

**Strengthening Adult Reading
Instructional Practices
200807-1830-001**

*Supporting Statement for
Request for OMB Approval of Data Collection*

Part A: Justification

Date Submitted:

TABLE OF CONTENTS

Introduction	1
1. Circumstances that Make Collection of Information Necessary	1
2. How the Information will be Used, by Whom, and for What Purpose	4
3. Use of Technical Collection Techniques and Information Technology	7
4. Identifying Duplication	8
5. Impact on Small Entities	8
6. Consequences if Study is not Conducted or if Conducted Less Frequently	8
7. Special Circumstances	9
9. Payment of Gifts	9
10. Assurances of Confidentiality	9
11. Questions of a Sensitive Nature	10
12. Estimate of Annual Hour Burden	11
13. Estimate of Total Cost Burden	13
14. Estimate Cost to the Federal Government	13
15. Reason for Program Changes or Adjustments	13
16. Publication Results of Data Collection	13
17. OMB Approval for Not Displaying OMB Date	20
18. Exceptions to the Certificate Statement	20
References	21

List of Appendices (Attached as individual Word files)

- Appendix A: WIA, Title II, Section 243
- Appendix B: Learner Background Interview Protocol, Form A
- Appendix C: Learner Background Interview Protocol, Form B
- Appendix D: Instructor Background Characteristics
- Appendix E: Class Observation Form
- Appendix F: Instructor Interview
- Appendix G: Instructor Log
- Appendix H: ABE Program Protocol
- Appendix I: Program Participation Letter
- Appendix J: Learner Consent for Participation in Research Form
- Appendix K: Introductory Letter to Program from ED

Supporting Statement for Request for OMB Approval of Data Collection

Part A. Justification

Introduction

This document presents Part A of the Supporting Statement for “Strengthening Adult Reading Instructional Practices” (SARIP), a study sponsored by the U.S. Department of Education’s Office of Vocational and Adult Education (OVAE).

A.1. Explain the circumstances that make the collection of information necessary. Identify any legal or administrative requirements that necessitate the collection. Attach a copy of the appropriate section of each statute and regulation mandating or authorizing the collection of information.

Need for the Study. The SARIP study is an initial investigation of whether the Student Achievement in Reading (STAR) training and materials are effective in developing adult basic education (ABE) instructors’ capacity to deliver evidence-based reading instruction and, consequently, in improving intermediate-level (4th-8.9th grade equivalence) adult learners’ reading skills. The Office of Vocational and Adult Education began STAR in 2005 as a pilot project to build state capacity to implement research-based reading reform in adult education classrooms. The STAR toolkit, which contains information and resources to improve reading classroom instruction in ABE, was developed to be used in training and in providing technical assistance to local ABE administrators and ABE instructors. The information used in creating the toolkit was based on the body of knowledge on effective reading practices developed by the Partnership for Reading and summarized in Kruidenier (2002).

STAR began with a pilot phase, during which OVAE worked with state adult education administrators, professional developers, local ABE administrators, and classroom instructors in six states to test the STAR toolkit and training. Forty-four ABE programs and 144 instructors were involved in the STAR pilot project. As a result of the pilot project, STAR has grown to a national initiative with a National Technical Assistance Team, an online professional development system, and national dissemination activities. As additional states have shown an interest in participating in STAR training, there is a need to collect preliminary information about the effects of STAR on learners’ reading development. These data can assist OVAE in planning future dissemination activities and in determining whether a more rigorous evaluation of STAR is appropriate at this stage of STAR’s implementation.

Study Design and Sample. The SARIP study will employ a quasi-experimental design to examine whether learners who are taught by ABE instructors that have been trained in the STAR methods and materials and have become proficient in these methods make greater gains in developing their reading skills compared to learners who have been taught by ABE instructors that have not participated in STAR. Criteria are being developed for determining high-implementing, STAR-trained instructors, and a sample of high-implementing, STAR-trained

instructors will be selected based on these criteria. The adult learners in the reading classes taught by the sample of high-implementing, STAR-trained instructors will constitute the treatment group for the SARIP study. The treatment learners will be compared to data from a matched sample of adult learners that have been taught by ABE instructors who have not participated in STAR training. This comparison group will be drawn from extant data that were collected in either of two previous studies conducted by the contractor (Abt Associates Inc.), which investigated intermediate-level learners' development of reading skills in ABE classes (*Study of Effective ABE Programs and Practices for First-Level Learners*¹ and *Building a Knowledge Base for Adult Decoding*²). The learner, instructor, class, and ABE program data in these two studies were collected using the same instruments that will be used in the SARIP study. Thus there will be comparable data for the treatment and comparison groups.

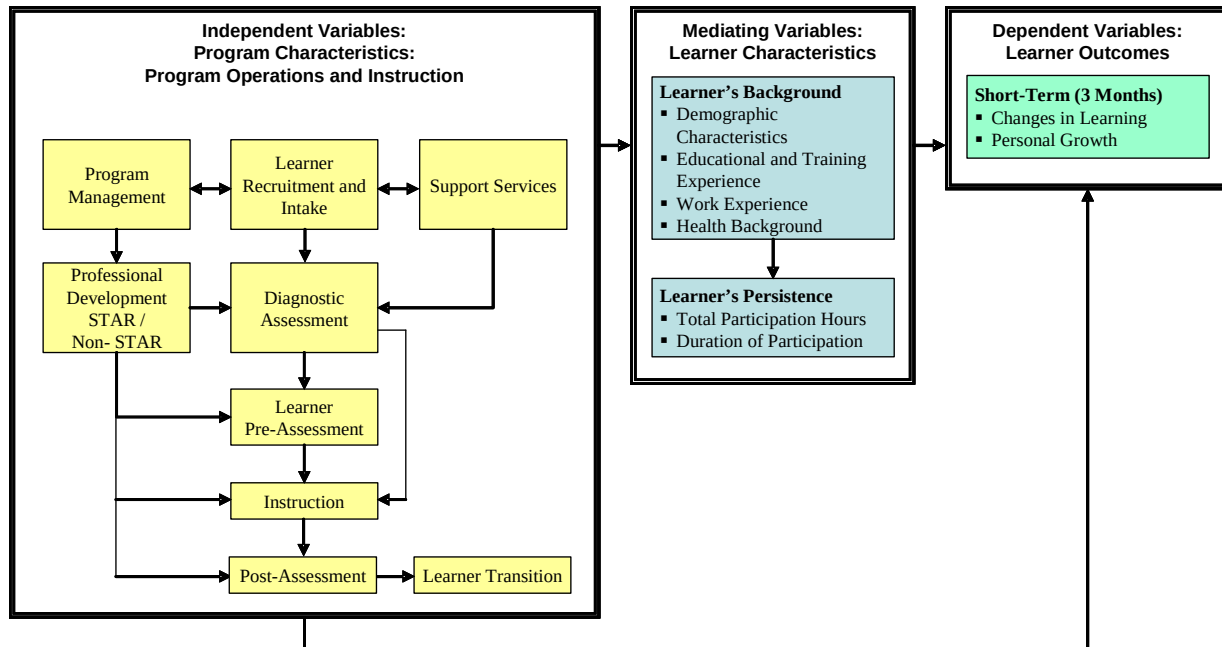
Study Framework. The contractor developed a conceptual “model of change” for ABE programs that is presented in Exhibit 1. The model specifies two broad categories of programmatic variables that are predicted to be related to the amount of learner change: overall program operations (e.g., whether or not instructors received professional development in STAR, whether instructors use diagnostic reading assessments to guide instruction) and instructional content and strategies (e.g., the content and strategies related to STAR). The study framework posits that program participation in STAR professional development will lead to changes at the ABE program level in the use of diagnostic reading assessments for the placement of learners in reading classes. It also posits that participation in STAR professional development will lead to changes in instruction.³ The model hypothesizes that the differences in classroom instruction, due to STAR, will lead to differences in reading outcomes, and that the differences in classroom instruction may interact with learner characteristics such that learners with particular demographic or personal characteristics, or that have higher or lower class attendance, may benefit in varying degrees from receiving instruction from high implementers of STAR. The model suggests three categories of research questions. These are: questions about differences in reading outcomes; questions about differences in instruction; and questions about differences in ABE program operations.

¹ The Study of Effective ABE Programs and Practices for First-Level Learners was conducted by Abt Associates Inc. and funded by the U.S. Department of Education/OVAE, and Office of Policy and Program Studies Services during 1995-2003. This was a descriptive study that investigated a range of reading instructional strategies for adult learners.

² Building a Knowledge Base for Adult Decoding is being conducted by the University of Delaware and Abt Associates Inc. under a grant from NIH/NICHD in partnership with OVAE and the National Institute for Literacy. This is an experimental study that is investigating the impact of the use of a decoding curriculum for intermediate-level adult learners that was developed by the University of Delaware.

³ The STAR website provides overview information about the STAR's key component and the training and technical assistance that are provided as part of STAR.

Exhibit 1 Model of Change



Study Questions. The SARIP study's primary research questions address learners' reading outcomes. These questions are the following.

1. What size reading gains do learners who are taught by high-implementing, STAR-trained instructors achieve?
 - a. What is the relationship between learners' background characteristics and their reading gains?
 - b. What is the relationship between learners' attendance and their reading gains?

2. Do the reading outcomes (i.e., alphabetic, fluency, vocabulary, and comprehension) of intermediate ABE learners who are taught by instructors that are high implementers of STAR differ from the reading outcomes of intermediate ABE learners who have participated in ABE reading classes taught by non-STAR trained instructors?

While the primary focus of the study is on learner outcomes, the contractor also will investigate differences between the types of reading instruction provided by the instructors in the treatment and comparison groups. This information will be useful in understanding the reasons for any differences in learner reading outcomes between the treatment and comparison groups. The study questions that address reading instruction are:

3. Is the instructional content provided in the reading classes taught by high-implementing STAR-trained instructors different from the content of the reading classes taught by non-STAR trained instructors?

4. Are the reading instructional strategies used by high-implementing, STAR-trained instructors different from the strategies used by non-STAR trained instructors?

A third topic that will be addressed concerns the operational characteristics of the ABE programs in the treatment and comparison groups. Since the STAR training includes guidance about selected operational characteristics of ABE programs, such as the use of diagnostic reading assessments to identify strengths and weaknesses in learners' reading skills that can guide targeted reading instruction, the study will examine the key characteristics of ABE program operations to understand any differences between the operation of the ABE programs in the treatment and comparison groups. This information may be helpful in understanding whether there are any program-level factors that affect differences in reading outcomes between the treatment and comparison groups. The study question regarding ABE program operations is the following:

5. Does the operation of ABE programs differ according to whether the program had instructors that are high implementers of STAR, or had only non-STAR trained instructors? Specifically, does the use of diagnostic reading assessments for organizing reading instruction for learners differ between the treatment and comparison programs?

Data Collection. The data that will be collected for the SARIP study are described in the response to Item A.2 below.

Timeline for Data Collection. The timeline for data collection is the following:

Learner Data Collection:

- a) Pre-reading assessments and background interview: Beginning of winter term 2009
- b) Attendance data from ABE program records: April 2009; June 2009
- c) Post-test assessments and background interview: End of spring term 2009

Instructor/Class Data Collection:

- a) Instructor observation and interview: winter/spring 2009
- b) Instructors' submission of instructor logs: bi-weekly during winter and spring terms 2009.

ABE Program Data Collection:

Interviews with program director and key program staff: winter/spring 2009

Legislative Authorization. This study is authorized under the Adult Education and Family Literacy Act, Title II of Public Law 105-220, Section 243, National Leadership Activities. Section 243 allows the Secretary of Education to establish and carry out a program of national leadership activities to enhance the quality of adult education and literacy programs nationwide. Appendix A contains Section 243, National Leadership Activities.

A.2. Indicate how, by whom, and for what purpose the information is to be used. Except for a new collection, indicate the actual use the agency has made of the information received from the current collection.

Described in this response are the types of information that will be collected in the SARIP study, and the ways in which, and by whom, the information will be used.

Information to be Collected in SARIP Study. In the SARIP study, data will be collected from treatment learners, instructors, and ABE programs. The contractor will train data collectors from the local communities in which the study’s ABE programs are located and will collect the learner data in the study. The contractor also will implement the quality control procedures for the learner data collection. The measures and instruments are described below.

Learner Data Collection. Learner data will be collected to address Study Questions 1, 1a, 1b, and 2. Three types of data will be collected about treatment learners: (a) reading skills using pre and post versions of standardized reading tests, (b) demographic and background information using a standardized interview protocol (pre and post versions) that was developed by the contractor (Abt Associates) and was used in the contractor’s two prior adult reading studies and has been well tested (see Appendices B and C, Amended Learner Background Interview, Forms A and B), and (c) learner attendance from ABE program files. Presented in Exhibit 2 are the measures, instruments, and data collection methods for the SARIP study’s learner data collection.

Exhibit 2 Learner Measures, Instruments, and Data Collection Methods		
Reading Measure	Instrument	Data Collection
Word recognition	Woodcock-Johnson-R: Letter-Word Identification	I (Individual Testing)
Word analysis	Woodcock-Johnson-R: Word Attack	I
Word recognition	WRAT-3: Word Reading	I
Fluency/word recognition	TOWRE: Sight Word Efficiency	I
Fluency/word analysis	TOWRE: Phonemic Decoding	I
Fluency	NAAL Passage Reading	I
Vocabulary	Nelson Reading: Word Meaning	G (Group Testing)
Reading comprehension	Nelson Reading: Reading Comprehension	G
Reading comprehension	Woodcock-Johnson-R: Passage Reading	I
Other Measures		
Learner’s background (demographics, education, employment, health, perceptions of instruction, skills learned)	Learner Background Interview: Form A (baseline) Learner Background Interview: Form B (post-test)	I
Learner Class Attendance	Program Records	Record Review

Instructor Data Collection. Data on teachers and their class instruction will be collected to address Study Questions 3 and 4. These data are: (a) background characteristics of instructors,

(b) documentation of instructors’ reading teaching activities in the study’s target class, and (c) information about instructors’ use of the STAR methods and materials. The forms that will be used to collect these data were developed by the contractor (Abt Associates) and were used in the contractor’s previous reading studies, and have been well tested. Sections of the Instructor Interview form and the Instructor Log have been slightly modified with information related to the SARIP study. Presented in Exhibit 3 are the measures and sources of data for the instructor data collection. The forms are found in Appendix D (Instructor Background Characteristics), Appendix E (Class Observation Form), Appendix F (Instructor Interview), and Appendix G (Instructor Log).

Exhibit 3 Instructor and Program Measures, Instruments, and Data Collection Methods		
Measure	Instrument	Data Collection
Instructor Characteristics (demographics, experience in adult education, reading training)	Instructor Background Characteristics Form	Interview
Instructional Approach in Reading Class (reading content, instructional methods, and stages of a lesson)	Class Observation Form	Direct Observation
Instructional Approach in Reading Class (reading content, instructional methods, and stages of a lesson)	Instructor Log	Form Completion— Instructor Self-Report
Instructional Approach in Reading Class (use of diagnostic instruments, lesson planning, integration of STAR into teaching, reading content, and instructional methods)	Instructor Interview	Interview
Program Operations <ul style="list-style-type: none"> ▪ Program management ▪ Program improvement ▪ Learner recruitment and orientation ▪ Learner intake, diagnostic assessment ▪ Learner pre-post assessment ▪ Support Services ▪ Learner Transition 	ABE Program Protocol	Interview

ABE Program Data Collection. Information about the operational characteristics of the ABE programs in the study's treatment group will be collected during the pre-test period for the study. The contractor will use an ABE Program Protocol in conducting face-to-face interviews with the program's director and two key program staff. This protocol was used in Abt Associates' previous two reading studies and has been well tested. The constructs measured in the instrument are listed in Exhibit 3; the ABE Program Protocol is contained in Appendix H.

Uses of Information. The U.S. Department of Education will use the data collected in the SARIP study to assess the preliminary learner reading outcomes from the STAR intervention and to determine whether a more rigorous evaluation (e.g., a randomized controlled trial) of STAR should be undertaken at this point in the implementation of STAR. The data collected in the SARIP study about the delivery of instruction by teachers trained in STAR will be used by the U.S. Department of Education to review the STAR training and to determine whether modifications may be needed in the STAR training. The information about ABE programs collected in the study will be used by the U.S. Department of Education and by state adult education offices to provide guidance to local ABE providers about the types of ABE program practices that may support the delivery of effective reading instruction.

Prior to the SARIP study, no reading outcome data have been collected from adult learners who have participated in reading instruction conducted by STAR-trained reading instructors. There is a growing interest among states to participate in the STAR training, and learner outcome data are needed to demonstrate that the STAR training and materials are an effective approach to developing the reading skills of intermediate-level adult learners. There also is a need to identify the characteristics of high-implementing STAR-trained instructors so that state adult education offices and local ABE program providers interested in using STAR can have information about the types of instructor skills and knowledge that are associated with instructors whose learners are able to develop their reading skills. Furthermore, there is a need to know whether there are critical aspects of the operation of an ABE program, such as the administration and use of diagnostic reading assessments, which can help to improve the reading outcomes achieved by adult learners. The SARIP study will provide data to address these needs that otherwise will not be available.

A.3. Describe whether, and to what extent, the collection of information involves the use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses, and the basis for the decision for adopting this means of collection. Also describe any consideration of using information technology to reduce burden.

Audio-taping will be used in the collection of two types of data in the SARIP study. Learners' responses to the seven individualized reading assessments used in the study will be audio-taped so that the responses can be scored by a second researcher as part of the inter-rater reliability procedures for the study. The contractor's (Abt Associates) face-to-face interviews with instructors also will be audio-taped to help ensure the accuracy of the information that is documented. The individual background interviews with learners will be recorded onto paper forms, since the majority of the questions in the interview have pre-coded responses on the form

and this process is the most efficient method to use in collecting these data. The contractor's observations of STAR instructors' teaching also will be recorded manually, since the study's budget does not allow for videotaping of instructors' classes.

A.4. Describe efforts to identify duplication. Show specifically why any similar information already available cannot be used or modified for use of the purposes described in Item 2 above.

This is the first time that objective data on the instructional and program implementation and learner outcomes from the use of STAR will be collected, so the new data collection for SARIP is not a duplication of any previously collected data.

A.5. If the collection of information impacts small businesses or other small entities (Item 5 of OMB Form 83-I), describe any methods to minimize burden.

The primary entities for this study are adult basic education programs located in community colleges and school districts. Burden is minimized for all respondents by requesting only the minimum information required to achieve the study's objectives. All data collection will be coordinated by the contractor. The SARIP study's contractors will carefully specify information needs; the data that will be collected from the programs will be restricted to generally available information maintained in programs' administrative records. It is anticipated that adult basic education program personnel will be able to retrieve and transfer data with minimal burden and with support from the study's contractors.

A.6. Describe any consequences to Federal program or policy activities if the collection is not conducted or is conducted less frequently, as well as any technical or legal obstacles to reducing burden.

If the proposed data were not collected, OVAE would be unable to evaluate the STAR program. The dissemination of STAR has been part of OVAE's Improving Program Performance (IPP) initiative to promote high quality reading instruction in adult education classrooms throughout the United States, and it is critical to collect information about the effectiveness of STAR and the IPP initiative. OVAE is involved in ongoing dissemination activities related to STAR, and the collection of data at this point in time is essential to assist OVAE in determining subsequent investments in STAR.

The schedule for the study is pre-post data collection over a five-month period. This is the minimum frequency of data collection that can be undertaken to address the study's questions. Since the primary study questions are investigating learners' development of reading skills, two data collection points are necessary to assess growth, and two terms of instruction (approximately five months) is the minimum amount of time between pre- and post- assessment in order to obtain reliable outcome data.

A.7. Explain any special circumstances of information collection.

This request fully complies with the following regulations. Information collection will NOT be conducted in a manner:

- requiring respondents to report information to the agency more often than quarterly;
- requiring respondents to prepare a written response to a collection of information in fewer than 30 days after receipt of it;
- requiring respondents to submit more than an original and two copies of any document;
- requiring respondents to retain records, other than health, medical, government contract, grant-in-aid, or tax records, for more than three years;
- in connection with a statistical survey, that is not designed to produce valid and reliable results that can be generalized to the universe of study;
- requiring the use of statistical classification that has not been reviewed and approved by OMB; or
- requiring respondents to submit proprietary trade secrets, or other confidential information, unless the agency can demonstrate that it has instituted procedures to protect the information's confidentiality to the extent permitted by law.

A.8. If applicable, provide a copy and identify the date and page number of publication in the Federal Register of the agency's notice, required by 5 CFR 1320.8(d), soliciting comments on the information collection prior to submission to OMB. Summarize public comments received in response to that notice and describe actions taken by the agency in response to these comments. Specifically address comments received on cost and hour burden.

The Federal Register notices were published on May 14, 2008, Vol. 73, No. 94, page 27811-27812 and on July 18, 2008, Vol. 73, No. 139, page 41346.

A.9. Explain any decision to provide any payment or gift to respondents, other than remuneration of contractors or grantees.

As an incentive to participate, learners participating in the study will be reimbursed for their transportation costs to travel to the adult basic education program to meet with the data collector and participate in the reading assessment and interview. Based on the contractor's (Abt Associates) prior studies, this cost is estimated at \$20 per learner for each of the two data collection meetings (pre- and post data collection).

A.10. Describe any assurance of confidentiality provided to respondents and the basis for the assurance in statute, regulation, or agency policy.

Assurances of Confidentiality. In the SARIP study, assurances of confidentiality will be provided to the participating program and to each participating learner. Prior to an ABE program's participation in the study, the program's director will receive a letter from ED describing the study (see Appendix K). After a program has decided to participate in the SARIP study, Abt Associates will send a letter to the program's director that describes the activities that the program will participate in as part of the study (see Appendix I). One of the stipulations in the letter is that "responses are protected under the Privacy Act, and learners will be asked to complete a Learner Consent Form prior to participating in data collection." Each individual learner who volunteers to participate in the SARIP study will meet with the study's data collector prior to the beginning of data collection to read and sign the "Learner Consent for Participation in Research" form (see Appendix J).

Basis for Assurance. The contractor will follow procedures for ensuring and maintaining data privacy, consistent with the Privacy Act of 1974 (P.L. 93-579, 5USC 552a), the Freedom of Information Act (5 USC 522), and related regulations, including but not limited to: 41 CFR Part 1-1 and 45 CFR Part 5b, and the Federal common rule or ED final regulations on protection of human research subjects.

Data to be collected will not be released with individual student or program identifiers. Data will be presented in aggregate statistical form only. All contractor staff involved in collecting, reviewing, or analyzing individual-level data will be knowledgeable about data security procedures and will be prepared to describe them in full detail to respondents. Respondents will be assured that all information identifying them or their adult basic education program will be kept private to the extent allowed by law. The privacy procedures adopted for this study during all rounds of data collection, data processing, and analysis include the following:

- All study respondents will be assured that rules will be followed to protect their privacy. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific program or individual. Information will not be provided that identifies programs or individuals to anyone outside the study team, except as required by law.
- Hard-copy data collection forms will be delivered to a locked area for receipt and processing. The contractor will maintain restricted access to all data preparation areas (i.e., receipt, coding, and data entry). All data files on multi-user systems will be under the control of a database manager, with access limited to project staff on a "need-to-know" basis only.
- Individual identifying information will be maintained separately from completed data collection forms and from computerized data files used for analysis. No respondent identifiers will be contained in public use files made available from the study, and no data will be released in a form that identifies individual program staff, program participants, or comparison group members.

A.11. Provide additional justification for any questions of a sensitive nature, including matters that are commonly considered private.

There are no questions of a highly sensitive nature are included in the learner questionnaire. Learners will be asked to provide only demographic (ethnicity, race, and age), income, educational, and work experience information. Such items may be sensitive to some respondents, but they are types of questions that learners are asked when they enroll in ABE programs, and they are important as variables that may be related to learners' reading gains.

The questions are worded in a sensitive, nonjudgmental manner and have been successfully pre-tested and used with over 1,500 adult learners in the contractor's (Abt Associates) previous two reading studies with no evidence of harm. Furthermore, survey responses will be confidential, as described above, and responses will not affect learners' status in the program.

A.12. Provide estimates of the hour burden of the collection of information.

Exhibit 4 presents the estimated burden for the SARIP project. A total of 4,734 responses are expected annually and for the whole project. Data collection will occur during one year of the study. For learner respondents, there will be two rounds of data collection. For instructor respondents, there will be one round of data collection for two instruments, two rounds for one instrument, and 15 rounds for the fourth instrument. For program staff respondents, there will be one round of data collection. The total response time for data collection is 1,431.42 hours.

No estimated costs are presented for instructor respondents, since they will be paid by the study for their time collecting data (retrieving attendance data from program files, completing 15 instructor logs, completing a short background survey, and participating in a post-observation interview). The cost for the instructors is included in the estimate for Item 14.

The estimated hourly cost for program staff is based on an average estimate of the hourly rate of ABE program directors (\$40/hr.) and ABE program staff (\$30/hr). These data are the contractor's (Abt Associates) estimates based on their recent adult reading study conducted with the University of Delaware, in which they worked with 24 ABE programs in 12 states.

**Exhibit 4
Respondent Burden Estimates**

Informant/ Instrument	Number of Respon- dents	Number of Rounds	Total Number of Responses	Average Time per Response (Hours)	Total Respondent Time (Hours)	Estimated Hourly Cost to Respondent (Dollars)	Estimated Total Cost (Dollars)
Learners	232⁴		4,180⁵		1,004		\$0
W-J-R: Letter Word Id., Word Attack, Passage Reading	232	2	418	.60	251	\$0	\$0
WRAT3: Word Reading	232	2	418	.16	67	\$0	\$0
TOWRE: Sight Word Efficiency, Phonemic Decoding	232	2	418	.32	134	\$0	\$0
NAAL Passage Reading	232	2	418	.16	67	\$0	\$0
Nelson: Word Meaning, Reading Com.	232	2	418	.66	276	\$0	\$0
Learner Background Interview	232	2	418	.50	209	\$0	\$0
Instructors	26		494		277.42		\$0
Background Characteristics	26	1	26	.17	4.42	\$0	\$0
Post-observation Interview	26	1	26	1	26	\$0	\$0
Instructor Log	26	15	390	.50	195	\$0	\$0
Attendance	26	2	52	1	52	\$0	\$0
Program Staff	60		60		45		\$1,575
Interview	60	1	60	.75	45	\$35.00	\$1,575
TOTAL	318		4,734		1,326.42		\$1,575

⁴ The number of learner respondents is 232, which is the total number of learners that the study expects to have for the pre-test phase of data collection. It is expected that 80% of the number of learners at pre-test will be available for the post-test phase.

⁵ The number of total responses for learners for each instrument is calculated based on the 232 learners at pre-test and the expectation that 80% (186) will be available at post-test, for a total of 418 learners.

A.13. Provide an estimate for the total annual cost burden to respondents or record-keepers resulting from the collection of information.

There are no direct costs to respondents.

A.14. Provide estimates of annualized costs to the Federal government.

The estimated total cost to the Federal government for conducting the Study on Strengthening Adult Reading Instructional Practices (SARIP)—including the contractors’ design of the study, identification of ABE programs and instructors to participate in the study, collection of program/instructor/learner data, processing and analyzing the data, conducting briefings to OVAE, and preparing reports summarizing the results—is \$863,223 (Year 1: \$251, 774; Year 2: \$611,449).

The cost of the data collection activities associated with this project is projected to be \$233,993. The study period is from September 2007 to September 2009, with data collection taking place from January 2009 to June 2009. Provided in Exhibit 5 is the budget for the data collection costs for the contractors JBL Associates, Inc. and Abt Associates Inc.

Exhibit 5		
Data Collection Costs (Year 2 of study)		
Price Category	Hours	Cost
JBLA Labor (Including fringe & overhead)	192	\$15,726
Other Direct Costs:		
Travel		\$36,459
Telephone		\$60
Postage		\$630
Subtotal Other Direct:		\$37,149
Subcontractor: Abt Associates		\$96,204
Consultants:		\$62,236
Total Direct Costs:		\$211,315
General & Administrative (Dir. Costs, excluding Abt Associates, x 14%)		\$16,116
JBLA Fee (Dir. Costs + G&A x 5%, excluding Abt Associates)		\$6,561
Total Cost:		\$233,992

A.15. Explain the reasons for any program changes or adjustments reported in Item 16 of IC Data Part 1.

This is a new study.

A.16. For collections of information whose results will be published, outline plans for tabulation and publication.

Schedule. The schedule of the project’s activities and deliverables is presented in Exhibit 6.

Exhibit 6 Project Activities and Deliverables		
Task	Deliverable/Activity	Due Date 9/25/07 start date
1. Conduct an Initial Planning Meeting with OVAE	1a. Agenda 1b. Draft project work plan 1c. Meeting at OVAE 1d. Meeting summary	By 10/9/07 By 10/16/07 By 10/16/07 By 10/23/07 (1 week after meeting)
2. Establish and Convene a Technical Working Group (TWG)	2a. List of proposed TWG members to the Department for approval 2b. Final roster of TWG membership 2c. Draft agenda 2d. Disseminate briefing materials to TWG 2e. First TWG meeting 2f. Second & Third TWG meetings 2g. Summary of TWG meeting minutes	By 10/25/07 By 11/15/07 By 11/22/07 By 12/4/07 By 1/28/08 By 9/19/08; 1/21/09 By 12/18/07 1 week after meeting
3. Develop and Implement a Data Collection Plan	3a. Draft data collection plan 3b. Final collection plan 3c. Conduct data collection 3d. Complete collection activities	By 1/10/08 By 1/24/08 8/08-4/09 By 6/25/09
4. Maintain Quality Control	4a. Interim report (include preliminary tabulations, tables, cross tabulations)	By 2/05/09
5. Ensure accountability through reporting	5a. Two oral presentations with written reports 5b. Monthly project reports 5c. Annual reports 5d. Draft technical report 5e. Final technical report	By 8/16/08; 8/15/09 By the tenth of each month for the period of the contract. By 9/26/08; 9/25/09 By 8/11/09 By 9/25/09

Study Reports. There are two analytic reports planned for the study. The first is an interim report that will provide descriptive data about the pre-test sample of learners participating in the study. These data will include: (a) demographic characteristics of learners, and (b) pre-test scores on nine reading instruments. The second report is the study’s technical report, which will include the results of all of the analyses conducted in the study. The following is the draft outline for the final technical report.

1. Overview of the study
 - 1.1. Context for Study
 - 1.2. Overview of the Report
2. Study Design and Methods
 - 2.1. Research Questions
 - 2.2. Design
 - 2.3. Methods
 - 2.4. Study Constructs, Measures, and Data Collection
 - 2.5. Analysis Methods for Five Study Questions
3. Results: Learners Participating in STAR Classes
 - 3.1. Description of Learners
 - 3.2. Reading Gains Achieved by STAR Learners
 - 3.3. Relationship of Learners' Background Characteristics to their Reading Outcomes
 - 3.4. Relationship of Learners' Attendance to their Reading Outcomes
4. Description of Comparison Group Learners
5. Results: Comparison of Learners Taught by High-implementing STAR Instructors to Learners Taught by Non-STAR Trained Instructors
6. Description of Instructors and Instruction Provided to Treatment and Comparison
 - 6.1. Characteristics of Instructors in Treatment and Comparison Groups
 - 6.2. Comparison of Reading Content Taught by Treatment and Comparison Instructors
 - 6.3. Comparison of Instructional Strategies Used by Treatment and Comparison Instructors
 - 6.4. Extent of Differences Between Instructors in Treatment and Comparison Groups
7. Description of ABE Programs in Treatment and Comparison Groups
 - 7.1. Characteristics of Program Operations in Each Group
 - 7.2. Extent of Differences Between Programs in Treatment and Comparison Groups
8. Study Conclusions

Data Tabulation/Analysis. The study's estimation procedures are described for Study Questions 1-4, which will involve statistical analyses.

Study Question 1: Describe the Distribution of Gain Scores for Treatment Students. To address Study Question 1, the contractor will analyze the gain score data in conjunction with data on learner attendance and learner background characteristics. These analyses will describe the amount of change in learners' reading skills from pre-test to post-test in each of the two groups. The data for the basic change analyses will be change scores obtained from reading test data

collected before and after the instructional treatments. The description of the distribution of change scores will include calculation of statistics and production of plots. The statistics will include the mean, standard error, minimum, maximum, median, and the 5th, 25th, 75th and 95th percentiles. Note that if the study were to find, for example, that the mean, median, 25th and 75th percentiles were all greater than zero, this would be strong evidence that students are learning, and would provide a good description of how much they were learning. The study will examine these distributional statistics for both groups of learners: those in classes with high-implementing STAR instructors, and those in non-STAR classes.

The study also will test the statistical significance in learners’ change scores. Paired t-tests will be conducted. If the average change is positive and the null hypothesis is rejected, this will be evidence that program participation has shifted the distribution of reading skills among the population of participants.

In addition, the study will examine the effect size of the pre-post changes in learners’ scores on all of the reading measures. The effect size is a measure of the mean change expressed in standard deviation units. An advantage of expressing mean change as an effect size is that the mean changes for each of the study’s standardized reading tests can be directly compared to one another, because they are each expressed in a common metric. Further, a commonly used rule-of-thumb for interpreting the magnitude of effect sizes allows one to discuss the effect sizes of 0.20 or lower as representing a “small” effect, an effect size of about 0.50 as “medium,” and an effect size of 0.80 or greater as representing a “large” effect (Cohen, 1988). An example table shell for displaying gains expressed as effect sizes with 95% confidence intervals is shown in Exhibit 7.

Exhibit 7
Test Score Gains for Learners of High-implementing STAR-trained Instructors
(Study Question 1)

Subtest	STAR Mean Gain (95% CI)
WJ-R Word Attack	
WJ-R Letter Word ID	
WRAT3 Word Reading	
Towre Sight Word Efficiency	
Towre Phonemic Decoding	
Nelson Word Meaning	
Nelson Reading Comp	
WJ-R Passage Reading	
Sample Size	n =

Study Questions 1a and 1b: Describe the Relationship Between Learner Gains and (a) Learners’ Background Characteristics and (b) Learners’ Attendance. To address Study

Questions 1a and 1b, the study will regress measures of test score gains on learners’ demographics and learners’ attendance in ABE training with high-implementing STAR-trained instructors. These models will control for learners’ pre-test reading scores, and will be modeled in a hierarchical linear modeling framework that will account for the clustering of learners within classes and programs. These models will be similar in form to the models that will be used to address Study Question 2, described below, but will be fit to a subset of data comprised only of the sample of learners with high-implementing STAR instructors, and will consequently have no term corresponding to the treatment/comparison group indicator. An example table shell for display of results of models that will be used to address Study Question 1a is shown in Exhibit 8. Model results corresponding to Study Question 1b will be displayed in a similar manner.

Exhibit 8
Test Score Gains as a Function of Learner Characteristics
(Study Question 1a)

	Regression Coefficient	Standard Error	P-value
Pre-test score			
Non-native born			
Prior ABE			
Learning Problem			
Current Employment			

Sample Size

Study Question 2: Estimating the Impacts of High-Implementing STAR-Trained Instructors on Learners’ Reading Skills. To address Study Question 2, the study will estimate models to measure the regression-adjusted differences in learners’ post-test scores between high-implementing STAR-trained instructors and the matched comparison group. As described and justified in Part B of the supporting statement, the estimates suggest that the study design will be sufficient to detect impacts on reading skills of .37 standard deviations, or an effect size of .37.

To estimate these differences, the study will use two-level or three-level hierarchical linear models (HLMs) where individual learners (level-1) are nested in classes (level-2) and classes are nested in programs (level-3). Since it is likely that there will be insufficient variation at level-3 to support the inclusion of a third level in the model, the analytical model will be specified as two-level HLM.

If there is sufficient variance at level-3 to support a third level, the analytical model will be a straightforward generalization of the model shown below⁶. Models will be of the form:

$$Y_{\Delta,ij} = (Y_{Post,ij} - Y_{Pre,ij}) = \beta_0 + \alpha_{0j} + \beta_1(Trt_j) + \beta_2(Y_{Pre,ij}) + \sum_{m=1}^M \beta_{2+m}(X_{m,ij}) + \varepsilon_{ij}, \text{ [Eqn 1]}$$

where:

- $Y_{\Delta,ij}$ is a pre-post change score on a reading assessment (e.g., WRAT Reading assessment) for the i^{th} student in the j^{th} class;
- $Y_{Post,ij}$ is a post-treatment score on a reading assessment;
- $Y_{Pre,ij}$ is a pre-treatment score on a reading assessment;
- Trt_j =1 if class j is a treatment class, =0 if comparison class;
- $X_{m,ij}$ is the m^{th} of up to M covariates measured at pre-treatment for the i^{th} student in the j^{th} class;
- β_0 is the grand mean intercept value;
- α_{0j} is a random intercept term for the j^{th} class, assumed to be normally distributed with mean = 0 and variance = τ^2 ;
- β_1 is the treatment effect, which is equal to the mean difference between treatment and comparison group in change scores, conditional on (controlling for) pre-treatment score and all other model covariates;
- β_2 is the effect of the pre-test score on the change score;
- β_{2+m} is the effect of the m^{th} covariate on the change score;
- ε_{ij} is the residual for the i^{th} learner in the j^{th} class, assumed to be normally distributed with mean = 0 and variance = σ^2 .

We note that this model is equivalent to the following model:

$$Y_{Post,ij} = \beta_0 + \alpha_{0j} + \beta_1(Trt_j) + \beta_2^*(Y_{Pre,ij}) + \sum_{m=1}^M \beta_{2+m}(X_{m,ij}) + \varepsilon_{ij} \quad \text{[Eqn 2]}$$

where β_2^* in Equation 2 is equal to $\beta_2 + 1$ in Equation 1. All other terms in the model result in estimates and standard errors that are identical (Allison, 1990). Thus, the treatment effect and its standard error are identical in these two models, but the model specified in Equation 1 is more convenient because it expresses the outcome measure in a metric that is of substantive interest – the change in reading ability between pre-treatment and post-treatment.

Potential model covariates will include learner demographic characteristics (including age, gender, and an indicator for whether the learner was born and educated outside of the United States),

⁶ An additional level of clustering is classes within instructors. Although there will be some instructors that teach more than one ABE class within a year, the planned analysis models will ignore the clustering of classes within instructors. This is because in the contractor's (Abt Associates) experience, data sets of the size proposed for the current study cannot support the partitioning of variance into the four levels that would be required to account for the clustering of classes within instructors. There is usually a greater correlation among measures from learners within a class, than there is among learners in different classes, but who share the same instructor. Thus, collapsing those two levels of clustering into one usually works well. The study will attempt to model the variation among classes within programs, but that level of clustering may also not be estimable for most outcomes, and may therefore have to be dropped from most or all of the models.

disabilities, health, and general functioning, prior participation in adult basic education, and measures of goals and expectations. All potential model covariates will be exogenous variables created from items measured at pre-treatment. Covariates satisfying a $p < 0.20$ criterion will be retained in the final analysis models. This criterion is a reliable indicator for whether a covariate either serves to control for confounding, or helps reduce residual variation, which will decrease the standard error of the treatment variable (Budtz-Jorgensen et al, 2007; Maldonado & Greenland, 1993).

Estimated treatment effects will be converted to standardized effect sizes by dividing the treatment effect estimate by the pre-treatment pooled standard deviation of the treatment and comparison groups.

The results from the analysis will be presented in tables with different rows for each outcome measure. A draft table to present results from the analysis designed to answer Study Question 2 is presented in Exhibit 9.

Exhibit 9
Difference in Test Score Gains between High-Implementing STAR-Trained Instructors and Non-STAR Trained Instructors (Study Question 2)

	High Implementing STAR Mean Gain ^a	Non-STAR Mean Gain ^b	Difference ^c	P-value
WJ-R Word Attack				
WJ-R Letter Word ID				
WRAT3 Word Reading				
Towre Sight Word Efficiency				
Towre Phonemic Decoding				
Nelson Word Meaning				
Nelson Reading Comp				
WJ-R Passage Reading				
Sample Size				

^a Unadjusted Mean Gain for learners taught by high-implementing STAR-trained instructors.
^b Mean gain for non-STAR learners, adjusted for differences in the mean characteristics included in the models shown in Exhibit 5.
^c Difference between mean gain of STAR and non-STAR learners, adjusted for differences in the mean characteristics included in the models shown in Exhibit 5.

Study Questions 3 and 4: Estimating the Differences between High-Implementing, STAR-trained Instructors and Non-STAR trained Instructors in (a) Instructional Content and (b) Instructional Strategies. To address Study Questions 3 and 4, the study will conduct analyses of the instructor-level data. Before addressing these questions directly, the study will produce descriptive statistics (means, frequency distributions) on the characteristics of instructors in the treatment and comparison groups. Instructor characteristics include gender, number of years of teaching adult basic education, highest degree completed, academic area of specialty, type of reading training completed, and use of formal lesson plans.

To address Question 3, the study will estimate the differences in instructional content provided by high-implementing STAR and non-STAR instructors using two-level hierarchical linear

models, where instructors (level-1) are nested in programs (level-2). These models will have measures of instructional content as outcome (dependent) measures, and an indicator for STAR vs. non-STAR on the right-hand side of the model (independent variable).

To address Question 4, the study will estimate differences in instructional strategies used by high-implementing STAR and non-STAR instructors using two-level hierarchical linear models, where instructors (level-1) are nested in programs (level-2). These models will have measures of instructional strategies as outcome (dependent) measures, and an indicator for STAR vs. non-STAR on the right-hand side of the model (independent variable).

A.17. If seeking approval to not display the expiration date for OMB approval of the information collection, explain the reasons that display would be inappropriate.

The expiration date for OMB approval will be displayed.

A.18. Explain each exception to the certification statement identified in the “Certification for Paperwork Reduction Act Submissions.”

No exceptions to the certification statement are requested or required.

References

- Allison, P.D. (1990). Change scores as dependent variables in regression analysis, In *Sociological Methodology*, C. Clogg (ed.). Oxford: Basil Blackwell, 93-114.
- Budtz-Jorgensen, E., Keiding, N., Grandjean, E.M., & Weihe, P. (2007). Confounder selection in environmental epidemiology: Assessment of health effects of prenatal mercury exposure. *Annals of Epidemiology*, 17, 27-35.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edition. Hillsdale, NJ: Lawrence Erlbaum.
- Kruidenier, J. (2002). *Research-based principles for adult basic education: Reading instruction*. Washington, DC: National Institute for Literacy.
- Maldonado, G., & Greenland, S., (1993). Simulation study of confounder-selection strategies. *American Journal of Epidemiology*, 138(11), 923-936.
- Student Achievement in Reading (STAR) Web Site. <http://www.startoolkit.org/>