements needed to achieve long-term accuracy goals.

Planning phase and criteria	Extent	Evidence summary
Conduct a knowledge and skills inventory to identify employees' current skills and knowledge	●	CARE had an inventory of assistors' current knowledge and skills based on a prescreening testing process.
Perform a gap analysis to determine the differences between current and needed skills and knowledge now and in the future	◑	CARE did an annual gap analysis, but because CARE had not performed a needs assessment for a period that extended beyond the next year, it had not performed a gap analysis to determine the differences between current skills and knowledge and skills and knowledge needed in the future.
Phase: strategy development and selection		
Develop and apply criteria to determine when to use training and development strategies to fill skills and knowledge gaps	○	CARE had no criteria to guide decisions on using training and development strategies, as opposed to other strategies, such as improving guidance, or hiring practices to address skill gaps.
Ensure training and development efforts target factors that affect accuracy and are linked to needed skills and knowledge	◑	CARE had practices such as employee testing, and studies conducted to identify the most frequent errors made on an annual basis, that allowed CARE to link training efforts to identified short-term skills and knowledge gaps.
		However, because of limitations in long-term gap analyses, CARE had limited ability to ensure training and development efforts were linked to long-term needed skills and knowledge.
Consider the anticipated costs and benefits of alternative training and development efforts, ways to mitigate associated risks, and the appropriate level of investment for training and development efforts	○	CARE did not have these analyses.
Establish a detailed plan for evaluating training, including performance measures, data, planned analyses, and how the analyses will be used	__	For our assessment, see "Evaluation Plan" phase in appendix X.
Establish a process to ensure that strategic and tactical changes can be incorporated into training and development efforts	◑	CARE had efforts to anticipate and react to potential short-term changes that could affect training, such as tax law changes or lack of experienced assistors due to staff turnover. However, CARE did not have a training process to identify potential long-term changes such as technological innovations or changes in the economy that might impact training and development.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Evaluation phase and criteria	Extent	Evidence summary
Phase: evaluation plan		
Establish an overall approach for evaluating the impact of training on accuracy (a level 4 evaluation)	◑	L&E had adopted a four-level model for evaluating training, the Evaluation Monitoring System-Integrated Training Evaluation and Measurement Services (EMS-ITEMS), which was based on the widely-accepted Kirkpatrick model. IRS's Human Capital Office L&E officials concluded that level 4 evaluations were appropriate. However, CAS and L&E officials had not agreed on whether to conduct level 4 evaluations. In addition, L&E officials had not documented an analysis of what level of evaluation was appropriate.
Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	○	Because L&E and CAS officials had not agreed to do a level 4 evaluation, L&E had not analyzed the feasibility and cost effectiveness of alternative level 4 methodologies.
Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	◑	Because L&E and CAS officials had not agreed to do a level 4 evaluation, L&E did not have a level 4 data analysis plan. However, EMS-ITEMS did provide general guidance for different levels of evaluation.
Phase: data collection		
Ensure that data are collected systematically and in a timely manner	◑	EMS-ITEMS had controls to help ensure systematic and timely collection of data for level 1 and 3 evaluations. EMS-ITEMS had guidance on the collection of data including responsible parties and timing. However, the Administrative Corporate Education System (ACES) database used to collect level 2 data was no longer in place, and as a result, there was no system to consistently collect and store level 2 data. In addition, because L&E and CAS officials had not agreed to do a level 4 evaluation, L&E had not collected the data necessary to do a level 4 evaluation.
Ensure that data are accurate, valid, complete, and reliable	◑	EMS-ITEMS had level 1 and some level 3 data and the system had controls in place to help ensure that data were accurate, valid, and reliable. However, because L&E and CAS had not agreed to do a level 4 evaluation, L&E did not collect level 4 data. The ACES database used to collect level 2 data no longer exists and there is no replacement system planned.
Phase: data analysis		
Analyze data collected to assess the impact of training on accuracy	◑	L&E conducted some level 1, 2, and 3 evaluations of participants' opinions, learning, and subsequent job performance. For example, L&E analyzed course evaluations, tests, and supervisory evaluations after employees completed coursework to identify needed improvements to training. However, L&E had done no level 4 evaluations.
Compare accuracy benefits to training costs	○	Because L&E and CAS had not agreed to do a level 4 evaluation, L&E did not do a comparison of benefits to cost.
Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	○	None.

Evaluation phase and criteria	Extent	Evidence summary
Analyze internal and external stakeholders' assessments of training to include impact on accuracy	◐	L&E and CAS analyzed stakeholder assessments of training to identify needed changes. For example, management councils and focus groups were used to determine if changes were needed to improve course material, training environment, timing, and objectives of training. However, because L&E and CAS had not agreed to do a level 4 evaluation, there were no efforts to collect and analyze assessments from internal and external stakeholders to assess training in terms of its impact on accuracy.
Phase: application of evaluation results		
Evaluate training program as a whole, in addition to individual course evaluations, and document the results	○	Evaluations were for individual courses, not for the program as a whole.
Incorporate evaluation results into the training and development program to improve accuracy	◐	L&E and CAS reported using evaluation results in planning to make improvements to training courses or to identify training needs for the upcoming year. L&E and CAS used the available level 1 through 3 evaluations to make planned improvements on individual courses. However, because L&E and CAS had not agreed to do a level 4 evaluation or cost-benefit comparisons, L&E and CAS's ability to make informed decisions on improving training to improve accuracy was limited.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◐

Conforms to guidance to little or no extent ○

Evaluation phase and practice	Extent	Evidence summary
Phase: evaluation plan		
Establish an overall approach for evaluating the impact of training on accuracy (a level 4 evaluation)	●	L&E had adopted a four-level model for evaluating training (EMS-ITEMS) that was based on the widely accepted Kirkpatrick model. In 2004, L&E and TEC planned a pilot test of a level 4 evaluation. L&E and TEC officials stated in an official document that they intend to conduct a level 4 evaluation in 2005.
Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	○	L&E had not analyzed the feasibility and cost effectiveness of alternative methodologies.
Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	●	L&E had a plan to conduct a level 4 evaluation in 2004. L&E planned to conduct a level 4 evaluation in 2005 similar to that done in 2004. Officials said they would use the 2004 plan as the basis for the 2005 evaluation effort.
Phase: data collection		
Ensure that data are collected systematically and in a timely manner	●	L&E had controls to ensure data were collected systematically and in a timely manner. Data collection practices for all four levels of evaluation in the level 4 pilot test evaluation included specific steps to be accomplished, status updates, and due dates. In addition, for the 2005 pilot, data on TEC accuracy in five technical categories and staff proficiency in those technical categories were available before the end of 2004.
Ensure that data are accurate, valid, complete, and reliable	◑	In 2004, TEC was concerned that the accuracy data in the five technical categories, based on test calls, did not reflect the types of calls taxpayers actually made. TEC and IRS's quality review staff subsequently worked together in an effort to fix the problem. EMS-ITEMS had level 1 and some level 3 data, and the system had controls in place to ensure that data were accurate, valid, and complete. However, the database used to collect level 2 data no longer existed, and there was no replacement system planned.
Phase: data analysis		
Analyze data collected to assess the impact of training on accuracy	◑	In the 2004 pilot level 4 evaluation, L&E did an analysis but was unable to assess the impact of training on accuracy, because the level of proficiency of TEC staff in the five technical categories (capital gains, depreciation and sale of business property, rentals, trust and fiduciaries, and international and alien) was insufficient. L&E decided not to measure the impact of training on accuracy until 80 percent of the employees were proficient in the five technical categories. In a survey of 133 TEC employees, the percent reported having preexisting skills or being fully proficient in the five technical categories ranged from 31 to 62 percent.
Compare accuracy benefits to training costs	○	Comparing the value of training to the costs of training was to be the third phase of the uncompleted 2004 pilot test.
Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	○	None.

Evaluation phase and practice	Extent	Evidence summary
Analyze internal and external stakeholders' assessments of training to include impact on accuracy	◑	L&E and TEC analyzed stakeholder assessments of training to identify needed changes. For example, TEC used information from employee surveys to target training materials on the types of calls assistors reported receiving. However, they did not successfully collect and analyze assessments from internal and external stakeholders to assess training in terms of its impact on accuracy.
Phase: application of evaluation results		
Evaluate training program as a whole, in addition to individual course evaluations, and document the results	◑	As discussed above, although L&E attempted to do a pilot test of a level 4 evaluation of the training program as a whole, the evaluation was unsuccessful.
Incorporate evaluation results into the training and development program to improve accuracy	◑	L&E and TEC reported using evaluation results in planning to make improvements to training courses or to identify training needs for the upcoming year. However, because there was no successful level 4 evaluation or benefit cost comparison, L&E and TEC's ability to make informed decisions on improving training to improve accuracy was limited.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Appendix IX: Assessments of Evaluation: Tax Law Walk-in Service by W&I CARE

Evaluation phase and practice	Extent	Evidence summary
Phase: evaluation plan		
Establish an overall approach for evaluating the impact of training on accuracy (a level 4 evaluation)	◑	L&E had adopted a four-level model for evaluating training, the Evaluation Monitoring System-Integrated Training Evaluation and Measurement Services (EMS-ITEMS), based on the widely-accepted Kirkpatrick model. IRS's Human Capital Office L&E officials concluded that level 4 evaluations were appropriate. However, CARE and L&E officials had not agreed on whether to conduct level 4 evaluations. In addition, L&E officials had not documented an analysis of what level of evaluation was appropriate.
Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	○	Because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E had not analyzed the feasibility and cost effectiveness of alternative level 4 methodologies. In addition, L&E and CARE officials did not look at methodologies specific to questions on tax law assistance. The approach for training and evaluating staff who provided assistance by answering walk-in customers' tax law questions and staff preparing returns were the same.
Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	◑	Because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E did not have a level 4 data analysis plan. However, EMS-ITEMS did provide general guidance for different levels of evaluation.
Phase: data collection		
Ensure that data are collected systematically and in a timely manner	◑	EMS-ITEMS had controls to help ensure systematic and timely collection of data for level 1 and 3 evaluations. EMS-ITEMS has guidance on the collection of data including responsible parties and timing. However, the Administrative Corporate Education System (ACES) database used to collect level 2 data was no longer in place, and as a result, there was no system to consistently collect and store level 2 data. In addition, because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E had not collected the data necessary to do a level 4 evaluation.
Ensure that data are accurate, valid, complete, and reliable	◑	EMS-ITEMS had level 1 and some level 3 data and the system had controls in place to help ensure that data were accurate, valid, and reliable. However, because L&E and CARE had not agreed to do a level 4 evaluation, L&E did not collect level 4 data. The ACES database used to collect level 2 data no longer exists and there was no replacement system planned.
Phase: data analysis		
Analyze data collected to assess the impact of training on accuracy	◑	L&E conducted some level 1, 2, and 3 evaluations of participants' opinions, learning, and subsequent job performance. For example, L&E analyzed course evaluations, tests, and supervisory evaluations after employees completed coursework to identify needed improvements to training. However, L&E had done no level 4 evaluations. In addition, there was no distinction between analysis of walk-in tax law and return preparation assistance.
Compare accuracy benefits to training costs	○	Because L&E and CARE had not agreed to do a level 4 evaluation, L&E did not do a comparison of benefits to cost.

Evaluation phase and practice	Extent	Evidence summary
Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	○	None.
Analyze internal and external stakeholders' assessments of training to include impact on accuracy	◖	L&E and CARE analyzed stakeholder assessments of training to identify needed changes. For example, focus groups were used to determine if changes were needed to improve course material, training environment, timing, and objectives of training. However, because L&E and CARE had not agreed to do a level 4 evaluation, there were no efforts to collect and analyze assessments from internal and external stakeholders to assess training in terms of its impact on accuracy.
Phase: application of evaluation results		
Evaluate training program as a whole, in addition to individual course evaluations, and document the results	○	Evaluations were for individual courses, not for the program as a whole.
Incorporate evaluation results into the training and development program to improve accuracy	◖	L&E and CARE reported using evaluation results in planning to make improvements to training courses or to identify training needs for the upcoming year. L&E and CARE used the available level 1 through 3 evaluations to make planned improvements on individual courses. However, because L&E and CARE had not agreed to do a level 4 evaluation or cost-benefit comparisons, L&E and CARE's ability to make informed decisions on improving training to improve accuracy was limited.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◖

Conforms to guidance to little or no extent ○

Appendix X: Assessments of Evaluation: Return Preparation Walk-in Service by W&I CARE

Evaluation phase and practice	Extent	Evidence summary
Phase: evaluation plan		
Establish an overall approach for evaluating the impact of training on accuracy (a level 4 evaluation)	◖	L&E has adopted a four-level model for evaluating training, the Evaluation Monitoring System-Integrated Training Evaluation and Measurement Services (EMS-ITEMS), based on the widely-accepted Kirkpatrick model. IRS's Human Capital Office L&E officials had concluded that level 4 evaluations were appropriate. However, CARE and L&E officials had not agreed on whether to conduct a level 4 evaluation. In addition, L&E officials had not documented an analysis of what level of evaluation was appropriate.
Systematically analyze the feasibility and cost effectiveness of alternative methodologies for evaluating the impact of training and development on accuracy	○	Because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E had not analyzed the feasibility and cost effectiveness of alternative level 4 methodologies. In addition, L&E and CARE officials did not look at methodologies specific to questions on return preparation. The approaches for training and evaluating staff that provided assistance by preparing returns and staff answering walk-in customers' tax law questions were the same.
Before implementing training, develop analysis plans, including data to be collected, analyses to be conducted, and how the analyses will be used	◖	Because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E did not have a level 4 data analysis plan. However, EMS-ITEMS did provide general guidance for different levels of evaluation.
Phase: data collection		
Ensure that data are collected systematically and in a timely manner	◖	EMS-ITEMS had controls to help ensure systematic and timely collection of data for level 1 and 3 evaluations. EMS-ITEMS had guidance on the collection of data including responsible parties and timing. However, the Administrative Corporate Education System (ACES) database used to collect level 2 data was no longer in place, and as a result, there was no system to consistently collect and store level 2 data. In addition, because L&E and CARE officials had not agreed to do a level 4 evaluation, L&E had not collected the data necessary to do a level 4 evaluation.
Ensure that data are accurate, valid, complete, and reliable	◖	EMS-ITEMS had level 1 and some level 3 data and the system had controls in place to help ensure that data were accurate, valid, and reliable. However, because L&E and CARE had not agreed to do a level 4 evaluation, L&E did not collect level 4 data. The ACES database used to collect level 2 data no longer existed and there was no replacement system planned.
Phase: data analysis		
Analyze data collected to assess the impact of training on accuracy	◖	L&E conducted some level 1, 2, and 3 evaluations of participants' opinions, learning, and subsequent job performance. For example, L&E analyzed course evaluations, tests, and supervisory evaluations after employees completed coursework to identify needed improvements to training. However, L&E had done no level 4 evaluations. In addition, there was no distinction between analysis of walk-in tax law and return preparation assistance.
Compare accuracy benefits to training costs	○	Because L&E and CARE had not agreed to do a level 4 evaluation, L&E did not do a comparison of benefits to cost.
Benchmark training cost, training outcomes (accuracy), and analytical methods against high-performing organizations	○	None.

Evaluation phase and practice	Extent	Evidence summary
Analyze internal and external stakeholders' assessments of training to include impact on accuracy	◑	L&E and CARE analyzed stakeholder assessments of training to identify needed changes. For example, focus groups were used to determine if changes were needed to improve course material, training environment, timing, and objectives of training. However, because L&E and CARE had not agreed to do a level 4 evaluation, there were no efforts to collect and analyze assessments from internal and external stakeholders to assess training in terms of its impact on accuracy.
Phase: application of evaluation results		
Evaluate training program as a whole, in addition to individual course evaluations, and document the results	○	Evaluations were for individual courses, not for the program as a whole.
Incorporate evaluation results into the training and development program to improve accuracy	◑	L&E and CARE reported using evaluation results in planning to make improvements to training courses or to identify training needs for the upcoming year. L&E and CARE used the available levels 1 through 3 evaluations to make planned improvements on individual courses. However, because L&E and CARE had not agreed to do a level 4 evaluation or cost-benefit comparisons, L&E and CARE's ability to make informed decisions on improving training to improve accuracy was limited.

Source: GAO analysis.

Key:

Conforms to guidance to a great extent ●

Conforms to guidance to some extent ◑

Conforms to guidance to little or no extent ○

Appendix XI: Information on the Kirkpatrick Model of Training and Development Evaluation

In recent years, a growing number of organizations have adopted a balanced, multilevel approach to evaluating their training and development efforts. Such an approach can help provide varied data and perspectives on the effect that training efforts have on the organization. One commonly accepted model consists of four levels of assessment.[1] The first level measures the training participants' reaction to, and satisfaction with, the training program or planned actions to use new or enhanced competencies. The second level measures the extent to which learning has occurred because of the training effort. The third level measures the application of this learning to the work environment through changes in behavior that trainees exhibit on the job because of the training or development program. The fourth level measures the impact of the training program on the agency's program or organizational results. The fourth level is sometimes split into two levels with the fifth level representing a comparison of costs and benefits quantified in dollars. This fifth level—often referred to as return on investment (ROI)—compares the benefits (quantified in dollars) to the costs of the training and development program.[2]

Not all training and development programs require, or are suitable for, higher levels of evaluation. Indeed, higher levels of evaluation can be challenging to conduct because of the difficulty and costs associated with data collection and the complexity in directly linking training and development programs to improved individual and organizational performance. Figure 2 depicts an example gradation of the extent to which an agency could use the various levels of evaluation to assess its training and development programs. For example, an agency may decide to evaluate participants' reactions for all (100 percent) of its programs, while conducting an ROI analysis for 5 percent of its programs.

---

[1]Donald L. Kirkpatrick (author of *Evaluating Training Programs: The Four Levels*) conceived a commonly recognized four-level model for evaluating training and development efforts.

[2]Jack J. Phillips conceived the fifth level in the model for evaluating training and development efforts as discussed in the book, *Measuring ROI in the Public Sector.*

**Appendix XI: Information on the Kirkpatrick
Model of Training and Development
Evaluation**

Figure 2: Example Agency's Training and Development Programs Assessed Using Each Level of Evaluation
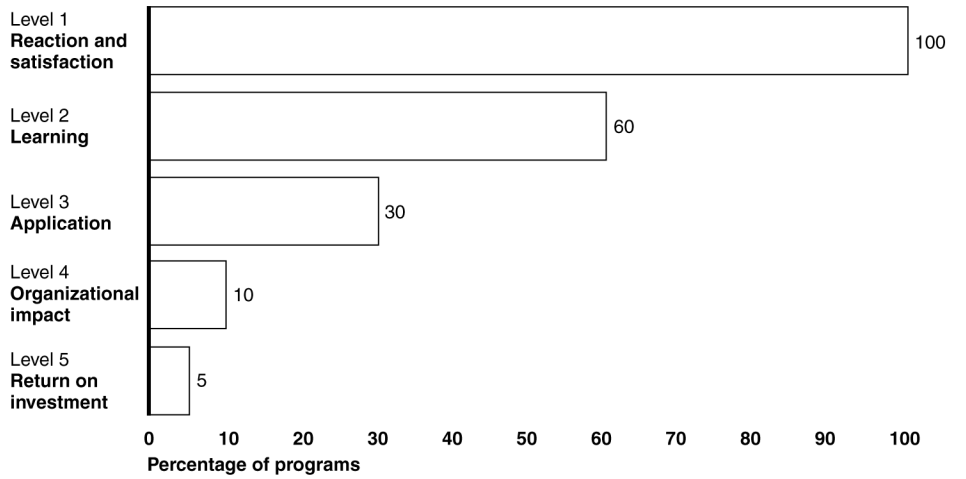
Source: Based on a table in "The public Sector Challenge: Developing a credible ROI Process," Patricia Pulliam Philips and Jack J. Phillips, *In Action: Measuring ROI in the Public Sector,* ed. Jack J. Phillips and Patricia Pulliam Phillips (Alexandria, Virginia: American Society for Training and Development, 2002).

Appendix XII: Comments from IRS

<div style="text-align:center">

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

</div>

COMMISSIONER

<div style="text-align:center">

July 6, 2005

</div>

Mr. James R. White
Director, Tax Issues
United States Government Accountability Office
Washington, D.C. 20548

Dear Mr. White:

Thank you for the opportunity to review your proposed report entitled *Tax Administration: IRS Needs Better Strategic Planning and Evaluation of Taxpayer Assistance Training* (GAO-05-782).

Your report provides an assessment of our training and development programs for employees who provide taxpayers tax law and return preparation assistance. Such assistance can be challenging given the complexities of the tax code. Despite these challenges, I believe our tax law and return preparation assistors conscientiously perform their duties based on the facts and circumstances as presented to them by taxpayers, and I am pleased that the accuracy of assistance provided to taxpayers is improving. A 14.5 percent increase (after two years of decline) in telephone assistance accuracy suggests we are on the right track. This increase will be fully validated upon examination of the relevant returns to ensure that data are complete and representative.

Your review of our training and development programs, and accompanying recommendations, offer valuable insight. As the first agency to be assessed against the Government Accountability Office's (GAO) criteria in its Guide for Assessing Strategic Training and Development Efforts in the Federal Government, we particularly appreciate the examples of planning practices that conform to strategic guidance. Your timely report will assist us in our current efforts to reengineer our learning and education business processes. The Learning and Education Analysis Project (LEAP) will result in a new training structure and achieve significant cost savings. We expect to implement the new processes by December 2006; this report has been shared with the LEAP manager.

The Internal Revenue Service's (IRS) training developers diligently implement training programs for assistors. These programs must be completed in time for the filing season, as well as incorporate numerous and frequent tax law changes made subsequent to the previous filing season. Generally, I believe our past training efforts in this challenging area have been well and appropriately planned and executed. Indeed, certain aspects have been recognized. For example, our Training Development Quality

2

Assurance System (TDQAS), which is our Servicewide approach to training and development efforts, has been benchmarked by the American Society for Training and Development and the International Society for Performance Improvement. We believe that the TDQAS is an educationally sound, comprehensive process to ensure that strategic and tactical changes are incorporated into our training and development efforts. It is incorporated into the Internal Revenue Manual (IRM) and there are extensive guidelines that supplement the system. We recognize, however, that our learning and education processes can be improved; as noted above, we will take into account your recommendations as we reengineer those processes.

If you have any questions, please contact Beverly Ortega Babers, Chief Human Capital Officer, at (202) 622-7676.

Sincerely,

Mark W. Everson

Enclosure

Our comments on the report's specific recommendations follow:

Recommendation: Establish a long-term goal for the accuracy of taxpayer assistance.

We will continue efforts to achieve further improvements in the accuracy of the assistance we provide taxpayers. We have begun the process of establishing long-term goals based on the needs of taxpayers, other benefits, and costs.

Recommendation: Establish goals and measures for training and development logically linked to accuracy.

The true measure of training effectiveness is skillful on-the-job application of knowledge. Training accuracy results in proficient performance. We are committed to Level 3 Evaluation of our tax law and return preparation assistance training programs. We developed the Evaluation Monitoring System (EMS) based on the Kirkpatrick Model as well as Integrated Training Evaluation and Measurement Services (ITEMS) which include extensive online guidance and functionalities. This is the EMS-ITEMS cited in the report; its use is covered by policy and the IRM. While I believe our established processes conform to guidance to a great extent, I agree that training evaluations should be in line with benchmarked practices and the results used to plan further improvements in the accuracy of tax law and return preparation assistance we provide taxpayers.

Recommendation: Determine and track the relative importance of the various factors, including training, that affect accuracy.

We will investigate all factors that affect the accuracy of assistance provided to taxpayers. In an environment that requires changes on short notice, our strategic approach must remain flexible.

Recommendation: Benchmark training and development programs against high-performing organizations.

We agree that benchmarking is a beneficial process, albeit a resource-intensive one. Nevertheless, we are committed to benchmarking with regard to our taxpayer assistance training programs to the extent practicable within resource limitations.

Recommendation: Conduct skills and knowledge gap analyses for all taxpayer assistance programs, to include identifying and comparing current skills to long-term skill needs.

We recognize the value of such assessments. For example, our walk-in assistance program has developed the Technical Assessment Battery (TAB) to conduct skills and

2

knowledge gap analysis for its taxpayer assistance program. This assessment tool was recently redesigned to compare employees' current skill levels to the long-term skills needed to meet customers' needs. The TAB includes competencies required for all tasks needed in the Taxpayer Assistance Centers including tax law, return preparation, account work, notice inquiries, and collection work. Our telephone centers conduct skills and knowledge gap analysis as part of their recurring planning cycles directed at the delivery of accurate customer assistance. This analysis is conducted annually and for each planning period during the year to align employee skills with customer needs. Employee skill enhancements necessitated by the introduction of new technologies or employee tools are incorporated in the implementation process for the initiative. When organizational or legislative changes occur, we conduct additional skills and knowledge gap assessments to the extent practicable within resource limitations.

Recommendation: Consider costs, benefits, ways to mitigate risks, and the appropriate level of investment for training and development efforts.

Our annual training plan formulation process allows business divisions to identify their training priorities based on strategic plans and the level of funding appropriated for that year. Every business division closely monitors training plan execution throughout the year to ensure delivery of training programs and to identify unfunded needs and funds that will not be utilized. This allows for maximum flexibility to meet changing business priorities.

Recommendation: Continue to pursue the level 4 pilot in TEC and, if that analysis is shown to be feasible, conduct level 4 evaluations for its other taxpayer assistance training and development programs.

Level 4 Evaluation as embodied in EMS was carefully developed with the assistance of expert consultants. Our approach comprises three phases: time to capability, training impact on organizational performance indicators, and establishing training worth. The approach is innovative, yet reasonable. It is consistent with GAO's criteria, except for Return on Investment, although we believe the methodology developed by the Office of Personnel Management for establishing the worth of training is more appropriate. The TEC Level 4 pilot should be completed this year. However, because our Small Business/Self-Employed Division will not be providing taxpayer assistance on tax law after this year, our Wage and Investment Division will undertake future efforts. The EMS Level 4 Evaluation approach is established as policy, and our learning and education functions intend to follow it.

Recommendation: Replace the defunct ACES database, which had been used to store level 2 data, with another database for this purpose.

We recognize the importance of Level 2 data in the training evaluation process and are looking into available options, including our new Enterprise Learning Management System, for storing such data.

Appendix XIII: GAO Contact and Staff Acknowledgments

GAO Contact

James R. White (202) 512-9110

Acknowledgments

In addition to the contact named above, Charlie Daniel, David Dornisch, Chad Gorman, Jason Jackson, Ronald Jones, Veronica Mayhand, and Katrina Taylor made key contributions to this report.

GAO's Mission	The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.
Obtaining Copies of GAO Reports and Testimony	The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."
Order by Mail or Phone	The first copy of each printed report is free. Additional copies are $2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to: U.S. Government Accountability Office 441 G Street NW, Room LM Washington, D.C. 20548 To order by Phone: Voice: (202) 512-6000 TDD: (202) 512-2537 Fax: (202) 512-6061
To Report Fraud, Waste, and Abuse in Federal Programs	Contact: Web site: www.gao.gov/fraudnet/fraudnet.htm E-mail: fraudnet@gao.gov Automated answering system: (800) 424-5454 or (202) 512-7470
Congressional Relations	Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400 U.S. Government Accountability Office, 441 G Street NW, Room 7125 Washington, D.C. 20548
Public Affairs	Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800 U.S. Government Accountability Office, 441 G Street NW, Room 7149 Washington, D.C. 20548