

Annual Mandatory Collection of Elementary and Secondary Education Data through *EDFacts*¹

This collection is the necessary first step in transforming how the Department collects, uses, and reports on a large volume of education information. *EDFacts* mark a decidedly different approach to how the Department collaborates and interacts with institutional data providers and users at the elementary and secondary education levels. It is justified by the Department's mission to "...ensure equal access to education and to promote educational excellence throughout the nation."

Once approved and functional, *EDFacts* will provide to the Department and its constituents more comparable, timely, valid, and accurate information to better meet our mission. The system provides a framework for assessing multiple information collections and the network provides information for better assessing multiple education program initiatives, as well as the condition of American education. It is paramount to highlight the fact that this transformation not only will result in better compliance with the Paperwork Reduction Act by limiting the 'quantity' of information collections; equally important is the fact that the 'quality' of information will increase.

A. Justification

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.

The Department of Education (ED) currently administers a budget of about \$88.9 billion per year -- \$57.6 billion in discretionary appropriations and \$31.3 billion in mandatory appropriations. ED operates programs that touch on every area and level of education. The Department's elementary and secondary programs annually serve more than 14,600 school districts and approximately 54 million students attending more than 94,000 public schools and 27,000 private schools.

¹ The Department of Education has changed the name of the Education Data Exchange Network (EDEN) to *EDFacts* in recognition of the system's expanded capabilities, as explained in this Supporting Statement and the notice of proposed rulemaking for State-administered Programs published in the *Federal Register* on April 27, 2006 (71 FR 24824). Except as needed to describe the historical context of the database, submission and query capabilities of the system, this Attachment uses "*EDFacts*" to refer to the expanded system.

Since its inception, the collection, use, and reporting of education information and data has served as an integral component of the ED mission. Previously, many information collections were developed unilaterally responding to individual program and office determinations. The concept for the Education Data Exchange Network (EDEN), which has been renamed *EDFacts*, marks a transformation in how ED collects, uses, and reports on a large volume of education information.

The Secretary has determined that complete, accurate, and reliable data are essential for effective decision-making and for implementing the requirements of the Nation's education laws. The Department's ability to collect, store, and manage education data efficiently through electronic means allows for easier submission by States and reduces duplication of collections and burdens on states. It also facilitates the efficient use of data for analysis by program officials and other interested parties. Implementation of the No Child Left Behind Act of 2001 (P.L. 107-110) (NCLB) requires that educators have accurate and reliable data and the knowledge needed to assess how close the Nation is to reaching the goal of ensuring that every child achieve to high academic standards.

To date, submission of data through the *EDFacts* system has been voluntary and states have been required to continue reporting data through dozens of existing data collections required by Congress and administered by the Department. This situation has created competing databases with inconsistent information about the status and progress of education throughout the Nation.

The Department has designed this collection to obtain the most commonly collected data elements so that states need only report these data once, through a centralized, electronic process. The Department's goal in requiring electronic submission of information is to reduce each state's reporting burden significantly and to streamline dozens of data collections currently required by the Department.

Nearly every state has submitted some portion of the data electronically that it will eventually be required to submit to the Department electronically through *EDFacts*. But *EDFacts* will only reach its full potential in reducing state reporting burden and increasing the ability of the Department and states to analyze and improve student achievement if all states provide their data through the system.

Beginning with the 2006-2007 *EDFacts* data sets in this collection, the Department is seeking OMB approval, pursuant to the *PRA*, of this information collection request including the mandatory requirement that States submit electronically through *EDFacts* the program and demographic information that States currently are required to report under separate and overlapping collections. In time, this approval of this data collection request will eliminate the need for states to submit reports under current separate and overlapping collection instruments. The Department will discontinue all existing data collections that seek to collect the data that is required by this information collection.

ED is in the process of collecting data for the 2003-2004, 2004-2005, and 2005-2006 school years as approved last year by OMB (1880-0541). This OMB approval number

was changed in April 2006 to 1875-0240. This proposed collection includes the 2006-2007, 2007-2008, and 2008-2009 school years. In the current *Federal Register* notice inviting public comment on this collection request (Attachment A), ED described the mandatory requirement to supply *EDFacts* data and specifically requested public feedback about the availability of the data in the next school year (2006-2007). To the extent that any one or more of these proposed data groups are not available in the coming school year, ED seeks to know if those data will be available in future years (2007-2008 or 2008-2009). **ED seeks a three-year approval for this collection.** As part of this approval, it needs to be understood that ED is authorized to collect the data about these school years over whatever amount of time is required to secure this data from each state, district, or school.

ED seeks OMB approval under the *PRA* to collect the elementary and secondary education data from schools, school districts, and states as described in Attachments B and C. In Attachment B, ED documents all of the data groups it intends to collect. In Attachment C, ED explains the differences between the data groups ED intends to collect and those data groups it collected during the past months for the 2005-06 school year. Those persons who are familiar with the current EDEN collections may want to start by first reviewing Attachment C.

This ED collection is the fourth year of a multiple year effort to consolidate the collection of education information about states, districts, and schools in a way that improves data quality and reduces paperwork burden for all of the national education partners. The first year (FY 2003) of this initiative collected data from November 2003 into March 2004 as part of a pilot test to establish the availability of elementary and secondary education data in the states and state to federal data transmission standards and guidelines. The data collected in the pilot test was for the 2002-03 school year.

The transformation of the Department's data collection process began with the Performance Based Data Management Initiative, which established EDEN. That process is now complete with the renaming of EDEN as *EDFacts*. This new name is intended to signal the increased capabilities of the system that will permit users to query the system as well as submit data to the system. This includes the *EDFacts* data repository containing performance information about schools and federal education programs. Instead of sending multiple, and sometimes redundant, data collections to ED program offices, states will periodically submit data to the *EDFacts* data repository. ED will negotiate protocols with the states for this periodic collection of data via state data submission plans. ***EDFacts* is implementing a fundamental change in the management of data at the Department** as it moves from a data form-based model to an electronic data element-based model of collecting information from the states.

Civil Rights Data Collection

A number of federal laws support the collection of civil rights data. Title VI of the Civil Rights Act of 1964, prohibits discrimination based on race, color, and national origin. Title IX of the Education Amendments of 1972, prohibits discrimination based on sex.

Section 504 of the Rehabilitation Act of 1973, prohibits discrimination on the basis of disability. The Age Discrimination Act of 1975 prohibits discrimination based on age. Title II of the Americans with Disabilities Act of 1990 prohibits discrimination based on disability. The applicable laws supporting the collection of the civil rights data are found at <http://www.ed.gov/about/offices/list/ocr/know.html>.

The 1980 Department of Education Organization Act conveys to the Assistant Secretary for Civil Rights the authority to “collect or coordinate the collection of data necessary to ensure compliance with civil rights laws under the jurisdiction of the Office for Civil Rights” (20 U.S.C. 3401, Section 3413(c)(1)). Any data collection that the Office for Civil Rights (OCR) has “determine[d] to be necessary” to ascertain compliance with these laws, is mandatory pursuant to the implementing regulations of the Federal civil rights laws. Every two years, for almost forty years, OCR has collected data about, and from, school districts and schools under the Elementary and Secondary Civil Rights Survey as approved by OMB as 1870-0500.

OCR and the EDEN team have collaborated to develop the EDEN Survey Tool. Last year, for the first time the EDEN data collection instrument asked states to submit, on a voluntary basis, the data previously collected through OCR’s Elementary and Secondary Civil Rights Survey (E&S Survey).

The development of the survey tool served as an excellent example of accomplishing multiple objectives identified by OMB for ED, NCLB, and OCR. The EDEN Survey Tool facilitated a full migration of the Civil Rights Survey from a stand alone program centered collection to the EDEN system and reduced duplicative web survey development efforts in ED that were intended to collect essentially the same education information. As this tool is used in future surveys, each of those surveys will be cleared under this approval number through OMB under the PRA.

OCR first collected data through the E&S Survey in 1968. In 2000, OCR conducted a universal survey of school districts and schools in the United States. In 2002, OCR resumed its regular biennial survey of a sample of school districts. In the 2004 edition, ED and OCR collected data from a sample of school districts using the EDEN Survey Tool. This tool will be used again in the 2006 Civil Rights Data Collection under *EDFacts*.

The tool was developed incrementally with increasing functionality and capability. Initially, the tool included such features as e-mail functionality for communication with the survey respondents, multiple options for replying with the data, forms that can structure the responses for control, and functionality that allows selective sample surveys. Its analytical capabilities will be increased as more data elements are included in the *EDFacts* data repository.

Although ED was not able to take full advantage of the ‘pre-population’ of data from EDEN’s core repository in the first (pilot) test of the EDEN Survey Tool, that was the option that the district respondents most widely selected. ED plans to improve the tool’s

usability and usefulness and be much more able to take full advantage of this option in the 2006 Civil Rights Data Collection. The applicable laws supporting the collection of the civil rights data are found at <http://www.ed.gov/about/offices/list/ocr/know.html>.

General Education Provisions Act Data Collection

Another collection that has been absorbed by the current *EDFacts* collection is the General Education Provisions Act (GEPA) 424 Data Collection (previously approved by OMB as 1875-0203) which detailed the distribution of ED funds in support of elementary and secondary education programs. This collection provides data for a biennial report required by section 424 of GEPA. These data have been required for over 20 years. Since this is information also being sought for *EDFacts*, the combination of these collections is appropriate. Because the data collected this past year was for FY2002 and FY 2003 (approximately the 2002-2003 and 2003-2004 school years), it needed to be separate from the electronic submission of EDEN data from the states. Future biennial GEPA reports will use *EDFacts* data without an additional data collection effort or the associated paperwork burden.

Office of Special Education Programs Data Collections [Office of Special Education and Rehabilitative Services (OSERS)]

Up until the 2005-2006 reporting year, states had to submit their IDEA data twice – once to EDEN and once to the Office of Special Education Programs (OSEP) office’s Data Analysis System (DANS). In an effort to reduce this burden on states, ED allowed some states to submit some of their DANS data to EDEN only. For the 2004-2005 IDEA exiting data (1820-0521) as well as the 2005-2006 IDEA child count (1820-0043) and environment data (1820-0517), a select number of states submitted portions of their IDEA data to ED directly through the EDEN data system. These states were authorized to use this new “EDEN-only” process based on a comparison, performed in the fall of 2005, between the data they submitted to EDEN and the data they provided directly to the OSEP program office previously.

The step-by-step procedure for state IDEA data managers to submit their data to the EDEN data repository for the 2004-2005 reporting year for exiting data and the 2005-2006 reporting year for child count and environment data follows.

The EDEN review team notified those states that were granted “EDEN-only” status for a specific IDEA data collection form. At the same time, the team notified ED-OSEP, Westat-OSEP, and the EDEN Partner Support Center (PSC) of which states had been authorized to provide data through EDEN and did not also have to continue to submit data to the DANS.

Once the “EDEN-only” states were notified, the EDEN team sent instructions to both state OSEP and EDEN data coordinators and also to the PSC. These instructions explained the new data submission process and the steps for submitting data to EDEN.

Data for “EDEN-only” states were processed within EDEN along with all other file submissions. As the IDEA data for an “EDEN-only” state came into EDEN, the data files passed through the standard EDEN internal validity checks. If the data failed these initial checks, the state EDEN coordinator received a standard EDEN system e-mail, checked the data file for errors, conferred with the state IDEA data manager, made adjustments as necessary, and resubmitted the data. If errors occurred following resubmission/update of the data, PSC was available to assist the state in submitting their data into EDEN. If Westat-OSEP received a submission from a state with “EDEN-only” status, they did not process the data. Rather, they contacted the state and PSC to request that the state data came through EDEN.

Once the data passed the EDEN validity checks, PSC sent the state EDEN coordinator and the IDEA data manager .PDF reports of the data as now resident in EDEN. That data was loaded into the EDEN Submission System’s (ESS) Staging Database. PSC checked submissions on a daily basis to see if those states with an “EDEN-only” status for a particular file had submitted that data file. When a file from an “EDEN only” state was loaded into the ESS Staging Database, it triggered PSC to run a report of the data in that file. This data was extracted into Excel, mapped to the OSEP DTS format, and emailed to Westat-OSEP.

Westat-OSEP notified the state IDEA data manager and the state EDEN coordinator within three to five days of the submission that they received their state data. Within ten days, Westat-OSEP prepared standard edit check materials that they then sent to the state OSEP coordinator with a copy to the state EDEN coordinator. Westat-OSEP then contacted the state IDEA data managers to resolve questions regarding their data.

Letters signed by John Hager, Assistant Secretary of the Office of Special Education and Rehabilitative Services, and Tom Luce, Assistant Secretary of the Office of Planning, Evaluation, and Policy Development were sent to states that successfully submitted comparable special education data to the EDEN system and to the DANS collection. Fourteen states were exempted from submitting the IDEA exiting data to DANS; 19 were exempted from submitting the IDEA child count data; and 13 were exempted from submitting special education data on educational environments. ED is currently reviewing 2004-2005 Exiting (1820-0521) and 2005-2006 Child Count (1820-0521) and 2005-2006 Educational Environment (1820-0521) data to determine which states can be granted a DANS exemption on the next cycle of OSEP collections.

Common Core of Data (CCD) Data Collection

In the 19th Annual Management Information System Conference (February 2006), Mark S. Schneider, Commissioner for the National Center for Education Statistics, announced that starting with the 2006-2007 school year, the Common Core of Data (CCD) collection (1850-0067) will be submitted through *EDFacts* as a pilot test and states would not be required to submit these data twice.

He said that NCES and *EDFacts* began serious conversations in early February 2006 on merging CCD and *EDFacts* and it was decided that the current technology of the two submission systems is compatible and the process of reconciling the content and collection periods appears to be feasible. He proposed that the submission of 2006-2007 school year data be a transition and, upon success, the merger will become formal for submission of 2007-2008 data.

In 2006-2007 state reports will come into a single interface that will direct the files simultaneously to the CCD online collection application and to *EDFacts*. States will face the same CCD edits and have the same data quality requirements for locking their submissions, as they do now. However, the data sets will be submitted on the *EDFacts* collection schedule. This means that the separate CCD data files will be submitted incrementally when they become available instead of being held and submitted as a package together when the last file is available. This also means that the “prior year” data collected by CCD will be submitted as part of the appropriate *EDFacts* data files.

In 2007-2008, the state CCD reports will come into a single interface that integrates the on-line edits of both *EDFacts* and CCD – one set of edit reports, one help desk. The files will go to the CCD and *EDFacts* databases. CCD will carry out further quality control and add NCES IDs, geocode information, standardize agency or school name and address, etc. The edited data will be forwarded to the *EDFacts* database and to the NCES Web applications.

NCES and *EDFacts* staffs are currently working to implement this vision. As part of this team effort the CCD Coordinators from each state will meet with the state *EDFacts* coordinators in September 2006 for joint discussion and training sessions.

Office of Indian Education Student Count

The *EDFacts* collection this year is also assuming the paperwork burden for the Office of Indian Education's student count collection (formerly approved as part of 1810-0021). The burden hours for the collection of these data are from approximately 10,000 parents completing the student registration form for Indian students each year, and 1,100 districts receiving the grants reporting the student count numbers to the Department of Education. You may review the student form in Attachment F.

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.

The primary customers for these education data will be the program managers and analysts at ED. These data will be used to evaluate, in a more integrated way, the effectiveness and efficiency of federal education programs, with the intent to improve program management and focus our budget on those federal education programs that provide the best education outcomes for the nation's students and their families. State

and district education agencies will also be able to use this information in evaluating their education status and progress. Eventually the public, including parents and students, will have access to this information and be able to display information about the condition of education in their neighborhoods. This public data will have all of the necessary privacy and security requirements enforced.

Except for this collection, there is no previous use of the data in this form and relational data format. However, almost all of the data collected in this effort have been collected and used by various ED program offices in their independent and separate management and analysis of their respective federal education programs.

The Civil Rights Data Collection data are used to enforce the laws to ensure equal access to education and promote educational access throughout the Nation.

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

The *EDFacts* system allows each State education agency (SEA) to provide the *EDFacts* data in the most automated, electronic format that the technology of that specific SEA can currently support. Provisions were made to receive data in multiple formats to ensure the least possible burden on each of the participants. Additional ED contracts were established to provide the SEA with funds, expert technical assistance, and training including the identification of current best practices and knowledge sharing opportunities practiced by the states.

The *EDFacts* Survey Tool is a Web-based, computer-assisted data collection system that works by storing the survey instrument, as well as relevant "preloaded" information within the computer and displaying questions and data entry fields for the respondent in program-controlled sequences on a computer screen. Through computer control of the data collection process and the monitoring of responses, the computer-assisted system offers the capacity for substantial improvements in data quality and data collection efficiency over the standard format of the survey conducted using paper and pencil. The incidence of missing and inconsistent data is greatly reduced since questionnaire skip patterns are computer controlled. Moreover, invalid entries, contradictory entries, or entries inconsistent with available data on the sample school or the sample school's local education agency (LEA) are questioned by the computer and must be resolved or confirmed by the respondent during the self-directed Web instrument data collection.

In this next collection cycle the Civil Rights Data Collection will accept paper forms submission only by exception. If a district requests a paper form exception and ED agrees, the district respondent will be provided a paper version of the final Web form. The policy of complete electronic submission of data is in force for the 2006 data

collection. Part of that electronic submission will include allowing states and districts to submit their civil rights data on a mainframe tape as they have in past years.

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.

For the first few collection cycles as states transition to *EDFacts*, there will be some redundancy of data collected as the programs continued to use their current collections and the *EDFacts* team worked to establish the data quality and validity of these data. In order to transform the culture of information collection within ED, it is paramount to assure that current functioning collections transition to the new *EDFacts* system. Abrupt terminations would only ensure failure. ED decided to use a separate, parallel collection of these data because none of the existing, current ED program sponsored collections could be effectively modified to meet the larger multi-program requirements.

ED determined that no current program collection would be replaced or modified until a certified shared data repository could be established. Federal program managers' and analysts' requirements for data have increased under NCLB and they must be allowed to collect the information they need. The *EDFacts* data repository that has been established will enable ED to begin to determine what amount of required education program data can be annually collected and shared among the federal program offices. As the data quality (including completeness, timeliness, accuracy, validity) is being established, an internal ED review of all elementary and secondary education program collections has been underway to determine which collections can be reduced or eliminated. More details about the program data collections currently under review and consideration for reduction or elimination are found in Attachment D. The schedule and strategy for reviewing and amending these collections is also provided. Much progress has already taken place with a number of ED program offices.

Collaboration with ED Program Offices

The progress made with OSEP, OCR and NCES on the CCD has been described in the response to question #1 above. The *EDFacts* staff has been working closely with other ED program offices to utilize all of the tools and data developed for *EDFacts* to reduce the burden of their collections. Although these collections were (and will be) cleared under separate OMB numbers, there has been much progress. The annual Consolidated State Performance Report (CSPR) required by the Office of Elementary and Secondary Education has been the most notable.

OMB recently approved (1810-0614) the CSPR to collect data for the 2004-2005 school year. The paperwork burden on this collection was reduced by providing the respondents with data in *EDFacts* that enabled them to use *EDFacts* data in their submission of the CSPR.

The Office of English Language Acquisition is committed to collect the data required by their Biennial Report (1885-0553) using existing and expected *EDFacts* data to reduce paperwork burden on their respondents as done with the CSPR.

The Office of Indian Education, in the Office of Elementary and Secondary Education, intends to collect the data for their Formula Grant to States (1810-0021) similar to the CSPR using *EDFacts* data to reduce the paperwork burden on their respondents.

The work of identifying potential duplication and poor data in existing ED collections and moving the sponsoring program offices into the *EDFacts* repository will extend over a number of years. OMB and SEAs need to understand that such burden reduction involves extensive training and some significant changes in the state and ED business processes that currently exist.

Collaboration with ED Staff Offices

Regulatory Information Management Service (RIMS) in the Office of Management in ED is in the beginning stages of implementing a transformation in the processes that govern ED approval of information collection requests by all of the ED program offices. The RIMS team has worked very closely with the *EDFacts* staff in the analysis of the common and unique characteristics of the current federal elementary and secondary data collections. The RIMS team, under their PRA enforcement authority, will continue this transformation process with all ED data collection activities.

The major focus of the RIMS transformation process is to ensure that the whole life cycle of the information is analyzed whenever program offices propose data collections. RIMS is committed to ensure that the practical utility of the data is established before granting their approval to automate current collections and that the storage, protection, use, and dissemination of the data are understood before the collection process is approved. The automation of any paper forms without applying these transformation principles would simply perpetuate an existing problem of collecting data that cannot be used effectively.

The *EDFacts* team has been working to ensure that planning and system development are consistent with the agency's OMB-approved enterprise architecture, so that alignment with the Federal Enterprise Architecture is maintained continuously. At least one person attends an agency-wide "data and XML steering committee" sponsored by ED's Enterprise Architecture team, which is pushing to align definitions and schemas across all programs and whose membership includes the major K-12 and postsecondary data standards organizations.

Further, the team has begun to relate its work to other drivers of burden reduction in K-12, such as the streamlined collection of discretionary grant applications via Grants.gov and the redesign of ED's financial system. The Office of the Chief Financial Officer expects that application for 100% of ED discretionary grants will be made through the Grants.gov "Apply" module, rather than through ED's financial system or via paper, by 2007 fiscal year end. OCFO is coordinating with the Grants.gov program management

office, which reports that its Grants Executive Board (GEB) is looking into grant performance-reporting alternatives, which may or may not involve Grants.gov.

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

The ED*Facts* collection is a state to federal exchange of data in which the individual abilities of each state are considered in the arrangements to transfer these data. It is not the size of the state but instead it is the level of development of the state's education information systems that determines the level of sophistication to be used in the ED*Facts* data transfer. As described in #3 above, and as reflected in the questions within the attached *Federal Register* notice, ED continues to tailor each exchange to the state's abilities.

6. Describe the 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.

Education in this nation is on an annual cycle where success and failure are measured and reported annually. Federal program funding is also determined annually. Collecting this education information less frequently than annually would greatly diminish the ability of program managers and analysts to use the information to measure education progress in support of the President's No Child Left Behind legislation. Over time, the Civil Rights Data Collection will have most of its data from ED*Facts* each year and only need to collect some minimal number of required data every other year using the ED*Facts* Survey Tool. As the survey tool is fully developed, ED program offices are expected to use the survey tool to collect necessary education information that will not be available in ED*Facts*.

7. Explain any special circumstance that would cause an information collection to 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 a statistical data classification that has not been reviewed and approved by OMB;**

- **that includes a pledge of confidentiality that is not supported by authority established in statute or regulation, that is not supported by disclosure and data security policies that are consistent with the pledge, or which unnecessarily impedes sharing of data with other agencies for compatible confidential use; 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.**

This information collection activity does not have special circumstances that would include any of the requirements listed above.

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.

Describe efforts to consult with persons outside the agency to obtain their views on the availability of data, frequency of collection, the clarity of instruction and record keeping, disclosure, or reporting format (if any), and on the data elements to be recorded, disclosed, or reported.

Consultation with representatives of those from whom information is to be obtained or those who must compile records should occur at least once every 3 years - even if the collection of information activity is the same as in prior periods. There may be circumstances that may preclude consultation in a specific situation. These circumstances should be explained.

Copies of the 60-day Federal Register notice and current 30-day Federal Register notice are in Attachment A. There is also a copy of the electronic invitation ED sent to all of the ED*Facts* state coordinators advising them of their opportunity to provide their comments during the first comment period and a second message that was sent notifying them of the second public comment period. There is also a copy of the April 28, 2006, Notice of Proposed Rulemaking that proposes to provide enforcement capabilities now that submission of ED*Facts* 2006-2007 data is mandatory. The ED*Facts* responses to the public comments received through these notices and our direct communications with the states are in Attachment E.

In order to ensure the best possible information and the least possible paperwork burden on the SEAs, the ED team visited every state, Puerto Rico, and the District of Columbia in 2003 and again in 2004. The prospective ED*Facts* data elements were shared with the information and program managers in each state to determine the availability of each

specific data element from each specific state agency for the first phase of the PBDMI collection. Documentation of these discussions provided evidence that some of the original list of data elements would not be available from a significant number of the States at this time. Some of those data elements were dropped from the *EDFacts* data collection.

The EDEN data collection is being repeated this year through *EDFacts*. The ED team seeks to expand this data transfer to include other important education data in an iterative process of data collection, data quality evaluation, and data use. ED has provided the states with tools to evaluate and document the availability of each of the proposed data groups.

There will be no “free pass” given to the states on any obtainable education data required to manage federal programs and meet the goals of the No Child Left Behind Act. On the other hand, ED does not intend to try to collect data that does not exist. The ED position on whether any particular data element will be collected in future annual submissions is that each data element will be evaluated for its “practical utility” to the government and its availability in the real world. ED will avoid asking for data that is not in current data systems since that data is more likely to be a “best guess” than it is to be an auditable fact. It is ED’s intention to identify the best possible data that is available for use at the time the data is to be provided. ED will use the current public comment period to invite state advice on when the proposed data elements will be available over the next few school years and collection cycles. This is significant because this collection will become mandatory in 2006-2007.

ED also recognizes that some of these data may be available from the school districts or schools even if they are not available from the SEAs. ED reserves the right to ask for additional useful data from the states and, as needed, from the school districts and schools in future *EDFacts* collections. All of these additional data elements will be subject to intense internal ED evaluation through the *EDFacts* and RIMS transformation process and also will undergo the established OMB review and approval process that includes two opportunities for public review and comment.

The *EDFacts* team has maintained close contact and strong communication links with its state partners through regular messages, a full-time Partner Support Center, and regular conference/training sessions. Another *EDFacts* conference was held in the District of Columbia on May 22-23, 2006. ED paid for each state to send two representative data experts. Some states chose to send additional participants at their own cost. In this conference, we discussed the states’ data collection and data availability issues, reviewed the paperwork clearance process, and invited the SEAs to participate by responding to the *EDFacts* 2006-2007 collection proposal. The two *EDFacts* representatives from each state will meet again in a conference in September 2006 to discuss how we will implement the OMB approved data collection proposed in this submission package. At the September conference, the state CCD coordinators will join the *EDFacts* coordinators to help facilitate the merging of CCD into the *EDFacts* collection.

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

There is no current remuneration for any state. In the past ED has committed funds for subcontracts with each of the fifty states, the District of Columbia, and Puerto Rico under the Cooperative Systems contract administered by the NCES. This was the same contract that has provided staff support for the National Assessment of Education Progress (NAEP) in the states. Our discussions with the state data authorities suggest that some states have used those funds for overtime work, some for temporary or part time staff work, and some used it for system changes.

ED is considering proposing federal funding for a federal elementary and secondary data coordinator for each state similar to the NAEP position in each state. This proposal is only in its initial consideration stage but the *EDFacts* team and *EDFacts* state coordinators strongly endorse this approach to support the *EDFacts* federal data collection mandate.

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

There has been no assurance of confidentiality provided to the respondents beyond the agreement to protect individual student information under the Federal Educational Rights and Privacy Act. This issue is specific to the amount of data found in a “cell” that might make the identification of an individual student or staff member possible. ED is committed to protect individual privacy by not making public any data in rows or columns where a single cell is below a certain threshold of size. ED will continue to review the submitted data for any other security requirements.

11. Provide additional justification for any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private. This justification should include the reasons why the agency considers the questions necessary, the specific uses to be made of the information, the explanation to be given to persons from whom the information is requested, and any steps to be taken to obtain their consent.

There are no questions of sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private. The education data collected and stored in *EDFacts* will be data about schools, school districts, and states.

12. Provide estimates of the hour burden of the collection of information. The statement should:

- **Indicate the number of respondents, frequency of response, annual hour**

- burden, and an explanation of how the burden was estimated. Unless directed to do so, agencies should not conduct special surveys to obtain information on which to base hour burden estimates. Consultation with a sample (fewer than 10) of potential respondents is desirable. If the hour burden on respondents is expected to vary widely because of differences in activity, size, or complexity, show the range of burden, and explain the reasons for the variance. Generally, estimates should not include burden hours for customary and usual business practices.**
- **If this request for approval covers more than one form, provide separate hour burden estimates for each form and aggregate the hour burdens in Item 13 of OMB Form 83-I.**
 - **Provide estimates of annualized cost to respondents of the hour burdens for collections of information, identifying and using appropriate wage rate categories. The cost of contracting out or paying outside parties for information collection activities should not be included here. Instead, this cost should be included in Item 14.**

Fifty-two SEAs will supply *EDFacts* data annually. Based on some conversations with the states, ED believes that a reasonable estimate of the average additional burden per state for the transmission of the general PBDMI data will be 2000 hours – or the equivalent of one full time person for a year. Although there is a large amount of data to be provided, ED has initiated a number of provisions to minimize the burden on the states. One of the main provisions of this initiative has been to establish what data is already being produced and maintained by each state and not require the collection of data that is not currently available. In this collection cycle, ED is asking for education data for the 2006-2007 school year in order to ensure that the requested data are currently available in state and district databases. As states continue to automate and improve their education data systems and establish standard data processing procedures for providing the *EDFacts* data, this burden will decrease and become part of their standard daily operating processes.

ED is counting this paperwork burden as a revision to a currently approved collection with a small addition to the total ED paperwork burden budget for the basic data submitted to *EDFacts*. The burden for the previous collection included 52 respondents and 297,000 total burden hours. Last year ED estimated total burden of 12,480 hours (240 hours times 52 respondents) for submission of state data to the *EDFacts* system. ED believes that is the best estimate of *EDFacts* paperwork burden hours for this next year. There are few data groups requested in 2006-2007 and many of those eliminated were school level data that most consider the most burdensome. The remaining burden from the previous collection (284,520 hours) was related to the 2004 Civil Rights Data Collection.

The additional data burden this year will be associated with making this collection mandatory. As part of the exemption process, states will have to develop and submit to the Secretary a data reporting plan that shows when the state will be able to provide each of the requested data groups. This plan will cover data to be submitted over the next

three school years and demonstrate how the state will move from where it is now to the point where it can supply all of the *EDFacts* data in a timely and complete manner. Since most of the work associated with this plan is part of the standard operating procedures in each organization, the *EDFacts* staff estimates the additional burden to provide and update these plans to average 20 hours per state. The total additional burden will be 1,040 hours.

Another additional burden with this collection will be the collection through *EDFacts* of the 2006-2007 school year data from the additional seven providers of the CCD data previously sent to NCES. The average burden associated with collecting CCD data is 102 hours. For the additional seven submitters of CCD data through *EDFacts* the total burden is 714 hours.

The *EDFacts* collection this year is also assuming the paperwork burden for the Office of Indian Education's child count collection. The burden hours for the collection of these data are produced by two activities. Approximately 10,000 parents complete the form for Indian students (see Attachment F) each year. With the average time to complete this form estimated at 15 minutes, the total annual burden is 2,500 hours. The districts receiving these grants must then report the child count numbers to the Department of Education. There are approximately 1,100 grantees submitting these data with an estimated average of .50 hours per grantee to submit the child count data. The total burden for the submission of the OIE child count data is $550 + 2,500 = 3,050$. The collection of these data and the calculation of this burden is now part of *EDFacts*.

The Civil Rights Data Collection will require additional burden this year and will be 462,000 total hours. Approximately 6,000 school districts and 60,000 schools will respond to the request for data. The average amount of time for each response is estimated to be seven hours. Since the number of schools per district varies so widely it should be noted that a district with only one school would, on the average, take seven hours for the district report and seven hours for the school report with a total of 14 hours. A district with 50 schools would take seven hours on the district report and average seven hours on each of the schools reports for a total of 357 hours.

	Respondents	Hours/Response	Total Hours
Previous – EDEN Data	52	240	12,480
Previous – Civil Rights Data			284,580
Previous – Total Burden	52		297,060
New – <i>EDFacts</i> Data	52	2000	104,000
New – Territories Submitting CCD Data	7	102	714
New – State Data Plans	52	20	1,040
New – Civil Rights Data	6,000	77	462,000
New – Indian Education Student Form	10,000	0.25	2,500
New – Indian Education Student Data	1,100	0.5	550
New – Total Burden	17,152		570,804

Please see the response to #15 below for a more complete explanation of this burden calculation. Any future additional ED program use of the *EDFacts* Survey Tool will be submitted to OMB for prior approval of the instrument and the burden.

Over the next few years, as RIMS implements its transformation of the current system, RIMS expects to ask OMB to adjust the burden hours on some ED information collection activities that collect elementary and secondary education data. Specific details about collections and extent of burden reduction are found in Appendix D and as mentioned in the response to #4 above.

The annualized cost is estimated to be \$ 22,832,160 (570,804 hours multiplied by an average wage of \$40 per hour). There is a wide range of hourly salaries associated with the professionals that will provide this data.

13. Provide an estimate of the total annual cost burden to respondents or record keepers resulting from the collection of information. (Do not include the cost of any hour burden shown in Items 12 and 14.)

- **The cost estimate should be split into two components: (a) a total capital and start-up cost component (annualized over its expected useful life); and (b) a total operation and maintenance and purchase of services component. The estimates should take into account costs associated with generating, maintaining, and disclosing or providing the information. Include descriptions of methods used to estimate major cost factors including system and technology acquisition, expected useful life of capital equipment, the discount rate(s), and the time period over which costs will be incurred. Capital and start-up costs include, among other items, preparations for collecting information such as purchasing computers and software; monitoring, sampling, drilling and testing equipment; and record storage facilities.**
- **If cost estimates are expected to vary widely, agencies should present ranges of cost burdens and explain the reasons for the variance. The cost of contracting out information collection services should be a part of this cost burden estimate. In developing cost burden estimates, agencies may consult with a sample of respondents (fewer than 10), utilize the 60-day pre-OMB submission public comment process and use existing economic or regulatory impact analysis associated with the rulemaking containing the information collection, as appropriate.**
- **Generally, estimates should not include purchases of equipment or services, or portions thereof, made: (1) prior to October 1, 1995, (2) to achieve regulatory compliance with requirements not associated with the information collection, (3) for reasons other than to provide information or keep records for the government, or (4) as part of customary and usual business or private practices.**

The collection of ED*Facts* data for the foreseeable future will require no additional systems development efforts by the states. The states are currently seriously involved in the development of state education information systems for their own use and for reports in response to the No Child Left Behind Act and other education legislation. The guidance, standards, and “best practices” developed by ED*Facts* will actually help reduce

the total costs associated with those systems development activities by providing cost effective common education information management solutions to the states and districts. Much of this work may be done under the recently funded Institute of Education Sciences (IES) discretionary grant for states to develop longitudinal state education data systems.

14. Provide estimates of annualized cost to the Federal government. Also, provide a description of the method used to estimate cost, which should include quantification of hours, operational expenses (such as equipment, overhead, printing, and support staff), and any other expense that would not have been incurred without this collection of information. Agencies also may aggregate cost estimates from Items 12, 13, and 14 in a single table.

The Performance Based Data Management Initiative (PBDMI) had a lifecycle development cost of \$44.1M while the steady-state maintenance costs are currently estimated to be \$4.1M per year. This investment enables estimated \$145M in lifecycle benefits at the federal level by eliminating redundant collection systems as well as an additional \$124M in lifecycle benefits at the state level by reducing the burden of data collections. The current budgets to support the collection of ED*Facts* data for FY 2006 and FY 2007 are, respectively 4.7 million and 6.6 million.

The estimated government cost for the next Civil Rights Data Collection is \$1,200,000.

15. Explain the reasons for any program changes or adjustments reported in Items 13 or 14 of the OMB Form 83-I.

Part of the difference in burden, an increase of 182,224 hours, is a program change. This is a result of: 1) reporting all of the burden for the Civil Rights Data Collection in the year the burden is earned (+177,420); 2) adding the burden of the 7 territories who will submit 2006-07 CCD data through EDEN this year (+714); 3) the requirement that each state develop a data submission plan as part of the mandatory submission requirement (+1,040), and 4) adding the Indian Education student count (+3,050). The rest of the increase in the estimated ED*Facts* paperwork burden (91,520 hours) is an adjustment.

This is a request for clearance for a currently approved collection with changes. All conditions for the current clearance are fully addressed in Attachment D. This includes a letter from Assistant Secretary Tom Luce to OMB on April 2006. Last year ED estimated total burden of 12,480 hours (240 hours times 52 respondents) for submission of state data to the ED*Facts* system. ED has increased this estimate of burden hours to the equivalent of one full-time employee for this next year.

The new additional hours for the development and submission of a state data reporting plan is estimated to be 1,040 hours. Additional hours associated with the seven additional submissions of CCD data is 714 hours.

EDFacts also includes the paperwork burden for the Office of Indian Education's student count collection (formerly approved as part of 1810-0021). The burden hours for the collection of these data are produced by two activities. Approximately 10,000 parents complete the form for Indian students (see Attachment F) each year. With the average time to complete this form estimated at 15 minutes, the total annual burden is 2,500 hours. The districts receiving these grants must then report the child count numbers to the Department of Education. There are approximately 1,100 grantees submitting these data with an estimated average of .50 hours per grantee to submit the child count data. The total burden for the submission of the OIE child count data is $550 + 2,500 = 3,050$. The collection of these data and the calculation of this burden is now part of *EDFacts*.

The Civil Rights Data Collection (previously 1870-0500) estimates 462,000 total hours to be incurred in the next twelve-month period. This represents an average burden of 77 hours for each of the 6,000 responding districts. It appears that most school districts complete both the district and school reports while some districts have their schools submit the school data.

The total *EDFacts* burden estimate is 570,804 hours. This burden estimate includes the whole Civil Rights Data Collection burden, declaring it all on the year when the collection is conducted instead of declaring only half each year as previously submitted. After the completion of the 2006 Civil Rights Data Collection, ED expects to stabilize the annual collection burden, including selected program office surveys, at a significantly lower annual number.

16. For collections of information whose results will be published, outline plans for tabulation and publication. Address any complex analytical techniques that will be used. Provide the time schedule for the entire project, including beginning and ending dates of the collection of information, completion of report, publication dates, and other actions.

At this time there is no definitive plan or an ED schedule for the publication of the data collected and stored in *EDFacts*. When the data quality is certified and the *EDFacts* database is determined to be the primary source of ED elementary and secondary education information, ED will return to OMB for discussions and approval of its publication plans.

It is expected that the GEPA data previously collected will be presented in narrative and tabular form as it has been in previous reports. ED will submit the next GEPA 424 report to Congress by August 15, 2006.

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.

This collection has displayed the OMB approval date in the transmittal documents requesting the information from the State agencies and in any written discussion or representation of the collection. There will be no paper *EDFacts* "form" for the general collection upon which to display the OMB number. The OMB number will be properly displayed on any Web form used by the *EDFacts* Survey Tool including the Civil Rights Data Collection.

18. Explain each exception to the certification statement identified in Item 20, "Certification for Paperwork Reduction Act Submissions," of OMB Form 83-I.

ED is requesting no exemptions from the Certification.

B. Collections of Information Employing Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results.

Under the No Child Left Behind Act, every school and school district must meet the education standards in the law. Sample data are inadequate to measure compliance with the law at that level.

This is a collection of elementary and secondary education information from the universe of schools, school districts, and states. It includes the basic demographic and assessment data that is needed to evaluate the status and progress of education in this nation. As a primary agency source of universal education information it may be used to efficiently draw samples from which statistical studies may be conducted in the future. When these studies are being considered, ED will return to OMB for discussions and approval.

Civil Rights Data Collection

The ED Civil Rights Data Collection, using the *EDFacts* survey tool, does employ statistical methods as described below.

Potential Respondent Universe

The design is a multi-state rolling stratified sample of public school districts in the nation, including charter schools. The sample was drawn, for each state, to include strata divided by size of district, and, within each stratum, high or low number of minority students, in order to ensure the necessary coverage for the preparation of state (and national) estimates, and to meet the need for data on specific districts.

In the sampling design for the survey, OCR merged the CCD universe maintained by the NCES with the Time Series Database from the E&S Survey to form a universe for the sample. The sample was drawn using a stratified random sampling methodology with

district size as the determinant for the six strata. In order to meet its data needs, OCR determined that a sample of 5,155 districts would be drawn through the rolling stratified sampling process that OCR continues to use for the ED Civil Rights Data Collection. In addition, to meet the need for the inclusion of specific districts in the sample, OCR will be sampling 750 districts with certainty, including: school districts with 25,000 or more students; school districts under Federal Court Order with the U.S. Justice Department; school districts recently released from Federal Court Order; and school districts in states with 25 or fewer schools. OCR will survey as a separate sample, approximately 160 of an estimated 1,155 cooperative, BOCES-type regional education centers functioning as school districts and their individual educational facility programs.

The total sample is approximately 6,000 public school districts and BOCES-type regional education centers providing full day educational services to children, and the approximately 64,000 schools corresponding to these districts.

Procedures for the Collection of Information

The ED Civil Rights Data Collection sample was the first cycle of a planned three-cycle rolling stratified sampling process—which is the sampling process used for the 1998 and 2002 Elementary and Secondary Surveys and the 2004 ED*Facts* Civil Rights Data Collection. The rolling stratified sampling process includes a disproportionate, random sample of districts within states. The basic sampling proportion is approximately 33%. States that have fewer districts had a proportionately larger sampling proportion and states with large numbers of districts had a small sampling proportion.

OCR first ranked states, in ascending order, by a vector, which is the product of the total number of districts in a state times the number of students in that state. In order to treat states with a smaller number of districts equitably with states with a larger number of districts, a descending proportion value was applied to each state, in rank order. This proportion value, or “sampling fraction” varied from 33.5% of all districts to 53.8% of all districts. States that occupy middle ranks received the general sampling proportion. The sampling proportion was increased progressively for the states with progressively fewer districts, and decreased progressively for states with larger numbers of districts. The proportion steps was .45 percent, so that no state had a sample, for districts not with certainty, of less than 33.5% of its districts selected.

The sampling proportion for each state was then additionally adjusted through a proportional adjustment so that the results of the original multiplication yield 5,155, the number desired for the sample of regular school districts. OCR will then stratify the districts by total enrollment in five ranges, as follows:

- 1 to 300 students;
- 301 to 3,000 students;
- 3,001 to 5,000 students;
- 5,001 to 25,000 students; and
- Districts to be surveyed with certainty.

OCR then calculated a total number of districts in each stratum, by state, and total number of schools in each stratum, by state. OCR also calculated the percentage of districts in each stratum, by state, based on the total number of districts in the state. To determine how the sample would be drawn for that stratum, the proportion in each stratum was multiplied by the number of districts sampled for the state. For example, a state that had 25% of its districts with total enrollment of 300 students or less would have 25% of its sample size drawn from the 300 or less stratum of districts in that state.

Calculations were performed to determine the percentage minority enrollment in each district within each stratum. Based on a median split of districts within each stratum, the districts were divided into two halves – high minority enrollment districts and low minority enrollment districts – with each set of districts receiving a proportionate distribution to the number of districts in each minority substratum. Procedures were then executed to minimize overlap with previous years of the rolling stratified sample, by dividing the sampling frame into four categories depending on selection probabilities and whether or not the district was selected in prior years. The sample was then selected with probabilities proportionate to the conditional probabilities.

The formula for drawing the sample was used in the majority of cases. However, there are two exceptions, based on the size of the district and the number of districts in each state:

- The first is that states with 25 districts or less, such as Hawaii, Delaware, Nevada, Maryland, and the District of Columbia, had all their districts selected;
- The second exception is for districts with an enrollment of over 25,000 students. These districts were placed in a separate stratum and selected with certainty.

The random sample will be drawn, by stratum, within each state.

Methods to Maximize Response Rate and to Deal with Non-Respondents

Historically OCR has had a very high response for the Civil Rights Survey. The 2000 E&S Survey was sent to a universe of all school districts and schools in the United States. The overall response rates were 97% of all school districts and 99 % of all schools. The overall response rates for the 2002 E&S Survey were 98% for school districts and 98% for schools. For the 2004 Civil Rights Data Collection, the response rates, including partial respondents to the data collection, were approximately 97% of all districts (95% of all districts responded completely with clean data-which is consistent with the 2002 Elementary and Secondary Survey), and 97% of all schools. If school districts fail to respond in a timely manner, the contractor for the data collection, with assistance from OCR as necessary, contacts the district until either the district responds, or it is designated as a non-respondent.

Tests to be Undertaken

In the past, OCR has used respondent comments and the results of edit checks of the most recent survey data to revise data items in future surveys. For the ED Civil Rights Data Collection, OCR is participating with *EDFacts* in a process for “usability testing” and pilot testing to further revise the data items.

Individuals Consulted on Statistical Aspects of the Design

Dr. Jerome Kravitz, Statistician, Office for Civil Rights, U.S. Department of Education, Washington, DC: (202) 205-9506

Dr. Adam Chu, Statistician, Westat, Inc., Rockville, Maryland: (301) 251-4326