United States Government Accountability Office

GAO

Report to Congressional Requesters

December 2004

EARNED INCOME TAX CREDIT

Implementation of Three New Tests Proceeded Smoothly, But Tests and Evaluation Plans Were Not Fully Documented

GAO

Accountability ★ Integrity ★ Reliability

GAO-05-92

December 2004

EARNED INCOME TAX CREDIT

Implementation of Three New Tests Proceeded Smoothly, But Tests and Evaluation Plans Were Not Fully Documented

GAO Highlights

Accountability·Integrity·Reliability

Highlights of GAO-05-92, a report to Congressional Requesters

Why GAO Did This Study

Research has shown that the Earned Income Tax Credit (EITC) has helped lift millions of individuals out of poverty. In recent years, the Internal Revenue Service (IRS) has paid approximately $30 billion annually to about 20 million EITC recipients. However, the program also has experienced a high rate of noncompliance. IRS estimated that EITC overclaim rates for tax year 1999, the most recent data available, were between 27 and 32 percent of dollars claimed or $8.5 billion and $9.9 billion, respectively.

We were asked to describe the three tests IRS has begun to reduce overclaims and how the funds appropriated for them were spent; assess how well IRS implemented the tests and describe planned refinements for the 2005 tests; and assess whether IRS's evaluation plans had sufficient documented detail to facilitate managerial review and stakeholder oversight and describe the status of the 2005 evaluation plans.

What GAO Recommends

GAO recommends that the Commissioner of Internal Revenue ensure the rationale for key decisions is documented; information on the quality and use of all types of taxpayer assistance is obtained; limitations are clearly stated when disseminating results; and development of detailed evaluation plans for the 2005 tests is completed. The Commissioner agreed with the recommendations.

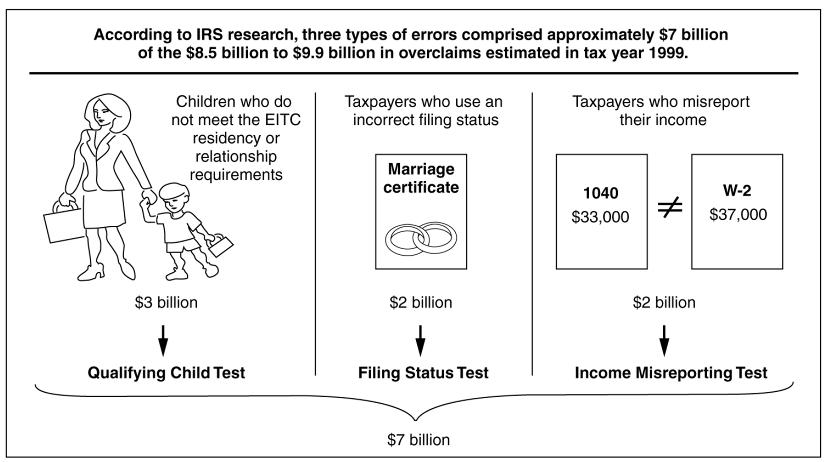
www.gao.gov/cgi-bin/getrpt?GAO-05-92.

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

What GAO Found

IRS implemented three tests in 2004 to address leading sources of EITC errors: a qualifying child test, where selected taxpayers were asked to document that their child lived with them for more than half the year in 2003; a filing status test, where selected taxpayers were asked to provide documentation to prove the accuracy of their 2003 filing status, and an income misreporting test, where a new screening process was used to select EITC returns that identify taxpayers likely to have the most significant changes in their assessments due to underreporting of income on their tax return.

Leading Sources of EITC Errors Contributing to Overclaims in Tax Year 1999



Source: GAO.

IRS's implementation of the tests proceeded smoothly and largely as planned. However, some information, such as a key change in the filing status test, was not well documented and the level and quality of some services provided to test participants were not measured. This lack of documentation hindered monitoring, oversight, and did not foster a common understanding of the tests. For the 2005 tests, IRS made key changes to the qualifying child test to encourage taxpayers to certify in advance of filing their return and to attempt to simulate what might happen with nationwide implementation. IRS also changed the sample selection criteria for the filing status test to better target noncompliant taxpayers.

IRS's plans for evaluating the 2004 tests generally lacked documentation and detail for many key issues, which undermined their value to managers and stakeholders. For example, IRS did not specify how it planned to analyze some qualifying child survey data. In essence, an evaluation plan is the management plan or roadmap for the evaluation endeavor and well-developed plans facilitate test management and oversight. Despite the importance of having evaluation plans prior to implementation, IRS had not completed its plans for the 2005 tests before two of the tests had begun.

Contents

United States Government Accountability Office
Washington, D.C. 20548

December 30, 2004

The Honorable Amo Houghton
Chairman, Subcommittee on Oversight
Committee on Ways and Means
House of Representatives

The Honorable Earl Pomeroy
Ranking Minority Member, Subcommittee on Oversight
Committee on Ways and Means
House of Representatives

Researchers generally consider the Earned Income Tax Credit (EITC), a federal program providing tax relief to low-income workers, to be a successful antipoverty program. The Council of Economic Advisers reported that because of the EITC, an estimated 4.3 million individuals—including 2.2 million children—were lifted out of poverty in 1997.[1] In recent years, IRS has paid about $30 billion annually to about 20 million EITC recipients. However, the EITC program has long experienced high rates of noncompliance. IRS's most recent EITC compliance study estimated that between $8.5 billion and $9.9 billion of the EITC claims filed for tax year 1999 should not have been paid. This amount represents an estimated rate of EITC overclaims—total erroneous claims less any amount that IRS recovered or expects to recover—of between 27 and 32 percent of EITC dollars claimed.

In February 2002, when the most recent compliance study was released, a task force of IRS and Treasury officials was convened to find ways of reducing EITC overclaims. The IRS/Treasury task force found that the three leading errors were responsible for about $7 billion of overclaims each year. These errors resulted from taxpayers (1) claiming children who were not a qualifying child, meaning they do not meet certain requirements, primarily that the child did not live with them for more than half the tax year ($3 billion in overclaims); (2) using an incorrect filing status of either single or head of household, when the correct status was married filing jointly or married filing separately ($2 billion in overclaims); and

[1]Council of Economic Advisers, *Good News for Low Income Families: Expansions in the Earned Income Tax Credit and the Minimum Wage* (Washington, D.C.: December 1998). The Council of Economic Advisers gives advice directly to the President and other senior members of the administration on fiscal policy.

(3) misreporting, primarily underreporting, their income ($2 billion in overclaims). In all three cases, these errors resulted in taxpayers receiving a larger credit than they should have received.

IRS received about $52 million in fiscal year 2004 for a new EITC five-point initiative to improve service, fairness, and compliance with the EITC program. Included in that amount were funds for three EITC tests to evaluate new methods for reducing the overclaim rate in the three leading problem areas identified by the IRS/Treasury task force—qualifying child, filing status, and income misreporting. After designing the tests and evaluation plans to guide the agency in its assessment of the tests, IRS implemented them beginning in fiscal year 2004. Although the tests and evaluation of them continued into fiscal year 2005, IRS issued a status report to Congress on the tests in August 2004[2] to comply with a legislative mandate.[3] IRS also decided to begin another round of EITC tests in fiscal year 2005 in the same three areas, but with some refinements. The qualifying child and income misreporting tests started in the first quarter of fiscal year 2005 and the filing status test will begin in the second quarter of 2005.

Because IRS's plans surrounding the tests have garnered much attention,[4] you asked us to review each of the fiscal year 2004 tests and (1) describe the tests and how funds appropriated for them were spent, (2) assess how well IRS implemented the tests and describe IRS's planned refinements for the fiscal year 2005 tests, and (3) assess whether IRS's plan for evaluating the tests contained sufficient documented detail to facilitate managerial review and stakeholder oversight and describe the status of IRS's evaluation plan for the fiscal year 2005 tests.

To describe the tests and determine how the funds appropriated for them were spent, we analyzed IRS documents and interviewed IRS officials. Our assessment of IRS's implementation of the tests and documentation of its evaluation plans for the tests was based on IRS's stated goals for them,

[2] Internal Revenue Service, *IRS Earned Income Tax Credit (EITC) Initiative: Status Report to Congress* (Washington, D.C.: August, 2004).

[3] Pub. L. No. 108-199, 118 Stat. 3 (2004).

[4] We reported on this topic in 2003, focusing on the test relating to qualifying child errors. See GAO, *Earned Income Credit: Qualifying Child Certification Test Appears Justified, but Evaluation Plan Is Incomplete*, GAO-03-794 (Washington, D.C.: Sept. 30, 2003).

which were that the tests reduce overclaims and, for the qualifying child test, also maintain EITC participation for eligible participants and minimize taxpayer and IRS administrative burden.

To assess how well IRS implemented the tests and describe IRS's planned refinements for the 2005 tests, we developed criteria to evaluate IRS's implementation (i.e., the execution and day-to-day management) of the tests. Those criteria included whether IRS sent correspondence to the taxpayers' correct address, where taxpayers could go to receive assistance about the tests and the quality of that assistance, IRS's hiring and training of staff, and other aspects of administering the tests. To assess the implementation based on these criteria, we analyzed IRS's status reports; reviewed policies and procedures; observed test operations in Kansas City, Missouri; Atlanta, Georgia; and Fresno, California and selected those locations based on a variety of reasons, including the location of key managers and work; interviewed IRS officials; and reviewed a judgmentally selected sample of cases for each test at those locations in order to review a variety of case types.

To assess whether IRS's plans for evaluating the tests contained sufficient documented detail to facilitate managerial review and stakeholder oversight, we used GAO guidance and the social science evaluation literature to identify key attributes of an evaluation. We shared these attributes with IRS officials and they generally agreed to their relevance. These attributes included the research design, outcome measures,[5] target and sample populations, data collection activities, analyses, and dissemination of results. We obtained all available documentation on IRS's evaluation plans for each of the tests and reviewed that documentation to determine whether we could understand from the documentation alone how IRS planned to address the key attributes. Where we could not, we interviewed IRS officials to further understand whether and how the officials planned to address those key attributes. Written documentation should be complete, facilitate tracing of events, and be readily available for examination to foster a common understanding of the program and

[5] Outcome measures are used to assess the tests' impact. Well-defined outcome measures clearly name and define a measure as well as the methodology used in its calculation.

facilitate oversight.[6] To describe the status of IRS's evaluation plan for the fiscal year 2005 tests, we primarily relied on interviews with IRS officials. Appendix I provides more detail on the scope and methodology, including a more detailed description of the implementation criteria used in conducting our work.

We reviewed documentation and interviewed IRS officials to ensure that the data that we received from IRS were reliable for the purposes used in this report and determined they were. As we conducted our work, IRS continued to collect and analyze significant amounts of data from each test. The tests continued into fiscal year 2005, primarily to finish cases where taxpayers had either not responded to IRS's request for substantiation or the substantiation provided was deemed insufficient by IRS. Thus, the analyses presented in this report are based on work that we performed while the tests were ongoing. Final conclusions about the impact of the tests on overclaim rates, burden, and participation should not be made until all data are collected and analyzed and IRS publishes its final report, which is due to Congress by June 30, 2005. We performed our work from October 2003 to December 2004 in accordance with generally accepted government auditing standards.

Results in Brief

IRS implemented three tests—qualifying child, filing status, and income misreporting—to evaluate the potential for reducing the EITC overclaim rate beginning in fiscal year 2004. For the qualifying child test, selected taxpayers were asked to document, when filing their 2003 tax return, that their qualifying child lived with them for more than half the tax year. The filing status test had selected taxpayers provide documentation to prove the filing status claimed on their 2003 tax return. For the income misreporting test, IRS employed a new screening process to select EITC tax returns from an existing program to identify taxpayers likely to have the most significant underreporting of income on their tax return and, therefore, the highest potential EITC overclaim amount. IRS reported spending about $17.5 million on the three EITC tests—about $3.2 million less than planned because expected workloads did not materialize. Because IRS spent less than planned on the tests themselves, it was able to

[6] See GAO, *Internal Control Management and Evaluation Tool*, GAO-01-1008G (Washington, D.C.: August 2001); GAO, *Standards for Internal Controls in the Federal Government*, GAO/AIMD-00-21.3.1 (Washington, D.C.: November 1999); and GAO, *Designing Evaluations*, GAO/PEMD-10.1.4 (Washington, D.C.: May 1, 1991).

fund some activities under the five-point initiative that otherwise would have gone unfunded.

Implementation of the tests generally proceeded smoothly. IRS addressed the major issues that arose during implementation and based planned refinements for the 2005 tests primarily on lessons learned from the 2004 tests. Some of the strengths of the tests' implementation include that they were conducted largely as planned and IRS hired and provided staff with training. An example of an issue that arose and IRS subsequently addressed was when some examiners mistakenly required taxpayers who were part of the filing status test to complete documentation required for the qualifying child test. Once this problem was identified, according to IRS officials, internal quality reviews helped ensure it was no longer occurring. Although implementation generally proceeded smoothly, some important information about the tests, including the basis for a key policy change, was not well documented and the level and quality of some services offered to taxpayers were not measured. As a result, it was difficult for management or staff to gain a common understanding of the program and rationale for changes and to fully monitor the implementation. Further, developing and documenting such information would enable others to review the methodology. IRS made refinements to all three tests for the subsequent round of fiscal year 2005 tests. Two of the most significant refinements were to the qualifying child test, including that (1) taxpayers would be encouraged to certify in advance of filing their return that their child met the EITC residency requirement and (2) a portion of the taxpayers would be drawn from a single community.[7] According to IRS officials, earlier certification could help them process cases more quickly and get eligible taxpayers their refund faster. Targeting one community would simulate what might happen if a certification requirement were imposed nationwide. Another key refinement was related to the filing status test, whereby IRS would change the sample selection criteria to better target noncompliant taxpayers.

IRS's plans for evaluating the three 2004 tests had critically important strengths such as identifying test goals and linking evaluation objectives and outcome measure to these goals. However, IRS's evaluation plans were generally not in sufficient detail to facilitate managerial review and

[7] Although IRS mailed letters to sampled taxpayers in this community, some uncertainty exists about implementation because of a lawsuit that was filed to block the test as we were finalizing our report.

stakeholder oversight and help ensure that the evaluations' results would be communicated in such a manner that others could understand and judge their strengths and limitations. In essence, an evaluation plan is the management plan or roadmap for the evaluation endeavor and would facilitate oversight. As such, the more completely a plan is developed, the more likely it will be useful to managers in ensuring that the evaluation is well-executed. IRS's evaluation plans lacked detail, for example, about key planned analyses that IRS intended to conduct. For example, IRS did not specify how it planned to analyze some qualifying child survey data. Despite the importance of having detailed plans prior to implementation, IRS has not completed its evaluation plans for the fiscal year 2005 tests even though two of those tests have begun.

We are recommending that the Commissioner of Internal Revenue ensure that the rationale for key policy decisions and other-significant events be documented; obtain some information on the quality and level of service for all types of taxpayer assistance provided as part of the EITC tests; when disseminating the test results ensure they clearly state test and evaluation limitations and complete the development of comprehensive and adequately detailed evaluation plans for the 2005 tests.

We requested comments on a draft of this report from the Commissioner of Internal Revenue. We received written comments, which are reprinted in appendix III. In his comments, the Commissioner said that IRS agreed with the recommendations. We further discuss the Commissioner's comments in the "Agency Comments and Our Evaluation" section of the report and in other sections where appropriate.

Background

The EITC, enacted in 1975,[8] was originally intended to offset the burden of Social Security taxes and provide a work incentive for low-income taxpayers. The credit has been modified several times since its introduction, and three laws[9] have been enacted in recent years aimed at resolving some concerns with EITC rules. Despite modifications, the original goal of the credit remains intact and the EITC continues to provide a substantial benefit to millions of American families.

[8] 26 U.S.C. Sec. 32.

[9] Pub. L. No. 107-16, 115 Stat. 38 (2001), Pub. L. No. 106-170, 113 Stat. 1860 (1999), and Pub. L. No. 105-34, 111 Stat. 788 (1997).

The EITC is a refundable tax credit, meaning that qualifying working taxpayers may receive a refund greater than the amount of income tax they paid for the year. EITC payments have a (1) phase-in range in which higher incomes yield higher EITC amounts, (2) plateau phase in which EITC amounts remain the same even as income rises, and (3) phase-out range in which higher incomes yield lower EITC amounts. The amount of credit a taxpayer receives is based on several other factors, such as the presence and number of qualifying children. In general, taxpayers with one or more qualifying children receive a higher credit than taxpayers without qualifying children. For tax year 2003, the amount of EITC that could be claimed with two qualifying children ranged from $0 to $4,204 per tax return filed, depending on income and filing status.

EITC requirements for tax year 2003 include rules for all taxpayers claiming the credit and additional rules that differ depending on whether or not a taxpayer has qualifying children (see table 1).

Table 1: EITC Requirements for Tax Year 2003

Rules for all taxpayers claiming the EITC	Additional rules for taxpayers with a qualifying child	Additional rules for taxpayers without a qualifying child
Must have a valid Social Security number	Income less than: If one child: $29,666 (or $30,666 if married filing jointly). If more than one child: $33,692 (or $34,692 if married filing jointly).	Income less than: $11,230 (or $12,230 if married filing jointly)
Cannot use married filing separately status	Child must meet age, relationship, and residency tests [a]	Must be at least 25 years old, but under 65
Must be a U.S. citizen or resident alien all year	Child can be claimed by one taxpayer only	Cannot be the dependent of another person
Cannot file form 2555 or 2555-EZ	Cannot be a qualifying child of another taxpayer	Cannot be a qualifying child of another taxpayer
Must have investment income of $2,600 or less		Must have lived in U.S. more than half the year
Must have earned income		

Source: IRS.

[a] In general, a qualifying child must be under 19 years old at the end of the tax year and have lived with the taxpayer for more than half the year. To meet the relationship test, the qualifying child had to be a son, daughter, adopted child, stepchild of the taxpayer, or a descendent of any such individual. Sisters, brothers, stepsisters, stepbrothers, and descendents of any such individuals also qualify if the taxpayer cares for the individual as they would their own child. In addition, a foster child can qualify for the relationship test if certain conditions are met.

IRS has periodically measured EITC compliance.[10] For tax year 1999, (the most current data available), IRS estimated the EITC overclaim rates at 27 to 32 percent of EITC dollars claimed, or $8.5 billion to $9.9 billion.[11] IRS has limited data on underclaims, which for tax year 1999 were estimated to be $710 million to $765 million. Because of the persistently high rates of noncompliance, we also have identified the EITC program as a high-risk area for IRS since 1995.[12]

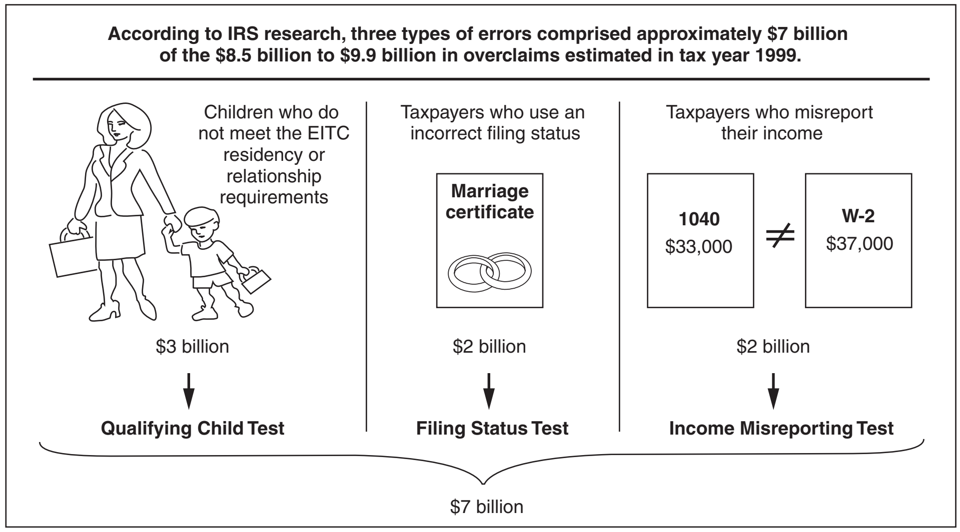
In February 2002, the compliance study was released and the Assistant Secretary of the Treasury, Tax Policy, and IRS Commissioner convened a joint IRS/Treasury task force to identify ways of reducing EITC overclaims, while maintaining participation among eligible claimants and minimizing taxpayer and IRS's administrative burden. The task force found that the leading causes of errors resulting in EITC overclaims were due to taxpayers (1) claiming children who were not a qualifying child, (2) using an incorrect filing status, and (3) misreporting their income. With this information, the task force designed what ultimately became initial versions of the three tests, as show in figure 1. As envisioned by the task force, even if fully implemented, IRS does not plan to apply the test requirements to the entire EITC population because IRS can use available data to verify the eligibility of certain taxpayers.

[10] IRS measured EITC overclaim rates in the past using differing methodologies. In 1997, IRS estimated the overclaim rate at 23.8 to 25.6 percent of EITC dollars claimed.

[11] Department of the Treasury, Internal Revenue Service, *Compliance Estimates for Earned Income Tax Credit Claimed on 1999 Returns* (Washington, D.C.: Feb. 28, 2002).

[12] Prior to 2001, EITC was part of a broader IRS tax filing fraud high-risk area that we defined. Beginning in 2001, the focus of that designation was narrowed to EITC. GAO, *High-Risk Series: An Update*, GAO-01-263 (Washington, D.C.: January 2001).

Figure 1: Leading Sources of EITC Errors Contributing to Overclaims in Tax Year 1999



Source: GAO.

Because a new analysis of EITC compliance using 2001 tax return information is not expected to be complete until spring 2005, IRS did not know whether compliance has significantly changed since 1999 when developing the EITC tests. However, IRS officials do not think EITC compliance has improved substantially since then. In October 2004, Congress passed a new law to make the definition of a qualifying child uniform in various IRS provisions, but those changes are not effective until tax years after December 31, 2004. In general, the revised definition appears to mainly affect other tax situations, such as claiming dependents, more than just affecting the EITC.[13] IRS is studying whether the change would affect any testing that may be done in 2006.

[13] Pub. L. No.108-311, 118 Stat. 1166 (2004).

Having a Complete Evaluation Plan Before Implementation Helps to Ensure Success

IRS completed its initial evaluation plans for the three EITC tests in December 2003. In September 2003, we recommended that IRS accelerate the development of its qualifying child evaluation plan to help ensure the success of the test. [14]

An evaluation plan ideally should be completed and disseminated for review and feedback before beginning the research activity (or in this case, test). As we reported, although an evaluation plan need not precisely identify all issues and how they will be evaluated before implementation, the more complete a plan is, the more likely the evaluation will be sufficient and support future decisions. IRS's Internal Revenue Manual[15] also recognizes the desirability of having an evaluation plan in place before a project is implemented; for example, it requires such plans before reorganizations.

IRS Implemented Three Tests on Leading Sources of EITC Noncompliance and Reported Spending Most of the Funding Received on the Tests

In an effort to implement the joint IRS/Treasury task force recommendations, IRS implemented three new tests—qualifying child certification, filing status, and income misreporting—in 2004. IRS reported spending about $17.5 million on the three EITC tests—about $3.2 million less than planned. Because IRS spent less than planned on the tests, it was able to fund some activities under the five-point initiative that otherwise would have gone unfunded.

Qualifying Child Certification Test Required Substantiation of Child Residency

The purpose of the qualifying child certification test was to evaluate the impact on the test goals of asking taxpayers to substantiate—when filing their tax return—that their qualifying child lived with them for more than half the tax year, as required by the EITC (see table 1). Under current rules, taxpayers are only required to substantiate that their child satisfied this residency requirement if they are being audited by IRS on this issue.

[14] GAO-03-794.

[15] See *Internal Revenue Manual* 1.1.4.

This test involved two random samples of 25,000 taxpayers who claimed one or more qualifying children for tax year 2002: a test sample, whose members were asked to substantiate their qualifying child's residency, and a control sample, whose members had similar characteristics to the test sample, but were not asked for any substantiation. Both samples were designed to include taxpayers (1) most likely to make errors and (2) whose qualifying child eligibility could not be verified from information available to IRS. IRS used prior research results to determine which taxpayers would be most likely to incorrectly claim a qualifying child.

The research showed that with those taxpayers most likely to make errors, the errors often correlated with the taxpayer's relationship to the child and the taxpayer's gender and filing status. Taxpayers most frequently making qualifying child errors included both fathers and males and females who were not the child's parents and who filed as single or head of household.

IRS also used available data to obtain evidence about taxpayers and whether their qualifying children met residency and relationship requirements. For example, a child's residency could be established with some certainty by using the Department of Health and Human Services' Federal Case Registry, and a child's relationship to the taxpayer could be established with some certainty using the Social Security Administration's KIDLINK.[16] When this available evidence supported the taxpayers' EITC claim that they had a qualifying child, those taxpayers were excluded from the qualifying child test.

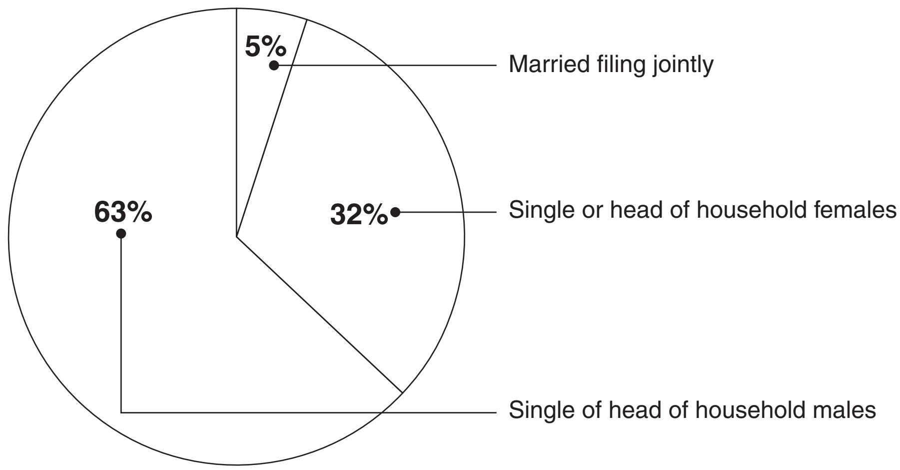
Prior research showed that taxpayers who comply with the residency requirement also comply in most cases with the relationship requirement. Thus, if a taxpayer's child met the residency requirement, there was a high probability that the relationship requirement would be met as well. Given this analysis and difficulties IRS encountered in identifying documents that taxpayers could readily obtain to prove their relationship to the child, any taxpayer whose EITC eligibility was not verified from available data became eligible to be selected for the qualifying child test in which they would be asked to substantiate that the child lived with the taxpayer for

[16] The Federal Case Registry compiles court and other records that indicate who is the custodian for a child. IRS assumes that children live with the custodian of record. KIDLINK ties parents' and children's Social Security numbers for children born after 1998 in U.S. hospitals.

more than 6 months during the year. Our September 2003 report contains a more detailed explanation on how the sample was designed.[17]

As figure 2 shows, males filing as single or head of household comprised the majority of the test sample. The control group had characteristics similar to the test group.

Figure 2: Characteristics of the Test Sample for the Qualifying Child Test



Source: GAO analysis of IRS data.

The qualifying child test had three components—a general test and two subtests. Under the general test, taxpayers received test documentation in English only[18] and could have provided substantiation in one or any combination of three ways—records, letters, or a Schedule A, also known as the general affidavit.[19] Records that a taxpayer could provide included school, medical, landlord, or child-care provider documentation. Letters were statements from certain individuals, such as a member of the clergy or a community based organization official, on official letterhead.

[17] GAO-03-794.

[18] Taxpayers were told on the forms that they could request that documents be provided in Spanish.

[19] Schedule A, referred to as the general affidavit, can be signed by parties including an attorney, child care provider, clergy, court official, employer, health care provider, Indian tribal official, landlord, school official, or social service agency or other government official.

Affidavits were legal documents in which an individual attests that the taxpayer's qualifying child resided with the taxpayer for a certain period of time. To be accepted, the document(s) had to contain various data, such as the names of the qualifying children and the dates the child lived with the taxpayer.

In response to concerns that taxpayers may have difficulty obtaining certification through the official sources cited on the Schedule A, such as through an attorney or landlord, and that English-only documents might weaken participation among taxpayers with limited English proficiency, IRS also implemented two subtests. The Schedule B, also known as the friends and neighbors affidavit subtest, for 1,000 of the 25,000 taxpayers, broadens the definition of the individuals allowed to certify the child's residency to include those who have personal knowledge of a taxpayer's circumstances, such as certain family members.[20] The purpose of this subtest was to determine whether such individuals could facilitate an increase in residency certification for eligible taxpayers. The taxpayers in the Spanish subtest, 1,000 of the remaining 24,000, received documents in both English and Spanish. The purpose of this subtest was to determine whether Spanish language documents would increase the number of taxpayers attempting to certify their child's residency. Table 2 describes the test and subtests.

Table 2: Description of the Qualifying Child Certification Test and Subtests

Test or Subtest	Description of Test
General Test (23,000 taxpayers)	• Documentation sent in English only • Only General Affidavit (Schedule A) provided
Schedule B Subtest (1,000 taxpayers)	• Documentation sent in English only • Friends and Neighbors Affidavit (Schedule B) provided
Spanish Subtest (1,000 taxpayers)	• Documentation sent in English and Spanish • Only General Affidavit (Schedule A) provided

Source: GAO analysis of IRS data.

[20] The Schedule B affidavit, referred to as the friends and neighbors affidavit, differs from the Schedule A affidavit in that it broadens the list of individuals who can sign the affidavit to anyone who has records or personal knowledge, other than the taxpayer's spouse, dependent, qualifying child, or parent of the qualifying child.

IRS sent the selected taxpayers[21] five documents in December 2003 informing them about the test, including:

1. Notice 84-A, a letter informing the taxpayer about the new certification requirements;

2. Form 8836, *Qualifying Children Residency Statement*, to be completed by the taxpayer and returned to IRS;

3. Schedule A or B (an affidavit) that could be used for certification;

4. Publication 3211M, *Earned Income Tax Credit Questions and Answers*; and

5. Publication 4134, *Free/Nominal Cost Assistance Available for Low Income Taxpayers*.

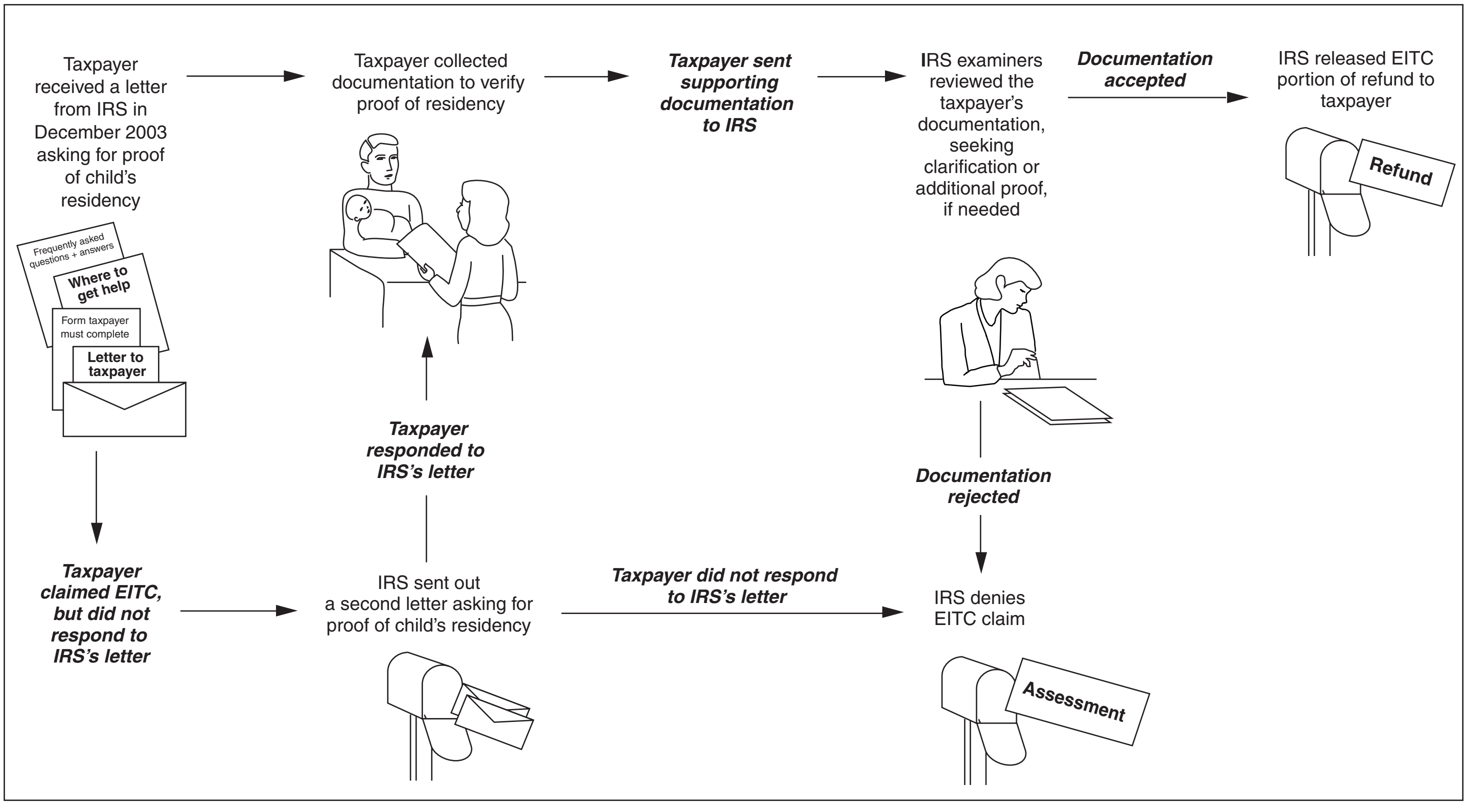
Under the test and subtests, once taxpayers received the documents from IRS, they were supposed to obtain supporting documents to prove their qualifying child's residency and send that documentation back to IRS in the same envelope as their 2003 tax return. IRS would withhold, or "freeze," the EITC portion of the taxpayers' refund until acceptable documentation proving a child's residency was received. Once IRS received the documentation, IRS examiners in Kansas City, Missouri, would review it and send a letter to the taxpayer accepting the claim, asking for additional information, or rejecting the claim. If the taxpayers provided acceptable documents, IRS would release the taxpayer's EITC portion of their refund.

If acceptable documentation was not provided or if no response was provided following a second notification letter, the taxpayer's EITC claim would be denied and the taxpayer would be informed of his or her right to appeal to the U.S. Tax Court.[22] This process is depicted in figure 3.

[21] Some taxpayers were removed from the 25,000 sample and not replaced because IRS learned that they had mitigating situations, such as serving in combat or living in a disaster zone. The true sample size (i.e., the number of taxpayers IRS required documentation from) was actually 24,711.

[22] IRS allowed for internal appeal hearings only if the taxpayer maintained contact with IRS and requested such a hearing. In the original contact letter, IRS told taxpayers they could request an internal appeal hearing.

Figure 3: The Qualifying Child Certification Test Process



Source: GAO.

Filing Status Test Required Substantiation That the Correct Filing Status Was Used

Another cause of EITC errors is when taxpayers claim an incorrect filing status. EITC filing status errors occur when married taxpayers incorrectly use single or head of household. Married taxpayers who incorrectly file individually as single or head of household could qualify for a larger EITC than they would otherwise be entitled to if they claimed the correct filing status. This is because, pursuant to statute,[23] IRS considers the combined income of married taxpayers who file jointly for purposes of determining the amount of EITC for which the taxpayer(s) qualifies.[24] Using combined

[23] I.R.C. Section 32(b)(2)(B).

[24] Taxpayers who file as married filing separate are not eligible for the EITC.

income may result in taxpayers exceeding the EITC income ceiling, therefore receiving no credit at all, or qualifying for a lesser credit amount. For example, in tax year 2003, married taxpayers filing jointly with $17,500 of income each, or a combined earned income of $35,000, and four qualifying children would not be eligible for the EITC. However, if each taxpayer incorrectly filed as head of household, claimed two qualifying children, and their $17,500 income, they would each receive a credit of $3,405 or a combined total of $6,810. IRS's databases offer limited ability to independently or systematically identify taxpayers who may be claiming an incorrect filing status.

The primary purpose of the filing status test was to evaluate the impact on overclaims of requiring taxpayers whose filing status has changed from married to single or head of household any time between 1999 through 2002 to substantiate the filing status they claimed on their 2003 tax return. To select the population for the filing status test, IRS started with a computer file of approximately 1.6 million taxpayer returns, or a 10 percent sample of all taxpayers who claimed the EITC with one or more qualifying children on their 2002 return. IRS eliminated the qualifying child and income misreporting test populations and, applied other exclusions, such as taxpayers subject to an audit examination, or taxpayers with more than one potential EITC-related issue. From that population, IRS selected taxpayers whose returns showed a filing status of married at least once in the previous 3 years. This resulted in a sample of 69,000 taxpayers, which IRS sorted by gender, zip code and filing status. Using a random sampling method, IRS selected 36,000 of these taxpayers for this sample who filed as single or head of household on their 2003 tax return. Females filing as single or head of household comprised 96.9 percent of the test sample.

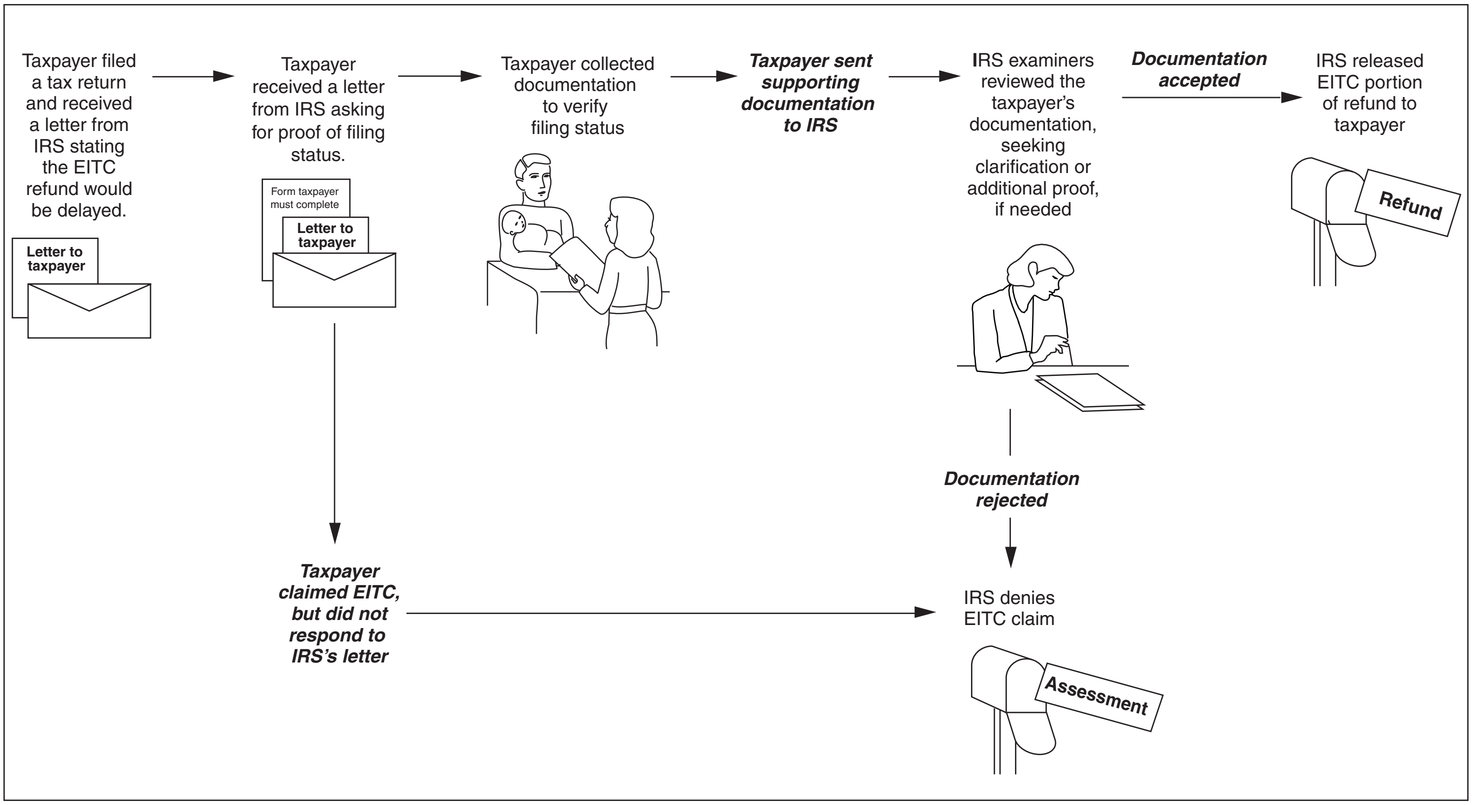
The taxpayers in the 36,000 sample who filed a 2003 tax return claiming the EITC received a letter from IRS about 2 weeks after filing their return informing them that the EITC portion of their refund would be delayed until IRS reviewed their return. Within 30 days, IRS sent a second letter asking taxpayers to verify their filing status, using the enclosed Form 886-FS, Filing Status Information Request and send it back to the IRS. This form requires taxpayers to provide documentation as to why they did not file as married for tax year 2003. Taxpayers were asked to provide IRS with documentation that they were divorced or legally separated as of December 31, 2003, they did not live with their spouse for the last 6 months of the year, the spouse was deceased, or some other reason existed to warrant a change of filing status.[25] IRS examiners reviewed the form and accompanying documentation and sought clarification or additional proof, if needed. If IRS examiners accepted the documentation, they released the EITC portion of the taxpayer's refund and closed the case.

If a taxpayer did not respond or IRS found the taxpayer's documentation unacceptable, then IRS sent the taxpayer a notice stating that IRS (1) changed the taxpayer's filing status from single or head of household to the married filing separate status, (2) disallowed the EITC, and (3) changed the taxpayer's standard deduction to the appropriate amount. In addition, IRS forwarded the taxpayers a letter informing them of their right to appeal the changes to U.S. Tax Court.[26] This process is depicted in figure 4.

[25] This is the same documentation IRS would request in an audit if filing status were an issue.

[26] Similar to the qualifying child test, IRS allowed for internal appeal hearings only if the taxpayer maintained contact with IRS and requested such a hearing.

Figure 4: The Filing Status Test Process



Source: GAO.

The filing status test also included a subtest to gather additional information on EITC claimants who used the head of household filing status. The IRS/Treasury task force found that taxpayers using the head of household filing status were more likely to misstate their filing status than taxpayers using a different one. IRS selected 500 taxpayers who filed as head of household on their 2003 tax return. The sample of 500 taxpayers showed 99 percent females and 1 percent males with the head of household filing status.

Unlike the test for the 36,000 sample, IRS did not ask taxpayers in the subtest of 500 who filed a 2003 return to provide supplemental documentation to support their filing status until after they had received their EITC refunds. And, unlike taxpayers in the 36,000 sample, where IRS had some information they had filed as married at least once from 1999 - 2002, IRS did not have such information on the taxpayers in the 500 sample.

In fact, IRS could not determine whether these taxpayers were ever married. As a result, IRS asked these taxpayers to confirm their eligibility for the head of household filing status, which they claimed on their 2003 tax return, by either (1) calling IRS on a special toll-free number and stating that they used the correct filing status or (2) completing a stub that was attached to the letter they received, checking yes or no, and mailing or faxing it to IRS. IRS did not ask these taxpayers to provide substantiation to support the filing status they claimed. This was, in part, because IRS had not identified any documentation that would be available to support a taxpayer's claim that he or she had never been married. If taxpayers indicated they were not eligible to use the head of household filing status, they could correct their filing status by sending in an amended tax return either by mail or fax.[27] IRS asked taxpayers to provide the information within 45 days from the date on the letter. All taxpayers who did not respond would be subject to an examination before their 2004 EITC refund would be released.

In another aspect of the filing status test, IRS planned to determine whether a third-party service that attempted to locate the address of taxpayers could be as reliable as the filing status test in identifying taxpayers who had used an incorrect filing status. The locator service used information from credit bureaus to determine whether taxpayers were living together and possibly married. The information from the locator service had no impact on taxpayers for this year's filing status test.

Income Misreporting Test Used New Screening Process to Find Cases Likely to Yield the Highest Assessments

Although some taxpayers could receive a larger EITC by over-reporting their income, misreporting of income for EITC is generally an understatement, according to IRS, resulting in the taxpayer receiving a higher credit amount than entitled. The purpose of the income misreporting test was to evaluate the impact on the test goal of a new screening process to select EITC tax returns [28] that identify taxpayers likely to have the most significant changes in their tax assessments due to underreporting of income on their tax return.

[27] IRS did not assess any penalties, interest, or additional tax due if the taxpayer sent an amended return. This decision was for the 2004 test only.

[28] Unlike the other two tests, the income misreporting test was based on 2002 tax returns. This was because IRS does not match information returns from third parties until several months after April 15th.

Income misreporting is a component of an existing program known as Automated Underreporter (AUR).[29] Under that program, IRS attempts to match income information as reported by the taxpayer on the tax return to information reported by third-party sources, such as a taxpayer's employer or bank. In instances where this matching process identifies discrepancies, IRS may assess additional taxes on the taxpayer. The annual AUR matching program identifies far more cases than IRS has staff to work. In determining which cases to work, IRS selects not only cases that it believes will generate the highest probable assessment, but also cases involving taxpayers who underreport different types of income (e.g., wages, interest). In the past, some of those cases—roughly 300,000 per year—involved the EITC. However, EITC was not one of the different types of categories from which IRS historically had chosen cases.

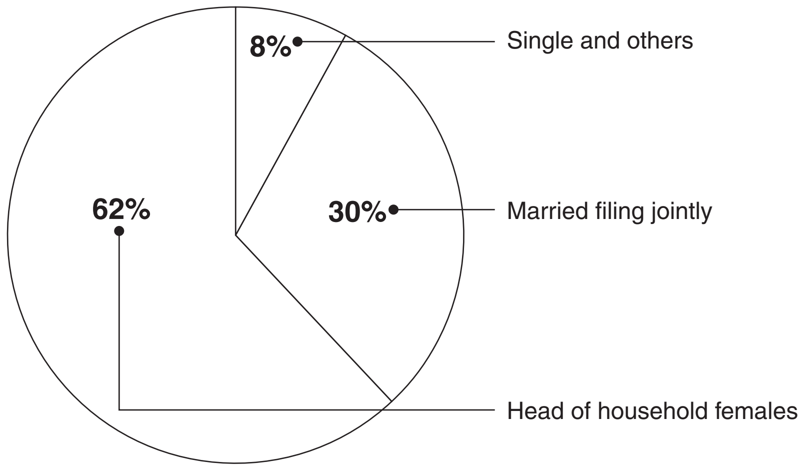
For the income misreporting test, IRS attempted to select---from all the EITC cases for which AUR found an income mismatch on 2002 tax returns—300,000 EITC cases expected to provide the highest EITC assessments. IRS employed a computer selection tool that used variables such as the taxpayer's filing history, filing status, and number of children to rank the cases in terms of the highest probable EITC assessments.

Additionally, IRS designed the test to determine whether certain characteristics of the selected cases made them more likely to yield higher assessments. Thus, IRS placed each of the selected cases in one of four groups: (1) "repeater egregious," cases in the same income category for the third year in a row and were assessed an additional tax for the previous 2 years; (2) "repeater worked," cases worked at least once during the last 3 years; (3) "repeater not worked," cases in the income misreporting inventory at least once in the last 3 years, but not worked; and (4) "other criteria," cases randomly selected from the other three categories and other criteria, such as first-time underreporters.

As figure 5 shows, the majority, 62 percent, of the taxpayers selected for the income misreporting test filed their return using the head of household filing status, while 30 percent claimed married filing jointly and 8 percent claimed a single filing status.

[29]During fiscal year 2003, the AUR program assessed taxpayers about $2.9 billion.

Figure 5: Characteristics of the Taxpayers in the Income Misreporting Test by Filing Status[a]



Source: GAO analysis of IRS data.

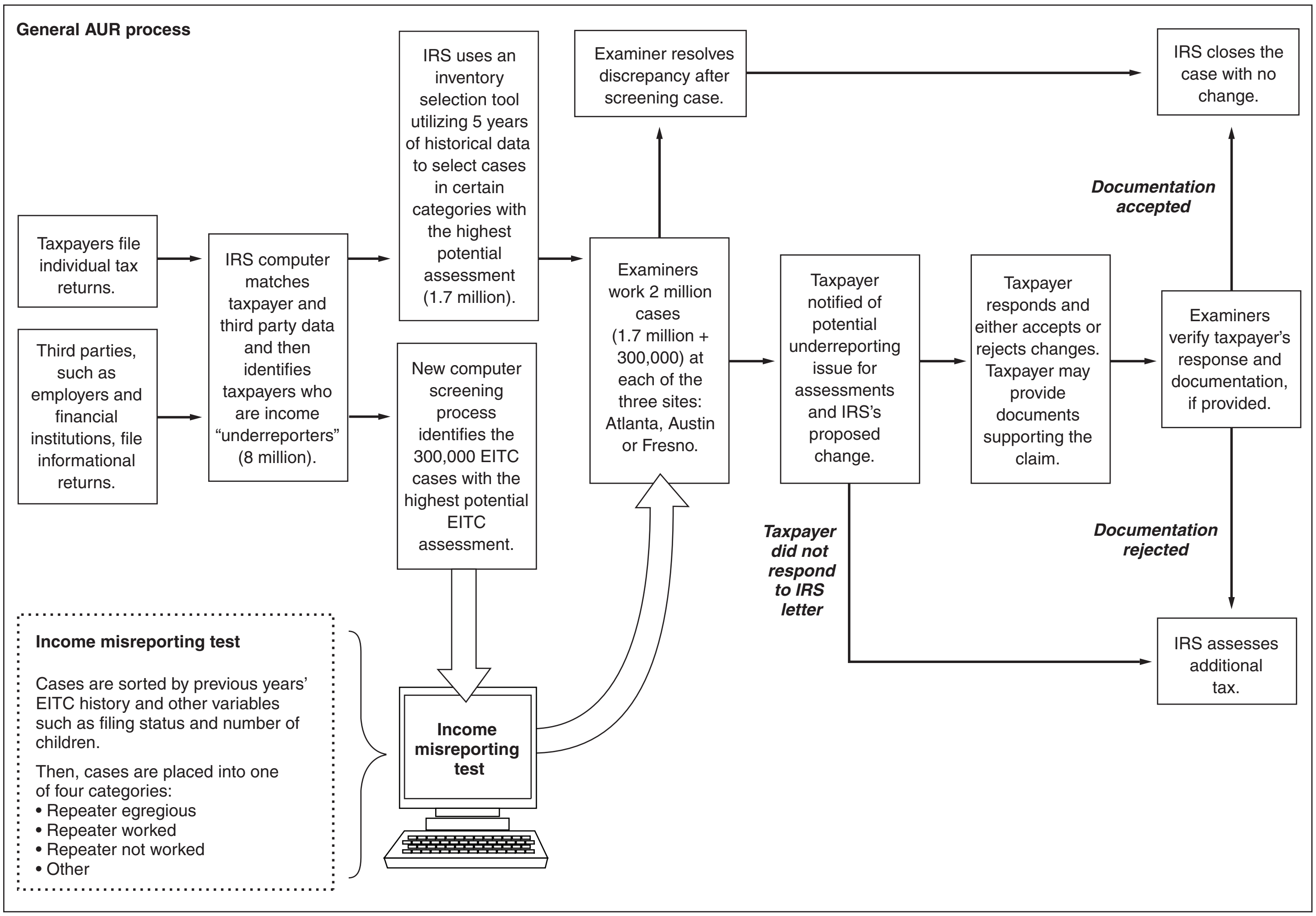
[a]IRS was unable to provide characteristics by gender.

IRS added the income misreporting test cases back into the general AUR inventory, and examiners in Atlanta, Georgia; Austin, Texas; and Fresno, California worked the test cases using the same processes as for all other AUR cases. Examiners manually screened all cases for simple math errors or errors that could not be picked up by a computer (e.g., placing an amount on the wrong line). If such an error was found and resolved, the tax return was accepted, and the case was closed. If the examiner could not resolve the discrepancy, the examiner sent a notice[30] to the taxpayer explaining that IRS found a discrepancy on his or her return. The taxpayer was given 30 days to respond to the notice. If no response was received, IRS sent another notice informing the taxpayer that the IRS had determined there was a deficiency in the return and the taxpayer must pay an assessment based on the deficiency or file a petition with the U.S. Tax Court within 90 days. If IRS received a response that took issue with IRS's assessment, the examiner would then determine whether the response was sufficient to support the taxpayer's original tax return.[31] If the response was sufficient, the examiner would close the case with no additional tax assessed. If the response was not sufficient or a response was not

[30] In most of these cases, a CP-2000 notice, "We Are Proposing Changes to Your Tax Return", is sent to the taxpayer; however, IRS sends a different notice in rare instances, such as when there was an income discrepancy of at least $100,000.

received, the IRS examiner would assess the taxpayer the additional tax. This process is depicted in figure 6.

Figure 6: The Income Misreporting Test



Source: GAO and Art Explosion (image).

[31] IRS officials said the Internal Revenue Manual guides this determination and that examiners receive months of training in how to make this determination.

IRS Reported Spending Less on Tests Than Anticipated

IRS reported spending about $17.5 million on the three EITC tests—about $3.2 million less than planned.[32] This funding was part of the Consolidated Appropriations Act of 2004,[33] which provided IRS with $52 million in fiscal year 2004 for a five-point initiative to improve service, fairness, and compliance with the EITC program. IRS announced the new initiative in June 2003. The initiative addresses:

- reducing the backlog of pending EITC examinations to ensure that eligible taxpayers whose returns are being examined receive their refunds quickly;
- minimizing burden and enhancing the quality of communications with taxpayers by improving the existing audit process;
- encouraging eligible taxpayers to claim the EITC by increasing outreach efforts and making the requirements for claiming the credit easier to understand;
- ensuring fairness by refocusing compliance efforts on taxpayers who claimed the credit, but were ineligible because their income was too high (or filing status was incorrect); and
- piloting a certification effort to substantiate qualifying child residency eligibility for claimants whose returns are associated with a high risk of error.

Of the $52 million budgeted, IRS reported spending or obligating $51.8 million in fiscal year 2004. Of that, IRS officials said they spent about $17.5 million on the tests---$7.4 million was spent on the income misreporting test, $5.6 million on the filing status test, and $4.5 million on the qualifying child test. IRS officials noted that, in some cases, the amounts they reported spending differed from what they budgeted. For example, IRS originally budgeted $7.2 million on the filing status test, but reported spending $5.6 million on direct costs for that test.

[32] We could not verify these figures because IRS does not have a cost accounting system. See GAO, *Financial Audit: IRS's Fiscal Years 2004 and 2003 Financial Statements*, GAO-05-103 (Washington, D.C.: Nov. 10, 2004).

[33] Pub. L. No. 108-199, 118 Stat. 3 (2004). The Act did not earmark a specific amount for the tests.

According to IRS officials, they spent about $3.2 million less than anticipated on the tests primarily because some planned work did not materialize. For example, for the filing status test, IRS originally planned to work more cases but about 10,000 taxpayers who were originally selected for the filing status test were not included for various reasons, such as they did not claim the EITC. IRS officials said that, as a result, they redirected funding to improvement projects within the five-point initiative that would otherwise have gone unfunded.

Tests Implemented Smoothly, and Refinements for the Fiscal Year 2005 Tests Made

IRS's implementation of the tests generally proceeded smoothly because of IRS actions including use of a detailed project plan and management involvement. IRS addressed most of the major issues that arose during implementation and released a status report to Congress in August 2004. IRS's plans for most refinements for the 2005 tests are based on the lessons that it learned from the 2004 tests.

Tests Were Executed Largely as Planned, Thus Meeting the Original Intent

The implementation plans for all three tests generally followed the recommendations of the IRS/Treasury task force, and IRS's only significant departure from those recommendations was based on an informed decision. The task force recommended that taxpayers claiming the EITC (1) provide IRS with documentation to prove a qualifying child's residency prior to payment of the credit (the qualifying child test), (2) submit additional data to establish that they are claiming the correct filing status (the filing status test), and (3) use a new screening process to select tax returns from an existing program to identify taxpayers likely to have the most significant underreporting of income on their tax return and, therefore, the highest potential EITC overclaim amount (income misreporting test). In all three tests, IRS gathered information needed to determine whether the task force recommendations have potential for reducing the EITC overclaim rate without undue adverse effects. It was important that IRS followed the task force recommendations; otherwise, the validity of those recommendations would remain unknown.

IRS made an informed choice in not implementing one recommendation. The task force had also recommended that taxpayers certify the child's relationship to the taxpayer. However, IRS determined that this was a lesser compliance problem than residency and that it could be difficult for taxpayers to provide some of the documentation that IRS planned to request for certification of the relationship. In addition, since both residency and relationship requirements had to be met to claim the EITC, if taxpayers fail residency certification, which is more likely according to the compliance study, there would be no need to test for relationship.

To implement each test, IRS prepared a detailed project plan with time frames for numerous action items such as developing notices, creating organization charts, hiring staff, developing training materials, working on systems needs, and determining samples. We found that IRS officials used these plans extensively. For example, initially, IRS managers checked the plan daily to determine if the schedule was being followed and less often as the tests progressed. For a task to be marked as completed, certain information had to be provided to the person in-charge of monitoring the plan, including validation from a senior manager that the task had been completed. According to IRS officials, the extensive use of the project plan helped them execute and effectively monitor the implementation of the tests.

Through Hiring, Training, and Management Actions, IRS Facilitated a Smooth Implementation

Implementation went smoothly, in part because IRS hired sufficient numbers of staff and provided adequate training to them. IRS hired about 410 staff, primarily examiners who processed cases and answered telephones, to implement the three EITC tests in total. About 260 of the staff were for the qualifying child and filing status test, while about 150 were for the income misreporting test. The majority of the qualifying child and filing status test staff were new to IRS, were hired on a temporary appointment, and worked in Kansas City, Missouri. The income misreporting staff worked in Atlanta, Georgia; Austin, Texas; and Fresno, California. According to IRS officials, these staffing levels were appropriate to manage the workload, thus contributing to the overall smooth implementation of the tests.

IRS provided specific training for the qualifying child and filing status tests. Among other things, the training included a history leading up to the tests, a description of the test processes, the roles and responsibilities of staff, several examples of how to determine whether taxpayer substantiation was acceptable, and information on how to use the Earned Income Credit Proof of Concept (EICPC) database, the database IRS used to manage the qualifying child and filing status tests. Examiners we met with in Kansas City told us the training was sufficient. However, there was a gap between the time examiners first received training and when they actually started doing the work. According to IRS officials, this gap caused the staff to lose some knowledge before they were able to apply it. To remedy this problem, IRS provided the staff refresher training and a staff-developed job aid. Examiners we interviewed said, that as a result, they felt confident in making decisions to accept or reject taxpayer substantiation.

IRS did not provide specific training for the income misreporting test, and instead relied upon the AUR program training because the process for working cases remained the same—only how IRS selected the cases changed. In our visits to Atlanta and Fresno, we found consistency in the training that staff received for the income misreporting test, including how the procedures were used when screening and working cases.

Management took steps to foster staff members' understanding of the importance of the tests. Once the current EITC program director was installed in late 2003, management oversight became more apparent for the tests. Senior IRS management responsible for EITC were involved in managing many details of implementation of the tests. To help garner staff support, when the tests first began, IRS managers held meetings with the examiners and took action based on their expressed concerns, such as making key revisions to the EICPC system. In addition, front-line managers with whom we spoke in Kansas City said the EITC director's involvement helped marshal staff support at that location. Managers said this was critical for smooth implementation of the tests, since they were the ones dealing directly with the taxpayers. The examiners we interviewed also said team meetings with managers helped them understand and effectively convey information about the tests to taxpayers.

IRS Monitored Undeliverable Mail and Attempted to Resend It to Corrected Addresses to Help Ensure Taxpayers' Response to Tests

IRS tracked undeliverable mail, mail that was sent to taxpayers and returned to IRS by the U.S. Postal Service, which was critical to the success of the tests. If taxpayers did not receive IRS's correspondence—letters, forms, and notices—they would not have known they needed to respond. And, had there been large volumes of undeliverable mail, the feasibility of the tests could have been undermined. Ensuring that those selected for the tests received the correspondence could have been particularly difficult because research has shown that some EITC claimants are highly mobile.

Although IRS used the most current address for test populations—in most cases the address taxpayers provided on their 2002 tax returns—IRS officials anticipated some mail being returned as undeliverable because the taxpayer no longer lived at that address. IRS first learned that it had a taxpayer's incorrect address when it received the undeliverable mail from the U.S. Postal Service. As it typically does for undeliverable mail, IRS employed a locater service to attempt to find the taxpayer's new address by using other kinds of information, such as addresses associated with any credit cards. When the locator service found a new address, IRS resent correspondence to the affected taxpayer.

IRS Provided Several Means to Help Answer Taxpayers' Questions and Found Strong Performance Where Data Were Available

IRS provided several means for taxpayers selected for any of the three tests to contact the agency for assistance. For example, in the initial contact letter for the qualifying child test, taxpayers were informed that they could get help from a toll-free telephone number where examiners could answer their questions, any local IRS office—commonly known as walk-in sites, and any of the approximately 200 low-income taxpayer clinics (LITCs) in the U.S. In addition, the National Taxpayer Advocate[34] was prepared to assist taxpayers selected for all three tests as needed.

Determining whether taxpayers received the correct information is an important aspect of implementation. IRS's performance in terms of the percentage of callers getting through to the agency and the quality of the answers given was strong and comparable to similar IRS operations. IRS received about 100,000 telephone calls from taxpayers about the qualifying child test and about 75,000 calls about the filing status test, as of September 30, 2004.[35] Common questions about the qualifying child certification test included "What documentation is acceptable?" and "When will my refund be released?" According to IRS's telephone data, about 86 percent of callers got through and received service. Based on historical data, IRS officials considered this level acceptable. Based on our annual reviews of IRS's telephone performance during the filing season,[36] we have reported comparable levels of service. For example, in 2003, 85 percent of taxpayer's calling IRS's main toll-free telephone lines got through and received service. IRS's internal quality reviews showed that, as of September 30,

[34]The National Taxpayer Advocate is available to help taxpayers when a hardship arises or when their problem has not been resolved in a reasonable time period.

[35] IRS received calls about the income misreporting test on the general AUR telephone line. IRS did not think it was necessary to track the test calls separately from the general AUR because taxpayers were treated the same under both the test and the program.

[36] GAO, *Tax Administration: IRS's 2003 Filing Season Performance Showed Improvements*, GAO-04-84 (Washington, D.C.: Oct. 31, 2003) and GAO, *IRS's 2002 Tax Filing Season: Returns and Refunds Processed Smoothly; Quality of Assistance Improved*, GAO-03-314 (Washington, D.C.: Dec. 20, 2002).

2004, test examiners provided accurate responses to taxpayers seeking assistance for the EITC tests via the telephone about 96 percent of the time, which was somewhat higher than the quality of IRS's responses on its toll-free telephone lines.

Because outside stakeholders expressed much concern about the tests, the National Taxpayer Advocate decided to assist any taxpayer selected for the tests, even if the assistance did not meet its established criteria.[37] The Advocate expected to assist about 2,600 taxpayers based on a needs assessment, which was rooted in historical data. However, as of September 30, 2004, the Taxpayer Advocate assisted a total of 837 EITC taxpayers participating in these tests. Most of the assistance provided included helping taxpayers receive an expedited refund due to a financial hardship. Internal quality reviews showed that the Advocate met its quality standards 100 percent of the time for the test cases selected as part of those reviews.

Steps Taken to Ensure Procedures Used by Examiners Led to Consistent Decisions

For each test, IRS took several steps that were designed to ensure consistency among the examiners making decisions about whether to accept taxpayers' substantiation. The qualifying child and filing status cases were worked in one location—Kansas City, Missouri—to make it easier to ensure consistency among examiners. The income misreporting test cases were worked in three locations—Atlanta, Georgia; Fresno, California; and Austin, Texas. IRS officials said they did not believe there would be a consistency problem in having the income misreporting test conducted across these locations because the test was not a significant departure from the general AUR work that had been done in these locations for the past several years. Other examples of steps IRS took to ensure consistent decision-making by examiners included holding multiple team meetings with staff, sending out notices to staff when errors were noted, having certain groups work only certain kinds of cases, preparing a job aid for examiners, and conducting extensive quality reviews. According to IRS managers and examiners we spoke with, these steps helped examiners make consistent decisions in the cases they were reviewing.

[37] The Taxpayer Advocate uses seven criteria to categorize and determine the severity and level of attention that taxpayers require to resolve their tax issue. For example, some issues, such as eviction or bankruptcy, are categorized as imminent hardships for taxpayers and receive the most prompt attention.

Another step IRS took to ensure consistency was to have managers in Kansas City review all those cases where taxpayers provided substantiation for the qualifying child test prior to filing their tax return—a total of about 800 reviews. This review helped IRS identify and correct problems that arose early in the tests. IRS officials stated that the review helped provide for a smooth implementation because it identified problems, which IRS corrected, and enabled IRS to issue supplemental guidance to ensure repeat errors did not occur. For example, for the qualifying child test, taxpayer substantiation did not always clearly indicate the exact dates of a child's residency with the taxpayer—for example, some may have shown "July through December 2003." Some examiners interpreted that to mean July 1 through December 31, giving taxpayers the time needed to certify for the qualifying child's residency. Other examiners interpreted this same information to mean July 1 through December 1, therefore not giving taxpayers the time needed to qualify their child. This review identified the inconsistent interpretation of dates, and IRS developed a policy decision and issued guidance on how to interpret the dates when the dates provided were vague.

IRS Internal Reviews Showed Few Significant Implementation Problems

The several on-going internal quality reviews during the tests generally found few significant implementation problems. IRS managers conducted reviews at the test sites, which examined accuracy, timeliness, and staff professionalism. For the qualifying child and filing status tests, these reviews showed generally high performance—case file documentation was correct 87 percent and 93 percent of the time, respectively, as of September 30, 2004. IRS did not capture this data for the income misreporting test; however, IRS data show that, as of September 30, 2004, 95 percent of all AUR cases, which included the income misreporting cases, contained correct documentation. The EITC Program Office also conducted a review that assessed whether policies and procedures for the qualifying child and filing status cases were being timely, accurately, and consistently followed. According to IRS, the review showed good results. For example, for the filing status test, the time between an examiner's decision to accept taxpayers' documentation and the issuance of the taxpayers' refund averaged fewer than 30 days.

IRS Addressed All But One Significant Problem That Arose During Implementation

Although several problems surfaced during implementation, particularly in the qualifying child test, IRS officials said that because they were able to take quick actions to address the problems, the problems did not adversely affect the tests or taxpayers selected for them to any great extent. It was not surprising that most of the problems involved the qualifying child test because it was a greater departure from past practice than were the other two tests. For example, although IRS had previously asked taxpayers to provide documents substantiating their qualifying child upon audit, IRS has not previously allowed taxpayers to provide affidavits to prove their claim; therefore, examiners had never reviewed such documents. In contrast, IRS considers the filing status and income misreporting tests expansions of existing IRS programs.

Examples of problems and IRS's actions to address them include:

- Early in the implementation of the test, some examiners advised taxpayers who had called about letters received from IRS to complete documentation for the qualifying child test even though they were selected for the filing status test. Examiners were instructed via an e-mail alert to use the EICPC database to determine the test for which the taxpayer was selected.
- Some qualifying child and filing status case files (paper and electronic) had documentation deficiencies, such as not getting a taxpayer's phone number for the case file or not obtaining complete/required information for cases where the taxpayer agreed with IRS's proposed changes. Through an e-mail alert, IRS officials reminded examiners of the procedures they must follow to properly document files.
- Some files were missing for the qualifying child and filing status tests. IRS established a new procedure that when a file could not be located within 2 weeks after the taxpayer had submitted correspondence, a new file would be created and marked "Possible Duplicate Folder."

In all three cases, IRS officials stated that the on-going quality reviews helped ensure that examiners followed the correct procedures.

Although IRS addressed problems that arose during the implementation of the tests, one significant problem still lingers. Some important information about all three tests, including a key policy decision regarding the filing status test, were either not well documented or not documented at all. Internal control standards state that significant decisions and events should be documented and readily available for examination.[38] When documentation is lacking, it is difficult for management or staff to gain an understanding of the program, refine work processes, or fully monitor the implementation. Further, developing and documenting such information would help ensure that test results are accurately determined and would enable others to review the methodology.

IRS developed various management documents to organize, direct, and monitor the test operations. However, while some important decisions about the tests were made after these documents were developed and after test implementation began, IRS did not explain the decisions by making additions to the documents. For example, IRS's initial plan required that the filing status subtest involving taxpayers who had never filed as married, but had filed as head of household on their 2003 return, include 5,000 taxpayers. Several months later, IRS reduced the sample to 500, but did not document the rationale for this decision until much later and at our request. Also, certain other key information, such as when and how information from a third- party locator service for the filing status test would be used, was not fully developed or sufficiently detailed. The Treasury Inspector General for Tax Administration (TIGTA) found similar deficiencies in IRS's documentation about the tests occurring during implementation.[39] As a result, this lack of documentation hindered not only test monitoring and oversight, but also did not foster a common understanding of the tests.

According to IRS officials, time or other priorities caused some significant decisions about the test not to be documented at the time those decisions were made. Further, they said that changes to tests are common during implementation and that they focused attention on ensuring the tests were carried out correctly, rather than on documenting the reasons for changes and other decisions as the tests proceeded. However, IRS officials acknowledged that documenting significant events was important.

In some correspondence to taxpayers about the tests, IRS referred taxpayers to LITCs or walk-in sites for assistance. However, IRS did not gather information on or measure the level or quality of assistance

[38] GAO-01-1008G and GAO/AIMD-00-21.3.1.

[39] Treasury Inspector General for Tax Administration, *The Statistical Sampling Method Used in the Earned Income Tax Credit Proof of Concept Test Appears Valid*, 2004-40-100 (Washington, D.C.: May 2004).

provided to test participants at LITCs or walk-in sites. IRS officials said they did not collect these data because they thought taxpayer use of this assistance would be minimal and there was no practical or cost effective way to gather the information. In his response to our draft report, the Commissioner echoed this sentiment, noting that because qualifying child test participants were randomly selected from around the country, efforts to measure services would be extremely difficult. Further, IRS officials did not view this as an implementation problem, but instead viewed it as a limitation of the tests. Whether it is a problem with implementation or test design, there are some important reasons why it would be useful to know the level and quality of services provided. For example, our prior work found that the quality of service IRS walk-in sites provided taxpayers was unknown.[40] Further, face-to-face assistance is costly, especially when compared to telephone services, which were used extensively in the 2004 tests. Recognizing that options for providing taxpayer assistance and outreach efforts are important, if IRS had data on the level and quality of service provided, it would be in a better position to determine the cost and benefit of providing this assistance. Officials recognize that information on use of these forms of assistance would be useful and indicate that they will collect information in conjunction with a planned 2005 simulation of a nationwide test. The simulation is discussed later in this letter.

Lessons Learned from 2004 Tests Prompt Most Refinements for New Round of Tests

IRS officials identified lessons learned from the 2004 tests that were implemented to help improve the 2005 tests. For example, IRS officials plan to use of its automated telephone responses, which is important because most taxpayers contacted IRS by telephone to obtain information about the tests.

Changes to forms and letters based on case reviews and examiner input. IRS officials told us that their modifications to letters and forms for the qualifying child and filing status tests to be used for the 2005 tests were primarily based on case file reviews and discussions with IRS examiners who interacted with the taxpayers selected for the 2004 tests. In April 2004, for example, EITC program officials reviewed case files and met with examiners to discuss common taxpayer errors and questions about letters and forms for the qualifying child and filing status tests. Examples of taxpayer questions were: "How do I prove I did not live with my spouse?"

[40] See GAO-04-84.

or “Who can fill out the affidavit?” Examples of taxpayer errors on forms included no signature, incomplete dates to prove residency, and either no Social Security number listed on the form or an incorrect number. As a result of their review, IRS officials revised the forms containing such information (Form 8836, Qualifying Children Residency Statement and the accompanying affidavit). For example, IRS changed the layout of the affidavit to help reduce taxpayer errors involving dates and the amount of time a child resided with the taxpayer. IRS did not make changes to the letters and forms for the income misreporting test because they were the same ones used under the general AUR program.

Improvements to key database based on examiner suggestions. Examiners in Kansas City, the site responsible for the qualifying child and filing status cases, suggested about 30 updates or other improvements to the EICPC system that they said would either reduce errors in the database, help IRS better manage the cases and workload, or improve subsequent data analyses. For example, examiners noted they were lacking computer fields to record certain information such as the taxpayer’s telephone number, whether the case was worked by the Taxpayer Advocate’s Office, or if an amended return had been received. As a result, examiners suggested ways to capture these data, which have been incorporated into the EICPC. IRS is continuing to update and make improvements to the EICPC.

Use of automated telephone voice response expanded. For the fiscal year 2004 qualifying child and filing status tests, IRS did not use automated responses to answer routine telephone calls and did not have a mechanism in place to obtain taxpayers’ views about telephone services provided. Both options were available for the income misreporting test and are available for users of IRS’s other toll-free telephone numbers. Officials recognized that commonly asked questions, such as “Where do I mail the required documentation?” or “Who can sign an affidavit?” could be answered via automated responses, and plan to provide this option for the fiscal year 2005 tests, leaving only the more difficult questions to be answered by an examiner. IRS also decided to implement a random feedback survey of taxpayers on the quality of service they received for the qualifying child and filing status tests when they called the toll-free number. The survey is a modified version of the one that IRS uses for its regular toll-free telephone operations.

Changes made to the qualifying child test encourage early certification and simulate implementation across the country. There are two major changes to the qualifying child test for 2005: (1) taxpayers will be encouraged to certify in advance of filing their return that their child met the EITC residency requirement; and (2) a portion of the taxpayers will be drawn from a single community—Hartford, Connecticut, while the rest will be drawn randomly from across the nation.

IRS officials contend that an early certification could help reduce delays in releasing EITC refunds because examiners would be able to validate cases before the start of the tax filing season when workloads reach their peak. Eligible taxpayers who provide acceptable documentation before the start of the tax filing season could get their EITC refund more expeditiously, IRS officials say, because the documentation would already be validated at the time taxpayers file their tax returns. IRS has some evidence that taxpayers are willing to certify in advance of the filing season because about 800 taxpayers did so as part of the 2004 qualifying child test, even though they were only asked to do so when filing their returns.

Regarding the targeting of the single Hartford, Connecticut community, IRS officials told us that they intend to simulate what might happen if an early certification requirement were imposed across the country. This change was the result of a recommendation from a contractor's review of the 2004 test's sampling methodology. As part of this test, IRS plans to mount an outreach campaign to include partnering with local governmental and community-based entities to provide taxpayers assistance.

Need for refinement prompts reduction in filing status sample size. Based on its experience with the sample selected for the 2004 filing status test, IRS decided to dramatically reduce the sample size for next year's test, while simultaneously trying to improve the criteria for selecting the sample. As this year's test was implemented, IRS officials realized that the test was yielding a high number of taxpayers claiming the correct filing status, suggesting that the criteria for selecting them could be improved and the burden on taxpayers to prove their filing status was high, relative to the benefits gained. As a result, IRS officials reduced the sample size from 36,000 to 5,000 for the 2005 test to minimize taxpayer burden as IRS works to improve the selection criteria.

IRS also is testing two refinements in the sample selection criteria for the 2005 filing status tests to determine whether the selection criteria can be improved. First, IRS plans to apply TIGTA's finding, which IRS officials said that they had also identified, that IRS could better use information it possesses to verify the filing status of some taxpayers, such as those whose spouses have died or those who have submitted an amended return. Any such taxpayers whose filing status could be verified using such information would not be included in the sample. Second, IRS also plans to refine the sample selection to not include taxpayers whose filing status of single or head of household can be corroborated by information from the third-party locator service, which was tested in 2004.

Income misreporting changes designed to improve sample selection. IRS has planned minimal changes for the 2005 income misreporting test because it found few issues that needed to be addressed. Changes were made to selection criteria to help identify cases with a potentially higher assessment amount. For example, IRS will no longer select cases where the taxpayer's adjusted gross income is over the maximum amount for claiming the EITC and EITC is claimed anyway because IRS found those cases yielded a lower assessment than other cases.

IRS's 2004 Evaluation Plans Lacked Sufficient Documented Detail to Allow for Oversight; Evaluation Plans for 2005 Tests Were Not Completed Before Two of the Tests Had Begun

IRS's plans for evaluating the three 2004 tests lacked sufficient documented detail to facilitate managerial review and stakeholders' oversight and thereby help ensure that the evaluation of the tests' results would be as sound as possible and the results would be communicated with full recognition of their strengths and limitations. For many aspects of IRS's evaluation plans, we were able to discern IRS intentions by piecing together information from multiple sources, including interviews with IRS officials. In essence, an evaluation plan is used to manage the evaluation endeavor. As such, the more completely a plan is developed, the more likely it will be useful to managers in ensuring that the evaluation is well-executed. Despite the importance of having detailed plans prior to implementation, IRS had not completed its evaluation plans for the 2005 tests before two of those tests had begun.

Evaluation Plans Had Strengths Including Linkage Among Test Goals, Evaluation Objectives, and Outcome Measures

Considering the written evaluation plans themselves, interviews with IRS officials, IRS's status report to Congress and other documents, we found that IRS's plans for assessing the three tests had important strengths. For instance, IRS's evaluation plans:

- had clear goals for each the three tests. The primary goal of all three tests was to reduce overclaim rates. There were additional goals for the qualifying child test—maintaining EITC participation for eligible participants and minimizing taxpayer and IRS administrative burden.

- linked evaluation objectives and outcome measures—which determine the extent to which the goals were met—to the test goals. For example, the income misreporting test had outcome measures that included the percentage of cases where an EITC claim was reduced or disallowed and the average amount of the change. These measures were clearly linked to the test's goal of reducing EITC overclaim rates.
- selected samples to provide information that could be generalized to the EITC population being targeted. Both TIGTA and an outside consultant reviewed the samples for the qualifying child test and found that the 23,000 sample for the general test was sufficient—a conclusion with which we also agree. TIGTA also reviewed the samples for the income misreporting test and found that it should provide reliable results.[41]

Lack of Detail and Documentation in Evaluation Plans Undermined Their Value

IRS's evaluation plans for the 2004 tests lacked sufficient documented detail for us to determine how IRS planned to conduct key aspects of the evaluations. When we were able to determine how key aspects of the evaluations would be conducted, we often did so based on interviews and analyses of various documents. The general lack of detail and documentation undermined the value of the plans by, for example, limiting IRS's and stakeholders' ability to oversee the evaluations, identify and address limitations in the evaluations, and ensure that limitations will be clearly communicated when the results are disseminated.

IRS's written evaluation plans for the three tests were essentially outlines that were not comprehensive, meaning that they did not fully document all key aspects of the evaluation. For example, IRS's written plans did not provide information on the sampling methodologies used in all three tests. These were not articulated until IRS issued its August 2004 status report to Congress. In addition to the status report, which also provided additional insights into the types of analyses IRS plans to conduct, we relied on multiple other sources to gain a complete understanding of IRS's planned evaluation activities. We interviewed IRS officials and reviewed other information and documents they provided, such as the contractor's report on the qualifying child test's sampling methodology. According to IRS officials, the lack of comprehensive and detailed written plans was due to

[41] TIGTA 2004-40-100.

other priorities, such as undertaking the numerous steps needed to implement the tests themselves.

While we recognize competing demands on the EITC program office, striking a balance between documenting evaluation plans and implementing and evaluating the tests helps ensure all parties understand the evaluations and the managers and stakeholders are able to oversee implementation and evaluations. Well-developed evaluation plans have a number of benefits, perhaps most importantly, increasing the likelihood that evaluations will yield methodically sound results, thereby supporting effective policy decisions. Such plans help (1) ensure that the agency has addressed the principal aspects of the evaluation, including the research design, outcome measures, target and sample populations, data collection activities, analyses, and dissemination of results, (2) officials monitor changes to tests and assess the impact of those changes on the planned evaluations, and (3) facilitate management and stakeholder review. Having comprehensive and detailed evaluation plans helps ensure that all those working directly on the evaluation have a common understanding of how data will be collected, analyzed and impacts assessed. Concerns or weaknesses can be identified and corrected, and plans can be updated to reflect any changes during implementation and afterwards, as the evaluation plan could be considered to be a "living document." Finally, a well-developed plan helps ensure that evaluation results can be communicated with appropriate recognition of the evaluation's strengths and limitations so stakeholders can better understand how to use the results when making decisions.

The following are illustrations of the overall lack of detail and documentation in IRS's evaluation plans for the 2004 tests.

- **Evaluation objectives were not documented in one place.** Although we found that IRS's evaluation plans had objectives linked to the test goals, the objectives were not identified in any single location for any of the three tests nor specifically identified as objectives. Thus, we pieced together the information from multiple sources, including interviews with IRS officials. For example, we had difficulty identifying the evaluation objectives pertaining to the use of the third-party locator service for the filing status test.

- **Key outcome measures lacked important detail.** IRS's evaluation plans lacked important information for all the key outcome measures, such as their definition and purpose, formula/methodology, data source and collection method. For example, IRS's evaluation plans for the qualifying child test did not identify the specific data that would be used to produce the outcome measure—the number of taxpayers who claim (or do not claim) the EITC. IRS has provided this type of information about its measures for other programs. For example, for its telephone and other operations, IRS annually prepares a comprehensive document known as a data dictionary, which includes items such as the definition and purpose of the measure and its formula/methodology. IRS officials agreed that providing such information in the evaluation plans could have been valuable in managing the EITC tests. Without knowing details on outcome measures, stakeholders do not have enough information about a measure to know whether it is valid and reliable.

- **Limited information was provided on planned analyses.** The evaluation plans also lacked specificity with regard to the key analyses planned and what those analyses were intended to accomplish. For example, IRS conducted a survey to obtain information about a taxpayer's experience with the qualifying child test. IRS originally planned to survey these taxpayers in April 2004. The survey was not conducted until September 2004, primarily due to delays in selecting a contractor and developing the survey instrument. The 5-month delay may substantially reduce the number of taxpayers who accurately remembered the actions they took and thus affect the quality of the responses (i.e., recall bias). The accuracy of individuals' survey responses declines the further away those responses are from the date of the actual events. IRS and the contractor are aware that such recall bias could exist and stated that they will consider it when analyzing the survey results, but no detail was available on how they would do so. This is critical because the potential utility of the survey results could be in question.

The lack of detail in IRS's evaluation plans also increased the risk that reports disseminating the results of the tests would not fully disclose the evaluations' potential limitations. In its August status report to Congress, IRS did not make clear that the qualifying child test results could only be generalized to taxpayers IRS had reason to believe were most likely to make an erroneous claim for the EITC when filing for the EITC in 2002. Absent such clarity, stakeholders might incorrectly assume that test results apply to all taxpayers claiming qualifying children for the EITC. Also, IRS did not describe potential limitations of the outcome measures, specifically, how non-respondents would be accounted for in measure calculations.

IRS officials recognize that their final 2005 report will need to include information on the evaluation limitations, and expect to provide sufficient detail and explanation of limitations in that report.

Evaluation Plans for 2005 Tests Not Completed Before Two of the Tests Began

As of early December 2004, IRS had not completed evaluation plans for the 2005 testing, even though the qualifying child and income misreporting tests began in November. According to IRS officials, they had not yet completed an evaluation plan for the 2005 tests because final decisions about the testing were still being deliberated in October. In their view, it was less important to finish an evaluation plan for these tests by the time testing began, because IRS could use the 2004 evaluation plans in the interim. IRS officials acknowledge that evaluation plans are important and have started to develop them for the 2005 tests.

IRS can build upon the 2004 evaluation plans for all three tests. However, IRS made substantial changes for the qualifying child and filing status tests, which would need to be taken into account in developing comprehensive and detailed evaluation plans for the 2005 tests. Therefore, while we recognize that there will be similarities with the 2004 evaluation plans, the importance of having evaluation plans in place as testing begins or soon thereafter is heightened because of planned changes to the test.

Conclusions

The EITC program lifts millions of low-income taxpayers and their families out of poverty. However, its high rates of noncompliance—overclaims for the credit—could potentially undermine the credibility of the program because billions of dollars are annually paid out that should not have been. IRS's three tests—qualifying child certification, filing status, and income misreporting—are major initiatives to reduce overclaims by addressing the leading errors taxpayers make. Given the importance of the EITC to many low-income households and concerns about high overclaims, these tests are being closely watched by numerous stakeholders.

Although IRS has generally implemented each of the tests smoothly, it did not fully document some key management decisions and other significant events. Documentation supports a common understanding among staff about the program they are administering—particularly one as complicated as the EITC—and helps managers monitor whether a program is implemented as planned. Having adequate documentation during the 2005 tests could help foster a better understanding of the tests, ensure results are accurately determined, and facilitate review and oversight. In addition, while IRS told taxpayers selected for the qualifying child test they could visit various physical locations for assistance, including LITCs and IRS walk-in sites, IRS did not collect information from those sites to determine the level and quality of services provided. Because officials believe relatively few taxpayers used these sites, collecting information from the sites may not have been practical. However, the single city simulation of nationwide implementation may offer an opportunity to gather some information on these services.

The evaluations that IRS is conducting of each test are likely to yield some useful information and results that will help IRS officials and other stakeholders judge whether and how to proceed with further implementation of the new approaches to reducing EITC overclaims. Nevertheless, the lack of detail and documentation in the evaluation plans impeded officials' ability to manage the evaluations as well as external stakeholders' ability to review and understand the evaluations' strengths and limitations.

As of early December 2004, IRS had not completed its 2005 evaluation plans, although testing was underway for the qualifying child and income misreporting tests. A well-developed and timely plan would help IRS to improve on the 2004 evaluation plans and take into account changes in the tests themselves.

Recommendations for Executive Action

The Commissioner of Internal Revenue should

- adopt a policy of documenting the rationale for key policy decisions and other significant events as the 2005 tests are implemented;
- develop a means of gathering information during the 2005 tests on the use of such locations as LITCs and walk-in sites on the level and quality of service provided by those sites, particularly in light of IRS's plans to draw its sample from a single community for the qualifying child test;
- ensure that reports disseminating the results of the 2004 and 2005 test evaluations clearly outline aspects of test design and evaluation shortcomings that limit the interpretation and utility of the results; and
- complete the development of comprehensive and adequately detailed evaluation plans for the 2005 tests.

These actions should be done as soon as possible, with any significant changes to the evaluation plan appropriately documented as the evaluation unfolds.

Agency Comments and Our Evaluation

In his December 22, 2004 letter, the Commissioner of Internal Revenue agreed with our recommendations. Regarding the issue of documenting significant policy decisions, he noted competing demands that can often affect the quality of documentation, which we acknowledge in our report, and that IRS has implemented a process to meet this recommendation. The Commissioner noted that providing taxpayers with assistance is a top IRS priority. As such, the Commissioner reported that IRS has plans to identify the level and quality of services provided to taxpayers at LITCs and walk-in sites in the single test community. Regarding dissemination of results, the Commissioner reported that IRS is committed to ensuring all aspects of the test design and evaluation will be clearly described to stakeholders. Finally, the Commissioner reported that IRS intends to complete the 2005 evaluation plans, in part, based on GAO's recommendations about what a plan should contain. He also noted that IRS may need to assess whether any modifications to the 2004 qualifying child test criteria are appropriate in light of public events and community leadership reaction in the single test community.

We are sending copies of this report to the Chairmen and Ranking Minority Members of the Senate Committee on Finance and the House Committee on Ways and Means. We are also sending 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 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.

This report was prepared under the direction of Joanna Stamatiades, Assistant Director. Other major contributors are acknowledged in appendix IV. If you have any questions about this report, contact Ms. Stamatiades at (404) 679-1900 or me on (202) 512-9110.

Michael Brostek
Director, Strategic Issues

Scope and Methodology

For all three objectives, we reviewed and analyzed documents including Treasury's EITC compliance study of 1999 tax returns; a joint IRS/Treasury task force report; monthly status reports for each of the tests; draft and final letters, forms, and notices for each of the tests; implementation and evaluation plans; our prior reports; status results of the tests reported by IRS and its contractors; and reports and EITC literature by external stakeholders. We also interviewed Department of the Treasury and IRS officials involved in the EITC tests, including the National EITC Director, National Taxpayer Advocate, Director of Research Analysis and Statistics, and other IRS officials involved with implementing the tests. Additionally, we interviewed external stakeholders such as individuals at the TIGTA, Center for Budget and Policy Priorities, and Urban Institute, and reviewed and analyzed their reports.

We took steps to ensure that the data we received from IRS were reliable for the purposes of this report and determined that they were. Some of those steps included interviewing IRS officials knowledgeable about the computer systems where the data we obtained came from and reviewing documentation, such as system manuals and flowcharts. We identified and assessed potential data limitations and compared those results to our data reliability standards, noting no significant weaknesses. [1]

In addition, to describe the three tests and determine how IRS was spending the money appropriated it for the tests, we interviewed managers and budget officials in the EITC Program office and reviewed and analyzed IRS's fiscal year 2004 budget request and compared its planned to actual EITC spending plan. Because IRS does not have an adequate cost accounting system, we could not verify the accuracy of the figures IRS provided to describe how funds appropriated for the tests were spent.

We identified attributes of sound program implementation based on reviews of the social science literature, our prior work, and interviews conducted with IRS research and program management officials and external stakeholders, such as the Urban Institute.[2] We tailored these

[1] GAO, *Assessing the Reliability of Computer-Processed Data*, GAO-03-273G (Washington, D.C: October 2002).

[2] The three key sources we used to develop our implementation and evaluation plan criteria were: P.H. Rossi, and H.E. Freeman. *Evaluation: A Systematic Approach*, 4th ed. Newbury Park, Calif.: Sage, 1989, P.H. Rossi, H.E. Freeman, and M.W. Lipsey. *Evaluation: A Systematic Approach*, 6th ed. Thousand Oaks, Calif.: Sage, 1999, and GAO/PEMD-10.1.4.

attributes to apply them specifically to IRS's tests as shown in table 3. Finally, to assess how well IRS implemented the tests and determine IRS's planned refinements for further testing in fiscal year 2005, we reviewed policies, procedures, and training documents; observed procedures and operations in Kansas City, Missouri; Atlanta, Georgia; and Fresno, California; and interviewed front line IRS managers and examiners in these locations. We reviewed several case files for each test. Additionally, we analyzed relevant interim reports prepared by IRS and its contractors; and identified key results, and discussed them with IRS officials.

Table 3: Criteria for Assessing IRS's Test Implementation

Did IRS implement the test largely as planned?
Did IRS provide the necessary staff and training to accomplish the workload associated with each test?
Did IRS monitor the delivery of taxpayer correspondence?
Were taxpayers provided with a way to contact IRS or another entity if they had questions and, if they did, did IRS assess the quality of the interactions and the accuracy of the information provided?
Did IRS implement quality control procedures to ensure the quality of decisions IRS staff made?

Source: GAO.

To assess whether IRS's plan for evaluating the tests contained sufficient documented detail to facilitate managerial review and stakeholder oversight, we used GAO guidance and the social science evaluation literature to identify key attributes of an evaluation. These attributes included the research design, outcome measures, target and sample populations, data collection activities, analyses, and dissemination of results. We obtained all available documentation on IRS' s evaluation plans for each of the tests and reviewed that documentation to determine whether we could understand from the documentation alone how IRS planned to address the key attributes. Where we could not, we interviewed IRS officials to further understand whether and how the officials planned to address those key attributes. Written documentation should be complete, facilitate tracing of events, and be readily available for examination to foster a common understanding of the program and facilitate oversight. To describe the status of IRS's evaluation plan for the fiscal year 2005 tests, we primarily relied on interviews with IRS officials.

Updated Results from the Income Misreporting Test as of September 30, 2004

In August 2004, IRS issued a status report to Congress, which was mandated by the Consolidated Appropriations Act of 2004.[1] The report presents an overview of each of the three EITC tests, along with the design, status, and preliminary findings as of June 2004. According to the EITC National Director, the report contained some of the types of information that will be needed to support future decisions about the full implementation of the tests. Additionally, the EITC National Director noted that IRS also used the status report to provide information on such items as the sampling strategy that have been lacking in other documents.

IRS had updated results for the income misreporting test as of September 30, 2004. Updated results were not available for the qualifying child or filing status tests. As IRS stated in its status report, which showed data as of June 26, 2004, it is important to note that because the results are interim, no conclusions should be drawn from the information provided and no analyses about the impact of the tests were included.

As table 4 shows, IRS has screened all 300,000 tax returns for the income misreporting test, and slightly more than half have been closed with taxpayer agreement.

[1]Pub. L. No. 108-199, 118 Stat. 3 (2004).

Table 4: Status of Income Misreporting Test, June 26, 2004 and September 30, 2004 Results

	June 26, 2004		September 30, 2004	
	Number	Percent	Number	Percent
Total number of cases in AUR EITC test	300,000	100%	300,000	100%
Number of notices sent	261,169	87%	261,188	87%
Total number of cases closed [a]	102,545	34%	196,600	66%
Number of cases closed with taxpayer agreement [b]	66,981	22%	155,446	52%
Number of cases screened out [c]	38,831	13%	38,812	13%
Number of cases closed with no change [d]	35,564	12%	41,154	14%

Source: GAO analysis of IRS data.

[a] Total number of cases closed includes the number of cases closed with taxpayer agreement and number cases closed with no change.

[b] Number of cases closed with taxpayer agreement includes cases when an examiner determines a discrepancy in the case, and after sending a notice to the taxpayer, an additional tax assessment is made and the taxpayer agrees with this assessment.

[c] Number of cases screened out includes cases when an examiner reviews the case and determines the computer mismatch (i.e., a math or other error) can be resolved and thus closes the case without a change to the tax. According to IRS officials, the number of cases declined slightly during the period shown because some of the cases were transferred to examination.

[d] Number of cases closed with no change includes cases when the taxpayer is able to provide substantiation and thus the change is closed without a change.

Comments from the Internal Revenue Service

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

COMMISSIONER

December 22, 2004

Mr. Michael Brostek
Director, Tax Issues
U.S. Government Accountability Office
441 G Street, N.W.
Washington, D.C. 20548

Dear Mr. Brostek:

I am responding to your draft report titled, *EARNED INCOME TAX CREDIT: Implementation of Three New Tests Proceeded Smoothly, But Tests and Evaluation Plans Were Not Fully Documented* (GAO-05-92). The implementation of these tests will provide us with valuable information about key components of our long-term vision for this critical program. As a result, we are pleased that your review showed that the implementation of these tests had proceeded smoothly.

We agree with your recommendation to increase documentation, though as we have noted, much of this problem stemmed from managing the competing needs of documenting our evaluation plans while actually planning, implementing and refining the tests. When issues arose, we responded to them quickly and decisively. As a result, documentation may not be as robust as it might otherwise be. Nonetheless, as you noted, the tests themselves and the evaluation efforts associated with them proceeded smoothly and will provide valuable information on a potential certification process.

Your report notes we did not measure the level and quality of services provided to test participants in the walk-in sites and Low Income Taxpayer Clinics. Because these 25,000 certification test participants were randomly selected from around the country, efforts to measure services and support provided to these specific taxpayers in each location would be extremely difficult and unlikely to produce meaningful results. However, we are capturing telephone and walk-in traffic at IRS' offices and plan to expand these efforts for future tests.

We agree that measuring "third party" impact – burden on community-based organizations and institutions (e.g., schools) – is an important aspect of certification testing. That is why we have included it in our evaluation plans for our 2005 test. That test includes a single community or geographic area in which all individuals who met the certification criteria are asked to certify.

In terms of the 2005 evaluation plan, we are working to finalize key elements. However, our efforts have been delayed by two factors. First, we want to assess whether any

2

modifications to the 2004 criteria are appropriate in light of public events and community leadership reaction to certification in our test community. Second, while much of our 2004 plan is quite strong, your report provides valuable insights and opportunities for improvement, which we want to incorporate prior to finalizing a 2005 version.

Responses to your specific recommendations are enclosed. I appreciate your observations on the implementation of our EITC tests. If you have any questions, please contact Floyd Williams, Director, Legislative Affairs, at (202) 622-3720.

Sincerely,

Mark W. Everson

Enclosure

Enclosure

Recommendation for the Commissioner

Adopt a policy of documenting the rationale for key policy decisions and other significant events as the 2005 tests are implemented.

Response

We agree that better documentation is needed to show the basis for all significant policy decisions, and in response, we have developed a process and implemented it for the 2005 tests. While historical documentation about some aspects of the tests was somewhat lacking, these policy decisions were made by senior IRS executives after careful consideration. We believe this degree of oversight was instrumental in successfully implementing the 2004 tests. To this end, we are pleased that your review also identified that these tests generally followed the recommendations of the IRS/Treasury Task Force with departures based on informed decisions.

Recommendation for the Commissioner

Develop a means of gathering information during the 2005 tests on the use of such locations as LITCs and walk-in sites on the level and quality of service provided by those sites, particularly in light of IRS' plans to draw its sample from a single community for the qualifying child test.

Response

Our EITC mission is twofold – reduce erroneous payments while increasing participation by eligible taxpayers. As such, providing taxpayers with the assistance they need to accurately complete the certification process has been, and will continue to be, one of our top priorities. We accomplished this during the 2004 qualifying child test, in part, by providing test participants with a dedicated, toll-free number to quickly and accurately obtain assistance. We also provided all of our employees who interact with taxpayers, including those in the walk-in sites, relevant information about the certification test. Since the 25,000 test participants lived throughout the country, it is unlikely that many walk-in sites received requests for this type of assistance. Since a portion of our 2005 certification test is geographically concentrated in a single community, we have expanded our tracking system to identify taxpayer contacts regarding the certification test. This tracking process includes both walk-in and LITC sites in the selected community. We will use this information as well as information from our quality assurance system to evaluate the level of customer service provided to test participants.

2

Recommendation for the Commissioner

Ensure that reports disseminating the results of the 2004 and 2005 test evaluations clearly outline aspects of test design and evaluation shortcomings that limit the interpretation and utility of the results.

Response

We are committed to engaging stakeholders throughout this process to obtain their suggestions and to address their questions and concerns. To this end, we will clearly describe all aspects of the test, from design through execution and evaluation, in our report of the 2004/2005 test results.

Recommendation for the Commissioner

Complete the development of comprehensive and appropriately detailed evaluation plans for the 2005 tests.

Response

The IRS developed a detailed methodology to evaluate the 2004 test results. This evaluation included independent analysis by contractors who reviewed key programmatic components, including the design of the qualifying child test and a follow-up survey of test participants. Many of the 2004 evaluation criteria will be carried forward to the subsequent test along with enhancements derived from lessons learned during the 2004 test. However, as previously noted, before we completed our 2005 evaluation plan we wanted to ensure that we included all components of your recommended evaluation criteria.

GAO Contacts and Staff Acknowledgments

GAO Contacts

Michael Brostek (202) 512-9110
Joanna Stamatiades (404) 679-1900

Staff Acknowledgments

In addition to those named above, Tom Beall, Evan Gilman, Veronica Mayhand, Susan Mak, Donna Miller, Libby Mixon, Chris Moriarity, Ed Nannenhorn, Cheryl Peterson, Michael Rose, and Robyn Trotter 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

Susan Becker, Acting Manager, BeckerS@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548

United States
Government Accountability Office
Washington, D.C. 20548-0001

Official Business
Penalty for Private Use $300

Address Service Requested

Presorted Standard
Postage & Fees Paid
GAO
Permit No. GI00