OMB Clearance Package
Interactive Multilingual Tax Application Usability Testing
Language Services Office

The following is a submission to test a multilingual educational tool about income tax returns developed by IRS. The audience is Spanish-speaking taxpayers who have difficulty understanding English.

IRS wants to conduct two different types of test events. One test, "the pilot," would be given to taxpayers waiting for their appointment at a Taxpayer Assistance Center. That test, considered a pilot, is intended to get a quick understanding of taxpayers' reaction to the tool. The second, the "in-depth" test, is a more rigorous study with a small group of dedicated testers. Participants will be recruited well in advance and it will be conducted in a special testing facility.

Both tests are described separately. The burden hours they will require are given separately and together.

Background

Executive Order 13166 (EO 13166), "Improving Access to Services for Persons with Limited English Proficiency" (LEP) requires Federal agencies to provide meaningful access to persons who have difficulty speaking and understanding English*. In response, IRS' Language Services Program assesses the needs of and develops educational material for the LEP. Our goal is to help LEP persons comply with their tax responsibilities and to capture all the revenue IRS is entitled to collect.

Through the Customer Base and Needs Assessment processes, IRS learned that 70% of the LEP are Spanish-speaking (about 9.4 million actual or potential taxpayers), and that a significant proportion of them (almost 70%) cannot read. This is an enormous population of taxpayers that might not understand their tax obligations and therefore may not satisfy them. The Language Services Program developed an interactive computer-based application as a tool for educating Spanish-speaking LEP taxpayers about their tax responsibilities and Federal income tax returns.

The application is available on DVDs, each having 2 complete versions: one in Spanish, and one in English. The application has video-format narrators who discuss six critical tax-related topics (such as who must file, how to file, and what to do if contacted by IRS). The narration is accompanied by supporting graphics, examples, interactive activities, glossaries, downloadable documents, quizzes, and hyperlinks. Users can listen to the lessons; there is also complete text of the narration for the hearing disabled.

_____

* based on primary language, not disability

The Spanish-speaking LEP population is dispersed throughout the United States. Their country of origin or heritage varies from Mexico to South America to Central America, to the Caribbean, Spain, and more. Customs and phraseology for people from each heritage differs widely. Making sure that Spanish-speakers from all heritages understand a single Spanish product is a challenge.

The Language Services branch needs to test the effectiveness and customer satisfaction of this application in two test events: one in the eastern US, and one in the western US. By testing in two geographically diverse locations, we will be likely to test participants from all Spanish-speaking heritages and cultures.

Purpose

This OMB request is to obtain approval for the Language Services Program test the English and Spanish versions of the interactive "Guide to Federal Income Taxes." One of the two test events will take place in an IRS Taxpayer Assistance Center (TAC). Both test events will solicit feedback from Spanish-speaking (only), bilingual, and English-speaking (only) customers. Key areas of testing include the application's:

- **Educational value** – do users learn from it and does it address the topics they are concerned about?
- **Presentation** – is the arrangement and presentation of the information in a form that enhances its educational value and enables answers to easily be found?
- **User interface** – Is the user interface set up so that navigation and options are clear and intuitive?

Tests

Two distinct types of testing events will be performed. One will serve as a "pilot," and the other will be more in-depth. The test plan for both types of events is attached.

Pilot test event

The pilot test event will involve actual TAC customers as they are waiting for service. Customers with differing levels of Spanish proficiency will be asked to find answers to tax questions using the application. Over a period of three days, approximately 30 TAC customers will spend about 20 minutes each participating.

They will explore the application to find an answer to a question. They may use one of their own questions for which they came to the TAC, or they will be given a question by the test team. The time it takes them to find the answer or give up looking will be measured. They will then be asked, subjectively, how adequate the answer was and what their experience was finding it. They will also be asked

if they thought the information they found in the application would reduce the amount of time they needed to spend with the TAC assistor.

In-Depth Test Event

The in-depth test event will take place in an IRS-provided space. The test area will be set up only for the test (and not for any other purpose, as was the pilot test in the TAC). Each participant will be given both quantitative and qualitative tests. Any one given test will not require the user to view the entire application, as this would take an unnecessary amount of time.

The quantitative test will be one or more scenarios for the tester to go through to evaluate the educational value, presentation, and user interface. The evaluation of the application will be measured, for example, in how many questions the tester answers correctly on the quiz that ends each topic. Testers will be given up to 2.25 hours to complete their scenarios.

The qualitative test will be an interview with each tester to discuss their reactions to the application. They will be asked if they learned something about taxes that they did not already know (and what), if the user interface enhanced the applications' effectiveness, if there were any parts they think can be improved, and follow-up feedback questions depending on their answers. Testers will be interviewed for up to 0.75 hours.

This overall test will take up to three hours for each participant.

Methodology

The testing will be given in two events. One event will be in the eastern US and one in the western US. Testing in several US regions enables IRS to capture feedback from testers of different heritage and cultures.

The testing will take place in IRS facilities. The pilot test will be in a Taxpayer Assistance Center (TAC), a free walk-in office that taxpayers can come to discuss tax matters with IRS employees. A TAC with Spanish-speaking bilingual assistors will be selected. The second facility will be either in a TAC if the pilot test event yields significant results, or in a different kind of IRS-provided facility if needed.

Both versions of the application, Spanish and English, will be tested.

Pilot test event

The pilot test event will involve actual TAC customers as they are waiting for service. Bilingual IRS staff will canvas them. The selected customers will represent different levels of Spanish and English proficiency. Customers will be asked to find answers to tax questions using the application. Over a period of three days during November 2010, approximately 30 TAC customers will be approached and:

- Asked their proficiency in English and Spanish (as a selection criterion)
- Asked if they would participate in a 20-minute test for which they will receive compensation
- Told that arrangements can be made so they will not lose their place in line
- Ensured that none of their personal information will be used or collected
- Ensured that their participation has no connection to or reflection on their status with IRS or any other government agency
- Offered a $15 stipend for their time.

After completing some minor paperwork confirming that they received the above orientation, they will be accompanied to a testing area. The testing area will have one or more computer workstations with the application running. Testers will be introduced to how the application works. They will then begin to explore the application to find an answer to a question. They may use one of their own questions for which they came to the TAC, or they will be given a question by the test team. The time it takes them to find the answer or give up looking will be measured. They will then be asked, subjectively, how adequate the answer was and what their experience was finding it. They will also be asked if they thought the information they found in the application would reduce the amount of time they needed to spend with the TAC assistor. Test completion will be marked by giving them their stipend.

Following their meeting with the TAC assistor, the assistor will be asked if the customer seemed well-prepared for their meeting because of having used the application.

Each day of testing will represent a different time of day: morning, afternoon, and evening (if possible).

Costs for IRS include travel, material, and stipend. Costs for employees to participate will be divided between long-distance and local travel. The pilot test will occur in the Washington DC area. The usability testing specialist will have to travel for a period of four to five days from Ogden, UT at an estimated total cost of $2000. Two Washington DC-based IRS employees will participate. Their costs are for local travel only, and estimated to be no more than $15/day for a total of no more than $60 (if four days are involved).

Material costs will be incurred for handouts and supplies. Printing costs for the response forms will be approximately $45 and other supplies (pens, etc) will be no more than $5.

Lastly, the cost for the stipends, 30 people @$15, would be $450.

Total costs for the pilot test event are estimated to be $2060 for local and long-distance IRS employee travel, $50 for supplies, and $450 for stipends = $2560.00.

In-Depth testing event

Twenty testers will be recruited to participate. They will represent a variety of levels of proficiency in Spanish. Ten testers will speak only Spanish, and little or no English. Six testers will be bilingual in Spanish and English. Four testers will be fluent in English and understand little or no Spanish. A recruitment firm will be engaged to identify the testers. They will select candidates who:

- Have the language proficiency required
- Have basic knowledge in operating a computer and interacting with a computer application
- Are over 18 years of age
- Are actual or potential taxpayers
- Have a "layperson's" understanding of taxes and tax law (i.e., they are not tax professionals)

The testing will take place over a five-day span during December 2010. Most tests will be given during the day, but evening testing can be scheduled if possible.

Three IRS employees will be present for all the testing. One will be a subject matter expert in usability testing. One will be the project manager who developed the applications. The third will be fluent in Spanish and familiar with IRS language services. These IRS employees will introduce the product and test to the participants, conduct the interviews, and be available to answer questions.

A room for the testing will be outfitted with computer workstations and writing materials. The DVD with the applications will be already loaded into the computers. A handout with scenarios will be given to each tester and their specific scenarios identified.

Testers will be given up to 2.25 hours to complete their scenarios and answer the questions about their understanding. That will be followed by 45-minute one-on-one interviews in which their experience will be discussed.

Testers will be given a $100-$125 stipend for their participation, depending on which city is selected. The stipend is higher than for the pilot test because participants may have to travel to the test, and the test will take longer.

The results will be captured on the scenario forms the testers fill out and IRS employees' notes from the interviews.

Costs for IRS include travel, material, and recruitment.

"Recruitment" includes all the activities to contact and bring in participants as well as pay them stipends. Those activities will all be performed by a contracted vendor. They incur the material and labor costs for making contact, sending out letters and planning materials, and paying stipends. The cost for IRS to use a vendor for these activities is $15,000.

Costs for employees to participate will be long-distance travel. The pilot test will occur in the western US, presumably in Southern California. All three participating IRS employees will have to travel at an estimated total cost of $6000. Material costs will be incurred for handouts and supplies. Printing costs for the response forms will be approximately $45 and other supplies (pens, etc) will be no more than $5. Shipping costs may be incurred, at an estimated $50.

Total costs for the in-depth test event are estimated to be $6000 for local and long-distance IRS employee travel, $50 for supplies, and $15000 for the recruitment firm = $21,050.00.

Total estimated costs for both events = $23,610.00.

Efforts Not to Duplicate Research

These applications are brand new and unique. Research on their effectiveness has never been tested.

Efforts to Maximize Response Rate

The Language Services Program designed the tests to maximize response rate and limit taxpayer burden. A monetary honorarium compatible with similar testing and appropriate for the region will be provided to participants.   Light refreshments appropriate for the time of day will also be offered.

Burden Estimate:

Pilot test event

(anticipating that 100 people will be screened to select 30 participants)

Screening time: 5 minutes per person
5 minutes x 100 people = 500 minutes (8.3 hours)

(anticipating that 30 people fill out paperwork, listen to a five-minute introduction, and be set up in their work stations)

Paperwork, introduction, and set-up time: 10 minutes per person
10 minutes x 30 people = 300 minutes (5 hours)

(anticipating that 30 will take the full twenty minutes to perform all tests and interviews)

Testing time: 20 minutes per person
20 minutes x 30 people = 600 minutes (6 hours)

Total burden for Pilot Test Event (screening and survey participation) = 19.3 hours

In-depth test event

(anticipating that 100 people will be screened to select 20 participants)

Screening time: 20 minutes per person
20 minutes x 100 people = 2000 minutes (34 hours)

(anticipating that 20 people might make a trip by specifically to participate in the test)

Travel time: 40 minutes per person to and from testing
40 minutes x 20 people = 800 minutes (13.3 hours)

(anticipating that 20 will pick up some refreshments, listen to a 20 minute introduction, and be set up in their work stations)

Introduction and set-up time: 20 minutes per person prior to testing
20 minutes x 20 people = 400 minutes (6.7 hours)

(anticipating that 20 will take the full three hours to perform all tests and interviews)

Testing time: 3 hours per person
180 minutes x 20 people = 3,600 minutes (60 hours)

Total burden In-depth test event (screening and survey participation) = 114 hours

Total burden both pilot and in-depth test events = 133.3 hours

Privacy, Security, Disclosure, and Anonymity

The IRS will ensure that privacy, security, disclosure, and anonymity is upheld when gathering and reporting results from the survey. The recruiting contractors and participating IRS staff will be told specifically not to reveal, either during the screening or when summarizing the results, information that will reveal personal information. During screening and testing, participant names will not be used. Collected information will be limited to that necessary to accomplish the testing objectives. Participants will be told not to reveal personally identifiable information or use their real tax-related information for the tests.

Other Information

The following information will be provided to Statistics of Income, Statistical Support Section within 60 days after completing the second testing event:

Findings: A brief summary of key findings.

Actions taken/Lessons learned: A brief summary of any action taken or lessons learned as a result of the findings.

Final Number of completed participants

Pont of contact:

Carolyn Schaffer, Management and Program Analyst
IRS, Language Services Branch
202-283-4528

Attachments:
1. Test scenarios
2. Interview Questions
3. Moderator's Guide
4. Screener's Guide