



**CRDC**  
Civil Rights Data Collection

# The CRDC Improvement Project

Recommendations for the 2013–14 and 2015–16 collections and beyond

**DRAFT**

---

*This page is intentionally blank.*

# **The CRDC Improvement Project**

**Recommendations for the 2013–14 and  
2015–16 collections and beyond**

**Prepared by**

**American Institutes for Research**

**Sanamatrix**

**July 2014**

# Contents

- Contents..... ii
- List of Exhibits, Tables, and Figures..... iii
- Introduction..... 1
- Goals of the CRDC Improvement Project..... 3
  - Goal 1 | Reduce reporting burden..... 3
  - Goal 2 | Achieve better respondent engagement through better communication..... 3
  - Goal 3 | Achieve better data quality through better data collection tools..... 4
  - Goal 4 | Make data more useful and accessible to CRDC stakeholders..... 4
- About the CRDC Improvement Project Research Tasks..... 6
  - Task 1 | Review of known issues..... 6
  - Task 2 | Expert review of CRDC design..... 7
  - Task 3 | Site visits..... 8
  - Task 4 | Cognitive interviews..... 10
  - Task 5 | Pilot test..... 10
- Recommendations..... 11
  - Reduce reporting burden..... 11
    - Recommendations for reducing burden..... 11
  - Achieve better respondent engagement through better communication..... 36
    - Recommendations for communication..... 36
    - Recommendations for the Partner Support Center (PSC)..... 47
  - Achieve better data quality through better data collection tools..... 50
    - Recommendations for the online data collection tool..... 50
    - Recommendations for data validation and error reports..... 57
    - Recommendations for flat file submissions..... 61
  - Make data more useful and accessible to CRDC stakeholders..... 66
    - Recommendations for reporting back to stakeholders..... 66
- Challenges for Implementation..... 69
- Proposed Timeline of Improvements..... 70
- References..... 71

# List of Exhibits, Tables, and Figures

Figure 1   Recommended distribution of work focus by goal for the 2013–14 and 2015–16 collections and beyond.....	5
Table 1   Number and role of site visit participants at the OCR field offices, SEAs, and LEAs, by size of LEA.....	9
Exhibit 1 Examples of time burden for CRDC reporting.....	24
Exhibit 2 List of data elements reported as difficult and reasons for difficulty, by LEA size.....	27
Exhibit 3 Illustration of types of LEA data systems.....	34
Exhibit 4   Sample CRDC Calendar.....	45

# Introduction

The Civil Rights Data Collection (CRDC) is a legislatively authorized mandatory survey that collects data on key education and civil rights issues in our nation's public schools. These data are used by the U.S. Department of Education's (ED) Office for Civil Rights (OCR), by the Institute of Education Sciences' National Center for Education Statistics (NCES) and other ED offices, and by policymakers and researchers outside the Department of Education.

The CRDC is a longstanding and critical component of the overall enforcement and monitoring strategy used by the OCR to ensure that recipients of the Department's federal financial assistance do not discriminate on the basis of race, color, national origin, sex, or disability. OCR relies on the CRDC data it receives from public school districts as it investigates complaints alleging discrimination, determines whether the federal civil rights laws it enforces have been violated, initiates proactive compliance reviews to focus on particularly acute or nationwide civil rights compliance problems, and provides policy guidance and technical assistance to educational institutions, parents, students, and others. To meet the purpose and intended uses of the data, the CRDC collects information on public school characteristics and about programs, services, and outcomes for students, disaggregated by race/ethnicity, sex, limited English proficiency, and disability. Information is collected on where students receive instruction so that the data OCR collects accurately reflects students' access to educational opportunities at the site where they spend the majority of their school day.

CRDC data have been collected directly from local education agencies (LEAs) covering each of the 50 states and the District of Columbia since 1968, primarily on a biennial basis (i.e., in every other school year). Recent CRDC collections have also included hospitals and justice facilities that serve public school students from preschool to grade 12. Rather than rely on samples, recent collections cover the universe of all LEAs and all public schools, as well as state-operated facilities for students who are deaf or blind and publicly owned or operated justice facilities that provide educational services to youth. The universe includes all public entities providing educational services to students for at least 50 percent of the school day. Data are currently collected using an online data collection tool.

## Need for improvement

Feedback from prior CRDC collections indicates that LEAs have experienced unacceptable levels of reporting burden. The issues documented fall into two categories: content and the data collection tool. Content issues that contribute to high levels of reporting burden include being asked to provide data already reported by the state education agency (SEA) and being asked to

provide data that are not maintained by schools and LEAs at the level of detail required by the CRDC. Another content issue is a lack of clarity in the definitions of key terms. Respondents also reported that the 2011–12 CRDC data collection tool had performance issues; in particular, that there were not enough built-in edits and some edit messages were unclear.

Based on this feedback, the purpose of the CRDC Improvement Project will be to develop a new data collection tool and processes that (1) reduce respondent burden, (2) improve data response, (3) improve data quality, and (4) make CRDC data more useful and accessible to CRDC stakeholders.

To begin, NCES commissioned a set of research tasks to gather information about the challenges that LEAs, SEAs, and schools have in responding to the CRDC. These research tasks consist of a review of known issues; expert review of the survey design; site visits to LEAs, SEAs, and schools; cognitive interviews about survey language and wording; and pilot testing of the new data collection tool. This report, at current writing, is based primarily on the review of known issues and the site visits.

### **Organization of this report**

This report presents an overview of the main goals of the CRDC Improvement Project, a summary of the research tasks, a description of the challenges encountered by LEAs that contribute to the excessive response burden, comprehensive recommendations for achieving the project goals based on known issues and site visits, key challenges for implementing improvements, and a suggested timeline for improvement activities.

# Goals of the CRDC Improvement Project

The CRDC Improvement Project aims to achieve four main goals: (1) reduce reporting burden; (2) achieve better respondent engagement through better communication; (3) achieve better data quality through better data collection tools; and (4) make data more useful and accessible to CRDC stakeholders. The justification for these goals is explained below. Within each goal there are actions that can be taken for the 2013–14, 2015–16, or future collection cycles. Figure 1 illustrates the CRDC collections during which primary efforts toward meeting each goal are recommended distribution of work focus for each goal by data collection year.

## Goal 1 | Reduce reporting burden

### Recommended focus collection: 2013–14, 2015–16, and beyond

A universal theme among site visit respondents and public comments in response to the Office of Management and Budget (OMB) submission was that participation in the CRDC required numerous hours of sometimes multiple staff members to report and validate data, thus making the process exceedingly burdensome. The process of gathering data, preparing data for submission, and submitting data was arduous and could take LEAs weeks to complete; there were over 10,000 calls received or made by the Partner Support Center in the last administration. One site visit respondent described the CRDC as “a huge burden...our largest non-funded mandate.”

The level of burden impacts data quality. Given that staff have a limited number of hours to divert from their intended tasks, the greater number of hours inputting and uploading data limits the amount of time for validation and impacts data accuracy and quality. Further, when respondents view a mandatory data collection as a burdensome requirement, they will provide the minimum information necessary to comply with the request, at the very last minute possible, and avoid important steps in planning and quality control.

## Goal 2 | Achieve better respondent engagement through better communication

### Recommended focus collection: 2015–16

All of the CRDC points of contact (POCs) at the sites visited noted that the CRDC was a mandatory data collection, they had to fill it out, and that there could be penal consequences if they did not respond. One site visit respondent described the purpose of the CRDC as “punitive” and another reported emails about the CRDC to their SEA as potential email “spam” because they sounded “ominous.” While compliance is a necessary outcome, moving the perception of the CRDC from a negative compliance request to a positive, salient, and useful data tool for



CRDC respondents and stakeholders is vital to improving the response propensity of LEAs and SEAs. Engaged respondents are more likely to respond early and plan better. Early response decreases the need for costly and time-consuming follow-up efforts, and planning for the data collection decreases missing data due to incomplete or uncollected data elements and other errors. Communication about the purpose and value of the CRDC is central to changing its perception.

Also critical to developing and sustaining better response propensities among CRDC respondents is a communication plan that adequately sets and maintains expectations about roles, responsibilities, and deadlines. Universally, very few LEA or school staff other than POCs who were involved in collecting data related to the CRDC (e.g., guidance coordinators and school principals) knew that they were providing data for the CRDC. Even the name “CRDC” was unfamiliar to them. Improving the consistency, reach, type, and timing of CRDC communications will help respondents understand, anticipate, and respond to the data request.

### **Goal 3 | Achieve better data quality through better data collection tools**

#### **Recommended focus collection: 2013–14**

The data collection tools for the CRDC are the core components of a successful data collection. These tools guide the respondent through the data reporting process and produce the final data files that OCR relies on to measure and monitor civil rights compliance in American public schools. It is crucial that these tools are easy to use, valid, and reliable for requesting and delivering high-quality, accurate data. However, LEAs have reported problems with the prior CRDC online data collection tools that suggest the tools did not effectively meet these criteria.

Best practices in survey methods and web survey design, along with more flexible and powerful software, can improve the use, validity, and reliability of the CRDC data collection tools—resulting in better data quality.

### **Goal 4 | Make data more useful and accessible to CRDC stakeholders**

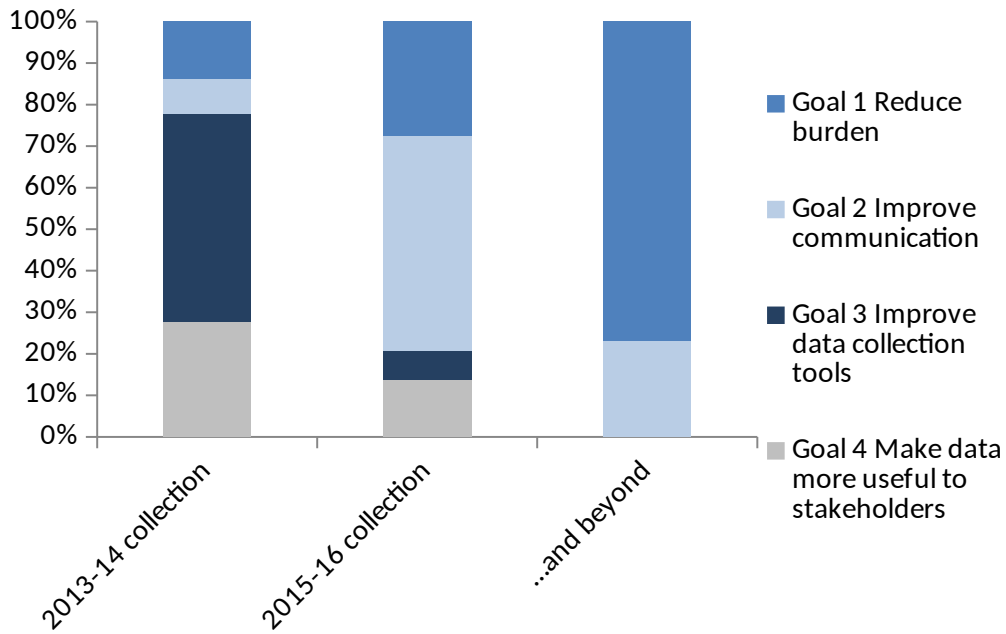
#### **Recommended focus collection: 2013–14**

Another way to assist SEAs and LEAs in understanding the value and importance of the CRDC is to raise awareness about the creation and release of final CRDC data files and reports and to return data back to SEAs and LEAs for their use in strategy and policy planning. Additionally, returning data to SEAs and LEAs can assist them in the review and quality assurance of their own data collection programs.

Most importantly, it is critical for CRDC respondents to feel that they are stakeholders in the data collection. Its value beyond compliance is determined by how helpful the data and data reports are to the people who can directly effect the changes needed to increase or maintain access to educational opportunities for all students.

Figure 1 below illustrates the recommended distribution of work focus for each goal by data collection year. The primary focus for the 2013–14 collection is to improve the data collection tools (Goal 3) and to make the data more useful to stakeholders (Goal 4). These goals were viewed as the priority due to the difficulties respondents experienced with the 2011–12 data collection system, and because developing a new system is the focus of the contract funding the CRDC Improvement Project for the first 2 years. Goals 1 and 2 are viewed as the priority for the 2015–16 and future data collections because, while important, addressing these goals requires more lead time than is available for the 2013–14 collection.

**Figure 1 | Recommended distribution of work focus by goal for the 2013–14 and 2015–16 collections and beyond**



# About the CRDC Improvement Project Research Tasks

NCES has commissioned a set of research tasks for gathering information about problems LEAs, SEAs, and schools have in responding to the CRDC. These tasks consist of a review of known issues; expert review of the survey design; site visits to LEAs, SEAs, and schools; cognitive interviews about survey language and wording; and pilot testing of the new data collection tool.

## Task 1 | Review of known issues

Reviewing known issues was the first step in the process of developing new CRDC data collection tools. The American Institutes for Research (AIR) and Sanamatrix CRDC Improvement Project team summarized known issues from materials provided by NCES. Materials for review included CRDC 2011–12 requirements documents, 2011–12 CRDC edit specifications, data quality analysis previously conducted by the American Institutes for Research (AIR) and other contractors, Question and Answer responses prepared by the 2011–12 Partner Support Center, the 2013–14/2015–16 OMB package, and public comments on the 2013–14 proposed data collection.

The summary of known issues documented problems or concerns in four broad areas: (1) the survey tool, (2) data quality, (3) data elements, and (4) survey methods.

1. Issues with the survey tool include performance issues such as edit and range checks, data fills, logic and skip patterns, user controls and navigation, and flat file submission.
2. Issues with data quality include inconsistencies with other NCES school and district data collections, universe coverage, problems with school IDs, outliers, and inability to accurately report data at the school level or other disaggregated levels.
3. Issues with particular data elements are those related to item-specific definitions, burden, and access to data.
4. Issues with the survey methods include those related to the procedures used to solicit, encourage, and assist reporting, such as communication, training, and tools to reduce burden.

## Task 2 | Expert review of CRDC design

NORC, at the University of Chicago, was asked by NCES to review the CRDC tool and recommend ways to improve the quality of the data collected by the program. NORC examined data from a variety of sources: stakeholders' feedback about the data collection tool obtained during a 2013 Management Information Systems (MIS) conference session, a CRDC Work Group meeting led by OCR, and a CRDC tool demonstration led by Acentia; the 2009–10 CRDC and the 2011–12 CRDC restricted-use data files with respondents' comments; and the 2010–11 Common Core of Data (CCD).<sup>1</sup>

More specifically, the research encompassed four major analytic activities:

1. Review and analysis of stakeholders' feedback on the CRDC tool from conferences, meetings and an independent NORC review. These results were used to make overarching recommendations for changes to the tool to improve accuracy and to reduce the edit failure rate and burden.
2. Review of CRDC respondents' comments from the item-level comment fields embedded within the tool. The review focused on understanding what issues respondents experienced and to make informed recommendations for improving the CRDC tool and content (e.g., change item wording or layout, provide more guidance/instructions, add definitions, add skip logic, add or revise the within-tool editing procedures, etc.).
3. Analysis of a subset of items collecting student enrollment and basic school characteristic data in Part 1 of the CRDC tool. Specifically, NORC explored whether a prior cycle of CRDC data and/or CCD data could be used in the current CRDC's edit check procedures.
4. Item-specific analysis of data from the 2011–12 CRDC for selected Part 2 items to identify possible outlier values that could be flagged for editing. As part of this research, NORC examined the extent to which the difference in the reference period for Part 1 (i.e., fall snapshot or point-in-time data) and Part 2 items (i.e., cumulative/end-of year data) contributed to edit check problems and recommended ways of adjusting the edits or the CRDC tool to minimize these problems.

Many of the design issues and recommendations identified in the summary of known issues were also confirmed in the NORC expert review.

## Task 3 | Site visits

---

<sup>1</sup> The Common Core of Data (CCD) is a program of the U.S. Department of Education's National Center for Education Statistics that annually collects fiscal and nonfiscal data about all public schools, public school districts, and state education agencies in the United States. The data are supplied by state education agency officials and include information that describes schools and school districts, including name, address, and phone number; descriptive information about students and staff, including demographics; and fiscal data, including revenues and current expenditures (<http://nces.ed.gov/ccd/>).

NCES, AIR, and Sanametrix conducted a process improvement and feasibility study that consisted of in-person site visits with LEAs, SEAs, schools, and OCR regional offices. The purpose of the site visits was to gather specific information about reporting procedures; periodicity of data availability; problems with specific data elements; suggestions about how the online data collection tool can assist in improving data quality; and the types of feedback reports LEAs, SEAs, schools, and other users would like to be receiving from the CRDC system. A key goal was to understand the process that LEAs use to complete a submission. A set of 13 research questions guided the information gathered from these site visits. The questions were as follows:

1. To what extent are data collected by the CRDC currently maintained as part of a state longitudinal data system? Do SEAs currently have the capacity to provide LEAs or ED with data that is collected by the CRDC?
2. When and how often do SEAs/LEAs collect CRDC-related data?
3. Who are the points of contact (POCs) and what is their data collection role?
4. What is the data collection cycle like?
5. What actions do LEAs need to take to complete their submissions?
6. What format do SEAs/LEAs store the data in?
7. What current CRDC tools are useful for POCs? (e.g., webinars, PSC, forms, templates, online tools)
8. Which data elements are difficult for LEAs or schools to report? Which data elements are easy to report?
9. What is the process for verifying and certifying LEA data? Are subject matter experts consulted for specific elements of the CRDC?
10. What are the reasons LEAs/SEAs collect the CRDC data? How are the data used? Are there specific data elements collected just for the purpose of the CRDC reporting?
11. How do schools report data to LEAs? What is the reporting process and cycle?
12. What can be improved about the CRDC communication process with POCs and other leadership?
13. What other general feedback do LEAs/SEAs/schools/OCR offices have?

Separate interview protocols were developed for LEA, SEA, school, and OCR staff based on these research questions. Participants for the site visits were selected from a list of 25 LEA and SEA POCs provided by NCES. Participants were recruited via telephone and email, and staff of each recruited LEA provided suggestions of schools to visit and school staff to interview. The two OCR field offices visited were selected from a list of six offices provided by NCES and were chosen based on convenience of scheduling with an LEA site visit. The majority of the interviews with LEA, SEA, school, and OCR staff were conducted in person, onsite over a 2-day period; however, due to scheduling constraints, some interviews were conducted via telephone.

Site visits were conducted in 14 different states at 15 LEAs, 11 SEAs, 9 schools, and 2 OCR regional offices. Visits were conducted between February and May 2014, with the majority occurring in February and March. Members from the American Institutes for Research (AIR) and Sanametrix team conducted the 2-day site visits in pairs. The sites visited varied by region, size, and level of sophistication of SEA and LEA data systems and programs offered.

**Table 1 | Number and role of site visit participants at the OCR field offices, SEAs, and LEAs, by size of LEA**

Role of site visit participants	Size of LEA visited and number of participants				
	Small	Mid-size	Large	Very large	Total
<b>OCR field office</b>	1	0	0	1	2
<b>SEA</b>	4	5	0	2	11
Data manager	3	4	0	1	8
Federal data coordinator	0	1	0	0	1
EdFacts coordinator	2	2	0	1	5
RA & QA supervisor	0	1	0	0	1
State liaison for CRDC	2	0	0	0	2
Civil rights compliance coordinator	0	2	0	0	2
IT staff	1	2	0	0	3
<b>LEA</b>	5	6	1	3	15
IT staff	2	4	0	2	8
Data manager	3	3	1	3	10
Office of Accountability staff	0	0	0	2	2
Administrative assistant	1	0	0	0	1
Guidance supervisor	0	1	0	0	1
General counsel	0	0	0	1	1
Athletics coordinator	0	1	0	0	1
Gifted & Talented coordinator	0	1	0	0	1
Research analyst	0	1	0	1	2
Charter school overseer	0	0	0	1	1
Director of Student Life & Services	0	1	0	0	1
Human Resources staff	0	2	0	1	3
Finance staff	0	0	1	0	1
<b>School</b>	3	4	1	1	9
Principal	0	2	0	0	2
Assistant principal	0	2	0	1	3
Data manager	0	0	1	0	1
Data processor	0	1	0	0	1
Secretary	3	1	0	0	4
Education specialist	1	0	0	0	1
Counselor	0	1	0	0	1

NOTE: The OCR field office was in the same region as an LEA visited, and SEAs were in the same states as the LEAs visited.

## Task 4 | Cognitive interviews

During May 2014, NCES, AIR and Sanametrix conducted telephone interviews with CRDC respondents (primarily LEAs) to gather information on which data elements on the CRDC school form are confusing to respondents and what information can be added to instructions, definitions, tables, and questions to make the data request easier for respondents. The cognitive interview protocol focused on specific data elements that NCES, OCR, and the site-visit research found to be problematic for respondents. Information about problematic data elements will be documented in a separate report.

Interviews were conducted with 20 participants. Each telephone interview was 90 minutes and up to three respondents who were involved in the reporting of CRDC data elements participated in the interview.

Additionally, modules consisting of topically related groups of CRDC data tables were developed and feedback on them was solicited from the cognitive interview participants. The AIR and Sanametrix team requested feedback on the data table groupings from site visit participants by email and from telephone interview respondents during the interview.

## **Task 5 | Pilot test**

Lastly, a pilot test of the new online data collection tool is scheduled for late-summer 2014. The purpose of the pilot will be to test the functionality of the new tool with 40–50 LEAs. The pilot period will last up to 3 weeks, during which time LEAs will respond to the survey. Pilot LEAs will be contacted to provide feedback and suggestions. A feedback mechanism that captures screenshots of user’s computer screeners will be implemented. At least two pilot LEAs will be designated to test the flat file upload. AIR and Sanametrix will evaluate the feedback and make recommendations to NCES/OCR for changes that are feasible to implement prior to the opening of the 2013–14 data collection period.

# Recommendations

## Goal 1 2 3 4

### Reduce reporting burden

#### Recommendations for reducing burden

**Challenge: Data elements gathered and stored through decentralized systems represent a large share of the reporting burden**

Recommendations:

Design the data collection tool to better align with the way LEAs collect, store, and report CRDC data

Evaluate the utility of difficult-to-report data elements relative to their reporting burden

Establish a task force or task forces to share promising practices for collecting data not typically housed in centralized systems

Engage the vendor community in developing tools to better support LEAs in responding to the CRDC

**Challenge: Expanding and changing data**

Reports from the site visits suggest that the amount of time and effort it takes to respond to the CRDC represents a significant number of staff hours allocated to gathering, reporting, and validating the CRDC data, making the process very burdensome. Exhibit 1 shows examples of time burden as reported by small, mid-size, and large districts that were able to provide this information. Burden as reported by site visit respondents varied across LEAs and LEA sizes and ranged from about 1 week to 6 months. Keep in mind that these examples are just for reporting activities during the reporting period. Many LEAs spend additional time planning, such as one small LEA whose CRDC contact trains all school registrars on the CRDC and leads a data group that meets every month for 2 hours.

[Format instruction - INSERT Exhibit 1 on following page (exhibits in the draft appear at the end of this section)]

Many of the recommendations for improved communications tools and data collection tools will also help reduce reporting burden, but there are additional strategic improvements that can be made to reduce the amount of time and effort schools and districts spend on the CRDC.

**Challenge: Data elements gathered and stored through decentralized systems represent a large share of the reporting burden**

Every LEA site maintained and utilized a student information system (SIS) to report at least some portion of the student count and program data for the CRDC. The primary purpose of these SIS systems was to meet reporting requirements of the SEA- or LEA-driven policies or practices. A secondary benefit was that data stored in the SIS were the easiest for LEAs to report accurately for the CRDC. CRDC data stored in the SIS system were all coded in a consistent manner, aligned with student demographic data to support most of the disaggregation required by the CRDC, and tied to a uniform set of school codes. For information stored in the SIS, LEAs typically run queries to aggregate and report the



elements inhibit LEAs from improving practices for gathering and reporting CRDC data

Recommendations:

Create a consistent, core set of CRDC items

Develop flexible special modules to explore policy directives

Evaluate utility of items not regularly used by OCR offices

**Challenge: LEAs are duplicating the reporting of data elements to their SEAs and OCRs**

Recommendations:

Engage SEAs in the reporting process to reduce district burden

Under take a strategic CRDC take the most time about 50% of response time is spent on 5% of the data, for example, AP results - Small LEA

**Challenge: LEAs perceive that ED is not using all its resources to lower LEA burden**

Recommendations:

Evaluate the best match between all ED surveys and non-core data elements

Populate available

information in the format that is needed for the CRDC. The SIS output is either submitted as a flat file or given to a clerk or administrator to enter into the online data tool. An actual screenshot example of an SIS query for Algebra I passing is shown below.

**CRDC - Alg 1 Passing Total Records:7102**

All Records						
Grade 09						
Gender F						
Race/Eth 1						
Gender	Race/Eth	Course	Grade	SPED	Couse	Grade
F	1	10104S1	09	0	20102S1	A
F	1	10104S1	09	0	20102S1	B+
F	1	10104S1	09	0	20102S1	D

The majority of data collected by the CRDC focuses on schools and their students. However, the CRDC also collects data on school staff and school expenditures, which are typically not stored in the SIS. All LEAs regardless of size had centralized staff and business information systems, and most of the information required for the CRDC was housed in these systems. Some examples of these systems were Alio, Quintessential, and PeopleSoft. The main causes of reporting difficulty were the use of contractor staff and having to report information at the school level that is only provided at the LEA level. To complete the CRDC submission, LEAs request data from the school HR or business departments, sometimes providing spreadsheets or templates for departments to complete. Once data has been returned to the LEA POC, the POC either provides this data to a school or submits the data on a school's behalf. While this collection process did not vary greatly between LEAs, districts did vary in their procedures for certifying this information. While in some cases the district's HR department was responsible for certifying the data, in other districts this responsibility fell to school principals or superintendents.

A number of SIS systems used by LEAs had vendor-supported tools that supported CRDC reporting. For example, one LEA mentioned that their vendor allows for customization of screens and fields, which allows the district to request specific data from schools. The vendor also works directly with the SEA and LEAs to provide the fields necessary for reporting, as shown in the screenshot below. The same vendor now has a CRDC component specifically designed to assist schools with CRDC reporting. The program will produce a completed CRDC survey in a PDF file that can then be manually entered into the CRDC website.

and matching data from ED Facts

**Challenge: Timing of data collection does not align well with LEA schedules**

**Recommendation:**  
Consider allowing LEAs to opt for two different submission models

Sometimes we'll have to contact school administrators regarding text in a behavior incident description to fully understand what transpired to know how to code the incident for CRDC – Mid-size LEA on difficulty reporting CRDC discipline data

**PowerSchool**

**Civil Rights Data Collection**

LEA-0001. LEA-Level Count for Schools and Students  
A. What is the count of students awaiting initial evaluation for special education programs and related services under IDEA?

LEA-0002. Is this LEA covered by a desegregation order or plan?

LEA-0003. Kindergarten Daily Length  
A. Does this LEA provide full-day kindergarten?   
B. Does this LEA provide part-day kindergarten?

However, there were cases of data elements included in the SIS system where reporting challenges still existed. The most frequent reporting challenge was unclear data definitions and CRDC definitions that differed from the definitions for SEA requirements. In these instances, respondents needed to create special queries and programs, recode data to fit the CRDC definitions, or go back to comments on school data-entry forms to ensure alignment with the CRDC definitions. This process added several hours to reviewing, recoding, and checking the SIS data element. An example of manual recoding of harassment incidents from one site is shown in the screenshot below. Specific definitional problems are detailed in a separate cognitive interview report.

F	F	White	W	N	N	N	N	N	N	Other	100%
M	F	African Americ	N	N	N	N	N	N	N	Other	Sexual harassment?
M	F	Asian	N	N	N	N	N	N	N	Other	Sexual harassment?
F	F	Multi-Racial	N	Y	N	N	Y	N	N	Other	not enough info
M	M	African Americ	N	Y	N	N	Y	Y	N	Other	not enough info
F	F	White	W	N	N	N	N	Y	N	Other	intimidation

Data not stored in the SIS or business and staff information systems, or data that are in these systems but not defined correctly for the CRDC, are the most difficult to report and sometimes even prevent LEAs from reporting at all.<sup>2</sup> These difficult data elements make up the majority of the excessive burden in both the data collection and validation process. **One LEA estimated that about 50 percent of their CRDC reporting time is spent on about 5 percent of the data.** The types of data that are difficult to report are presented in Exhibit 2, along with site visit respondents’ descriptions of difficulties for districts of different sizes.

[Format instruction - INSERT Exhibit 2 on following page]

In many LEAs, different departments demonstrated different processes of gathering and inputting data into centralized systems (e.g., hard copy

<sup>2</sup> In previous CRDC submissions, LEAs could request an exemption for reporting data that they did not collect. Additionally, site visit respondents mentioned that they had reported zeros in the absence of complete or accurate data to report.

forms entered by a school secretary, data entry screens that were custom designed by district IT departments) and ad hoc processes of gathering data not stored in the SIS system (e.g., Excel spreadsheets maintained by guidance counselors on AP exams; rosters of interscholastic athletics participants maintained by coaches; and Harassment, Intimidation, and Bullying (HIB) coordinators maintaining incident forms).

As these ad hoc responses to data requests accumulate, a decentralization of data responsibility, described by one mid-size LEA as “data silos,” results. Data silos were more common in mid-size and small LEAs, but were also present in large and very large LEAs. If decentralized LEA data can be viewed as a data silo, centralized LEA data, like an SIS or finance system, can be viewed as a data “barn.” Data silos typically contain data that are not required to be reported to the SEA and/or contain data elements that are managed at the school level, rather than at the LEA level (school athletics, for example). Exhibit 3 is an illustration of the data barn/silo structure.

[Format instruction - INSERT Exhibit 3 on following page]

The challenges to the CRDC data collection resulting from data being collected in silos are (i) there is no consistent system for coding data for aggregation, (ii) there is no link to student records for breakdowns by race or other CRDC categories, and (iii) in some cases, the raw information (e.g. narrative text data) is never coded at all. Data are sometimes collected using a content-specific commercial software vendor, sometimes collected using an in-house software system, sometimes collected using Excel spreadsheets, and sometimes collected using handwritten forms—and these methods can also vary by school, even in very large districts. With harassment or bullying data, it is often the case that a discipline incident form is filled out by the school personnel reporting it, in narrative format, and this might be the extent of the documentation. A redacted screenshot of an actual example narrative is shown below. When the CRDC requests aggregated data about bullying and harassment in this scenario, the HIB coordinator requests information from the school. The school staff or the HIB coordinator must read all of the narratives, incident by incident, school by school; code each incident; and then aggregate the incidents to meet the CRDC request.

This submission is a huge burden. It is our largest non-funded mandate – Very large LEA

## Statement

Victim: [REDACTED]

Date Received: 4/20/201

2013 2:13 PM: [REDACTED] s stating that" he is being Bullied by [REDACTED] approached [REDACTED] and started to inform him of the current situation brewing between himself and [REDACTED] and asked for his assistance. He told [REDACTED] that he waited this long to come to him to discuss this issue because he was very embarrassed and just wanted to protect himself, but felt he was now very afraid of [REDACTED] and [REDACTED] s bullying. Please read the attachment of [REDACTED] statement of events

[REDACTED] was in gym class in the basement where the ping-pong table was housed. He was playing a game with a classmate during gym (another 9<sup>th</sup> grader). After the game was completed they started fooling around by pushing and bumping into each other. [REDACTED] fell into the table and it collapsed. It was at that point in time when [REDACTED] <sup>teacher</sup> walked into the classroom and both students admitted to breaking the table.

[REDACTED] then walked into the room and asked what had happened. A few of the students in the room told him that [REDACTED] had broken the table. [REDACTED] then walked over to [REDACTED] and said he had to pay for breaking the ping-pong table or he was going to "beat him up and break him". [REDACTED] also threatened to give [REDACTED] a swirly and attempted to push [REDACTED] into the boy's room. [REDACTED] engaged in a bunch of bullying and intimidation unknowingly.

For LEAs where data are stored electronically in decentralized databases, it is a bit easier to merge these data with the data in their SIS, but it is also time consuming. For example, one LEA explained that to complete the CRDC they need to combine data from three other databases, but each database has a different way of identifying their schools (e.g., it's a three-digit numerical code in one database, the name of the school in another, and a state code in another). Some LEAs reported that they developed SQL code to merge the different databases together to reduce their burden, but others explained they did not have the resources (financial or staff capacity) to do this, so they combined data using Excel spreadsheets and pivot tables.

Successful CRDC responders are those who have fewer data silos or who integrate CRDC information that comes from data silos into the SIS or centralized staffing/finance systems.

### **Recommendation: Design the data collection tool to better align with the way LEAs collect, store, and report CRDC data**

Align the design of the CRDC tool to the way CRDC data elements are gathered, stored, and used within an LEA and support the standardization of the collection of data elements gathered through ad hoc processes. **Allow for multiple respondents** in the online data collection tool and group data elements into **topical modules** to make it easier for LEAs to assign out data elements to the appropriate staff or centralized databases. This recommendation is discussed under Goal 3, Recommendation: Allow for multiple users and permissions

### **Recommendation: Evaluate the utility of difficult-to-report data elements relative to their reporting burden**

If the CRDC definitions don't align with state data collections, we don't feel confident in the data submission – SEA respondent

We recommend that, where possible, NCES/OCR evaluate the utility of difficult data elements and eliminate them from the collection where the burden outweighs the utility. Difficult data elements include those that are housed outside of the SIS, data elements with definitions that do not align with those of the LEA, data elements that are not required for reporting by the SEA, and data elements that require extensive disaggregation.

**Recommendation: Establish a task force or task forces to share promising practices for collecting data not typically housed in centralized systems**

For those elements deemed essential, we recommend that NCES/OCR **establish a task force or task forces** to develop a plan for assisting LEAs in collecting, managing, and reporting data that are difficult to obtain in order to share promising practices and suggest ways to standardize the collection of data. This concept was suggested by one LEA who was reporting school expenditure data. The suggestion was to organize a workgroup made up of school finance staff to help understand the complexities of these data, and clarify finance-related data elements (e.g., develop better definitions) to reduce the burden on LEAs. This concept could be expanded to not only understanding the required data, but also to understanding ways to better collect it, as well as ways of achieving agreement on solutions or helpful tools that NCES/OCR could consider developing.

**Recommendation: Engage the vendor community in developing tools to better support LEAs in responding to the CRDC**

In some cases, the difficult data are obtained from outside of the LEA. For Advanced Placement (AP) exam data, some sites suggested NCES/OCR should specifically **work with the College Board** to develop templates or file specifications for delivery of AP and other data from the College Board to LEAs/SEAs. Additionally, other sites suggested that NCES/OCR actively engage with the vendor community to develop better tools to gather and report CRDC data. A software tool for tracking bullying/harassment incidents, called HIBSTER, was used by one LEA, but this tool did not offer all of the reporting formats needed for the CRDC. NCES/OCR could consider making a list of potential data providers for these difficult data elements and **communicate CRDC needs to those vendors.**

**Challenge: Expanding and changing data elements inhibit LEAs from improving practices for gathering and reporting CRDC data**

In almost all of the site visits, we heard that changes and additions to the

CRDC exponentially increase the amount of time sites spend on the data collection. One LEA explained that the process to create the SQL queries was very time consuming, but it was done to reduce burden in providing data for future CRDC submissions—however, if the data requirements change from submission to submission, it takes a long time to update/recreate their SQL code and this increases their burden. Additionally, if new items added to the CRDC collection are not in the SIS, there is increased burden for the LEA to report these data elements.

**Recommendation: Create a consistent, core set of CRDC items**

We recommend that NCES/OCR identify consistent, core items from the 2013–14 and 2015–16 collections. This approach will maintain consistency in the collection for the next two cycles, which will ensure that LEAs and SEAs do not need to modify existing programs and plans for meeting the current core content. By **committing to a set of consistent, core items** and **core disaggregation** (ideally aligned with ED*Facts* race/ethnicity disaggregations) that form the main body of the collection over time, a level of stability to the data collection will be apparent to LEAs and allow them to develop long-term plans and increase the utility of adding core items to centralized data system.

We further recommend that NCES/OCR agree to strict rules and approvals for modifying core content, followed by a lengthy minimum development time and cognitive research or pilot testing to support new item development.

**Recommendation: Develop flexible special modules to explore policy directives**

The ability of the CRDC to respond to emerging policy directives affecting civil rights is also important. To address the competing concerns of responsiveness and stability, we recommend that NCES/OCR develop a timeline and plan to create **flexible special modules** for new data demands.

Special modules can be handled separately from the core items and do not need to be repeated in every collection. They should also follow a guiding set of rules and approvals to ensure that only required and relevant content becomes a module; and they should undergo pre-implementation testing.

**Recommendation: Evaluate utility of items not regularly used by OCR offices**

Another means for reducing items is to **eliminate data elements that are not used regularly by OCR offices**. We visited two OCR offices



serving several states. The OCR respondents identified the most useful and least useful data elements. Both offices noted that their priority data reporting is dependent on the legal compliance review issues that are occurring and other compliance directives from the OCR main office.

The most useful CRDC data elements from the last collection, as identified between the two OCRs were

- School and district characteristics
- Harassment and some discipline data linked to harassment; in-school and out-of-school suspensions; enrollment linked to discipline; and number of law enforcement officials
- Restraint and seclusion
- Gifted and Talented
- Algebra 1—Data are more useful when broken out by seventh and eighth grade separately.
- AP data—It is useful to see whether an AP course is available to students.
- Special education data
- General enrollment demographic information—General enrollment data is widely used, but it is not ideal that the data are 2 years old.

The least useful CRDC data elements were

- Number of school days missed for out-of-school suspensions is only helpful if they know how many school days there are in a particular district.
- Preschool data—One OCR does not use preschool data because it is not required in the OCR office’s states.
- Interscholastic athletics—Data have not been used recently because there have been no complaints, and such data have not been a focus for the OCR administration.
- HR data—One OCR does not have an understanding of what the data could be useful for or what it means, because they have not yet had a case that needed FTE information, and they have not yet gotten into issues of teacher equity.
- School finance—These data can get confusing (e.g., some items are state and local, while others are federal, state, and local) and have not been useful recently.
- Distance education courses—There is confusion about interpretation of this data element in regard to how students in these courses are counted. For example, in some states, a student can leave an LEA to do a distance education course, and some special education students spend time in distance education and in a physical school.
- Discipline as disaggregated by disability. Prefer aggregation
- Suspension data that is reported by number of students and not

The reporting burden is high. If the SEA gets involved it just shifts [the] burden to them but does not reduce [the] burden overall - SEA respondent

number of suspensions; and number of days missed due to suspension, because it is only useful if they know the total number of school days in a district

We anticipate that additional recommendations about specific data elements to consider for deletion will be generated from the cognitive interview research.

**Challenge: LEAs are duplicating the reporting of data elements to their SEAs and OCRs**

A common point from LEAs was that CRDC data had been already reported to SEAs for other purposes.

SEAs expressed an interest in supporting LEAs in their reporting of data to the CRDC, but said that additional communication would be needed between the OCR, SEA, and LEAs to be successful. Because this data collection is not coordinated at the SEA level, SEAs need to be included in communications between LEAs and CRDC so that they are aware of deadlines, what has changed, and who is responsible at the LEA. The current lack of knowledge hinders SEAs from providing good data support, prevents SEAs from knowing whom to contact, and inhibits SEAs from answering relevant LEA questions.

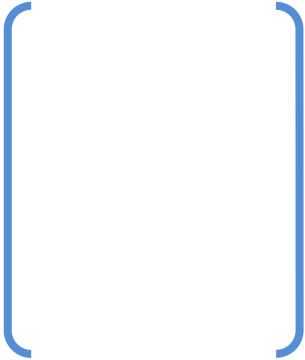
Several SEAs expressed a desire to be more active in providing their LEAs with data. One SEA specifically noted that there was a significant overlap between CRDC data and data reported to the state that could potentially be returned to the LEA for CRDC submission. However, the LEA has not had the staff or financial resources to put a system in place. Another SEA mentioned that receiving the specifications for the data collection well in advance would better allow the SEA to plan the process of feeding the data back to the LEAs.

Some SEAs already report on behalf of their LEAs (e.g., Florida) or offer complete or partial data files to their LEAs (e.g., Iowa, Wisconsin, Kansas) for submission to CRDC. Some of the SEAs we spoke to indicated that anywhere from 25 percent to 60 percent of CRDC data for their states could be prepopulated using existing databases. The actual percentages cited were “25%,” “40%,” “40–50%,” “60%,” “60%,” “60,” and “a large percent.” Other SEAs have indicated that they could provide data to their LEAs for some, but not all, CRDC data elements.

**Recommendation: Engage SEAs in the reporting process to reduce district burden.**

States that have engaged in prepopulating the CRDC from existing data





do so because they want to reduce burden on the LEAs and increase the quality of the data collection. For example, during one SEA interview, we found out that the majority of LEAs in the state serve less than 1,000 students and one individual does all of the administrative work. This means that the CRDC data collection falls on this individual, so the state has stepped in to help. In states that provide data, the districts are able to review and revise the data they are provided before final submission to CRDC. In the states in which the prepopulation of data takes place, we found that one or several individuals are deeply involved in the CRDC task force group and receive support from the SEA by providing staff time and other resources to the CRDC task. SEA assistance with the CRDC submission also helps to ensure alignment between the data that the LEA reports to the OCR and the data that the LEA reports to the SEA, which provides a foundation for a consistent set of publicly reported data across federal, state, and local governments.

**Recommendation: Undertake a strategic state-by-state campaign to engage SEAs**

We recommend that NCES/OCR undertake a strategic state-by-state campaign to engage SEAs in data reporting for the CRDC. In all of the states we spoke to who are assisting LEAs, there was considerable planning involved, and we would expect this to be similar for any new SEA assistance program. In each case there was more than one person involved at the SEA. While some individuals fulfilled multiple roles, SEA teams were typically composed of a “policy champion” who advocated to reduce burden on LEAs across the state, context experts responsible for mapping SEA data elements to CRDC definitions, and technical experts who created files for use by LEAs. NCES/OCR should be prepared to offer as much assistance as possible. Suggested steps for engaging SEAs are provided in the box below.

It is unlikely that all SEAs will be willing to provide the same amount of support and involvement. However, any SEA involvement can help reduce burden, so SEAs should be encouraged to provide whatever level of support they are willing to give. In anticipation of this variation in degrees of help that SEAs will provide, we recommend that NCES/OCR develop suggested “levels” of SEA involvement. For example, “Level 1” SEA involvement may simply be assisting with LEA contact and communication for the CRDC, whereas “Level 3” may be assisting with communication, data preparation, and providing a “help desk” during the CRDC reporting period.

**Challenge: LEAs perceive that ED is not using all its resources to lower burden**

Similar to the use of data submitted to SEAs for other reporting

purposes, both SEAs and LEAs argued that the CRDC could be partially prepopulated with data that SEAs submit to NCES for other reporting purposes, mainly the ED*Facts* data collection. SEAs and LEAs do, however, realize that the data submitted to ED*Facts* does not always align in either definition and/or disaggregation with the data elements in CRDC. They wish for this to change.

**Recommendation: Evaluate the best match between all ED surveys and non-core data elements**

Clarifying the rationale, purpose, and importance of the data elements may address the concerns about item redundancy in federal and state reporting. These rationales for inclusion on the CRDC could be linked from the online tool, similar to definitions. Enrollment data was often mentioned as redundant across sources.

**Recommendation: Prepopulate available and matching data from ED*Facts* and better align CRDC to ED*Facts*:**

- Directory information for LEAs and schools (this includes LEAID/NCESSCH, name, address, phone number)
- School type information (alternative school, charter school)
- Membership data by grade, by sex, by race/ethnicity

As NCES/OCR plans to address burden reduction by developing core content, SEAs and LEAs have pleaded for these data elements to align with ED*Facts*. This was specifically requested by nine sites and was documented in the summary of known issues. SEAs and LEAs would like to see entities within the Department of Education work together to align definitions of data elements across data collections and align levels of aggregation/disaggregation of data element counts. Alignment among data collections within the Department on data definitions and levels of aggregation would decrease the overall federal reporting burden on LEAs and SEAs by **reducing the number of methods by which like data items have to be enumerated.**

A few sites listed specific CRDC data elements that they feel should align better with ED*Facts*; these are listed below. Overall, however, respondents wanted ALL information that is similar across CRDC and ED*Facts* to be aligned, namely,

- Enrollment data
- Financial data
- Number of incidents that occurred at this school, robbery with weapon, robbery with firearm
- Injury data

**Challenge: Timing of data collection does not align well with LEA**

Fall is very bad. We are way too busy on state requirements to do anything for the federal government – Mid-size LEA on timing of the CRDC

## schedules

Many sites provided feedback on the timing of the CRDC survey and its effect on reporting burden and data availability – **there was no consensus on the best timing for the CRDC data collection period.** LEAs’ primary data reporting activities center around SEA requirements; reporting requirements vary by state in terms of the number of times during the year data need to be submitted and when the data need to be submitted. However, there was wide **agreement that staff did not have time to work on the CRDC collection at the beginning of the school year**, when LEAs are usually busy preparing SEA data submissions.

The summary of known issues also notes that some LEAs do not have a good understanding of when the data collection takes place and are concerned about the single deadline for submission, because not all CRDC data, such as school expenditures, may be finalized in time.

### September–October (beginning of school year)

There was general consensus that the beginning of the school year is a bad time to initiate the CRDC data collection. LEAs and SEAs are extremely busy at the beginning of the school year.

### November–December

One small and one large LEA expressed that November and December were preferred months for CRDC data reporting. Both indicated that fall and spring/summer are not good times for data reporting, because these times conflict with other reporting requirements for their respective states. Another small district expressed that December is a problem because they are busy responding to data requests for *EDFacts* at this time.

### January–February (winter) or spring

One mid-size district expressed a preference for a winter (explicitly citing the months of January and February) or a spring data collection. The preference for a spring data collection conflicts with the preferences of the two districts above who requested data collection in November/December.

### Summer

Another mid-size LEA expressed a preference for conducting the CRDC collection in August.

### Multiple due dates (*EDFacts* model)

One very large LEA suggested that NCES/OCR adopt multiple due dates for the CRDC collection—the survey could be divided into conceptual groups, and each grouping of data could have its own due date. This model is similar to what is done for the *EDFacts* data

collection.

**Recommendation: Consider allowing LEAs to opt for two different submission models**

Using multiple due dates for various sections of the survey could diminish burden to those SEAs and LEAs that currently have difficulty reporting data during the current CRDC collection period; conversely, it might add burden to SEAs and LEAs for whom the current collection cycle/calendar aligns well with their needs. We recommend that NCES/OCR consider allowing LEAs to opt for two different submission models—a single-date model and a multiple-date model.

**Main challenges for implementing Goal 1**

Maintaining consistency while making changes to the content and design of the CRDC are conflicting recommendations for reducing reporting burden. To tackle this problem, we recommend that the pace of change for the CRDC be slow and deliberate, particularly for the content. We also recommend that NCES/OCR inform LEAs and SEAs of the forthcoming plans for change and when the change will happen.

**Exhibit 1 | Examples of time burden for CRDC reporting**

Site	Time burden description	Summary of CRDC reporting activities
Small LEA	It takes a few weeks to compile all of the data.	<ul style="list-style-type: none"> <li>-one person at district level who receives, inputs, and submits all CRDC data</li> <li>-extracts data from six databases</li> <li>-state not involved in submission</li> <li>-majority of data stored in system, just need to extract it.</li> <li>No timeline for data entry</li> <li>-school personnel enter data into “master bridge” system and then POC extracts for CRDC. Manually enters into CRDC</li> </ul>
Small LEA	It takes approximately 30 minutes per school and a total of 5 hours to enter data for all the schools in their district into the CRDC website. Small parts of the CRDC take the most time, 50% of time for CRDC is spent on 5% of the data needed for CRDC. For example, AP results take a long	<ul style="list-style-type: none"> <li>-SEA involved in submission</li> <li>-data constantly being collected in SIS</li> <li>-runs canned queries to get CRDC numbers</li> <li>-numbers are manually entered into the site</li> </ul>

	time.	
Small LEA	It takes 52 hours.	<ul style="list-style-type: none"> <li>-uses SEA data system</li> <li>-SEA does not collect data for CRDC</li> <li>-LEAs aggregate data from paper reports</li> <li>-data put into spreadsheets, pivot tables are run to create aggregate counts</li> </ul>
Small LEA	Elementary schools: about 2 days—it’s easier and shorter, there are fewer questions. Middle schools take longer than elementary schools. High schools take longer than middle schools.	<ul style="list-style-type: none"> <li>-state not currently involved in submission</li> <li>-LEA data team (school registrars) retrieve data and give it to POC, who manually enters it (they meet for 2 hours every month, have trainings on definitions, data elements)</li> <li>-data team uses SIS to run queries</li> <li>-bullying/harassment data come from counselors</li> <li>-after about 2 weeks, registrars, HR, and finance provide POC with output and she enters manually into CRDC</li> <li>-downloads and gives to schools to review, and then makes edits</li> </ul>
Mid-size LEA	It requires 40 or more hours to collect all the data and compile it.	<ul style="list-style-type: none"> <li>-SEA provides 60% of data</li> <li>-one POC, schools not involved</li> <li>-school personnel enter data into SIS all year long</li> <li>-requests data from HR/finance departments separately, plus incident reports/harassment/ bullying data</li> <li>-merges data from SIS, HR, finances, SEA</li> <li>-uploads them to CRDC</li> </ul>
Mid-size LEA	About a week.	<ul style="list-style-type: none"> <li>-SEA not involved</li> <li>-runs queries from SIS and then consults subject experts to fill in remaining data</li> <li>-schools collect data on daily basis, stored in SIS</li> </ul>

		-data then keyed into survey tool
Mid-size LEA	It took about 3 weeks.	-SEA involved -school-level data submitted to LEA on regular basis -data for state collected and then converted to meet CRDC -experts consulted for data not included in the SIS (HR, athletic data, AP, LEP, vocational)
Large LEA	Creating the flat file takes approximately 2 weeks.	-single person -SEA not involved -school-level personnel enter data in SIS on an ongoing basis -POC merges data using previous queries to produce single flat file for submission -no special time frame, no interaction with schools
Very large LEA	It takes 6 months to do the CRDC.	-state does not currently provide any data to the LEAs -only LEAs are involved in submission -schools enter data into SIS all year long and then LEA contacts schools with questions if needed -coding system not detailed enough, narratives and not codes for incident report, -four different databases -schools report discipline differently and the LEA has to align the data

Source: CRDC Improvement Project Site Visits, 2014.

**Exhibit 2 | List of data elements reported as difficult and reasons for difficulty, by LEA size**

Data element	Site visit respondent description of difficulty, by LEA size
Athletics data	<p><b>Large</b> LEA: Athletics directors or coaches provide information. Rosters of students must be coded and aggregated to meet the CRDC requirements. Rosters often do not include a student’s race, ethnicity, gender, or student IDs, so the LEA must try to match the student names with the information in the SIS, which is very time consuming. School: Collected in hard copy; coaches for this school report to state, not district, so numbers are not accurate at district level</p> <p><b>Mid-size</b> LEA: They rely on coaches’ lists of individuals who play for them and it is difficult to get accurate data</p> <p><b>Mid-size</b> LEA: Asks schools for their athletic data; not good at keeping it in SIS</p> <p><b>Small</b> School: Athletics supervisor keeps athletics data and is it not in SIS</p> <p><b>Mid-size</b> LEA: Schools put data into SIS, but unsure how accurate</p> <p><b>Small</b> LEA: Coaches respond to data requests for athletics by hand</p>
Financial data	<p><b>Large</b> LEA: Often need more clarification of who falls into which category. Finance data come from own data source and are pulled by a colleague (senior coordinator for finance) and put in a spreadsheet. The data center manager then pulls the information needed and manipulates the spreadsheet for CRDC</p> <p><b>Small</b> LEA: Central office staff aren’t tracked by amount of time spent at each school LEA: Finance staff report that their system is not set up to track information needed for CRDC</p> <p><b>Mid-size</b> LEA: Expenditure at the LEA level might not accurately show what the dollars were spent on by school</p> <p><b>Very large</b> LEA: Finance department provides these data in PDF format, and it is time-consuming to enter such data into the CRDC school by school</p> <p><b>Small</b> LEA: Need definitions distinguishing support staff from administrative staff and noninstructional support staff</p> <p><b>Very large</b> LEA: Need clarity about whether preschool personnel salaries are disaggregated by funding source; question of whether salaries are needed for particular point in time or for teachers who were present for</p>

	<p>whole period of collection time; and whether CRDC is seeking budget amount or actual expenditure for money questions</p> <p>LEA: The district doesn't know teacher's overall experience in teaching, just the amount of time teaching in district</p>
Security guard, school resource, law enforcement officer	<p><b>Mid-size</b></p> <p>LEA: Security guards, school resource officers, and law enforcement are difficult to track because they contract for services, and it is hard to know who is working at a given time</p> <p><b>Mid-size</b></p> <p>LEA: Sworn law enforcement officer is not tracked in student information system—it would need to be hand-counted</p> <p><b>Small</b></p> <p>LEA: Have juvenile corrections officers who may serve an overlapping or similar function</p> <p><b>Large</b></p> <p>LEA: Often use city police and do not track information about city police</p>
Nonpersonnel expenditure	<p><b>Large</b></p> <p>School: District does not have accurate numbers because federal and state grant funds go directly to her school, not through district</p> <p><b>Small</b></p> <p>LEA: Need clear definition of what is included in nonpersonnel expenditure</p> <p>LEA: Schools are on different schedules for equipment replacement, so it may look like there is inequity within a single year</p>
Dual enrollment in courses	<p><b>Large</b></p> <p>LEA: Not in system as requested by CRDC, but can be calculated</p> <p>School: Information could be found in the system but system may not have separate codes for these; hard to capture</p>
Credit recovery	<p><b>Large</b></p> <p>School: In system as a description but not a separate code, so it may be hard to capture</p> <p><b>Small</b></p> <p>LEA: Need to be highly defined; students doing credit recovery are in a class with other students who are taking it for the first time, so it is hard to measure</p> <p><b>Small</b></p> <p>LEA: They do not report on this element because of the timing of classes</p> <p><b>Mid-size</b></p> <p>LEA: Need to be sure of the definition to see whether their programs qualify</p>
Advanced Placement (AP) exam data	<p>Many LEAs of all sizes expressed difficulty in reporting AP data because the LEAs do not receive the data from the College Board in a format that facilitates reporting to the CRDC (e.g., the AP data do not</p>



	<p>include enough identifying information to match with student SIS records).</p> <p><b>Large</b> LEA: Counting who took exams is hard; probably the most difficult aspect of all; a challenge, because it depends on whether you get the information from the College Board or whether someone cleans it up and matches it to our student IDs School: Data may not exist for all schools; comes from College Board; knowing whether students are allowed to self-select would be difficult to determine</p> <p><b>Small</b> LEA: Data are collected in SIS, but must be verified by high school guidance counselor</p> <p><b>Small</b> LEA: Must be obtained from guidance counselor in spreadsheet separately</p> <p><b>Mid-size</b> LEA: College board does not use student ID number, so it must be matched to student data by student name, which is cumbersome; College Board provides the data as a PDF and data must be manually entered into Excel files by LEA, and then pivot tables are run for aggregations</p> <p><b>Very large</b> LEA: Data exist in other offices and require a matching process; LEA has no control over the accuracy of the data</p> <p><b>Small</b> LEA: Passing APs is not something the LEA tracks; what constitutes passing?</p>
SAT and ACT	<p><b>Small</b> LEA: College Board data results don't include student ID numbers so can't easily be entered into system</p> <p><b>Mid-size</b> LEA: Data not in SIS; data are gathered by school guidance counselor and sent to CRDC POC</p> <p><b>Small</b> LEA: Students sign up for SAT/ACT independently, and the data do not go through the schools</p>
Number of students absent for more than 15 days	<p><b>Large</b> LEA: Is it for entire school year, or just within specific school student attended at end of year?</p> <p><b>Small</b> LEA: Does "absent" include suspensions, field trips, etc.? How do you define "all day absent" (number of classes missed? the missing of a bridge period?); different definitions of "full-day absent" are used by different schools</p>

	<p><b>Mid-size</b> LEA: Clear definition needed</p> <p><b>Very large</b> LEA: Students may be absent for more than 15 days, but days spread across multiple schools attended—this would require calculation</p>
Referral to law enforcement agency/official	<p><b>Large</b> School: Not sure how it data element entered/coded in system if it is not an arrest (not a black-and-white element)</p> <p><b>Mid-size</b> LEA: They do not track these data</p>
Arrests	<p><b>Mid-size</b> LEA: Difficult to provide because data element must be requested from schools. It is so rare that it is not thought to be reported in any certain way</p> <p><b>Mid-size</b> LEA: Don't keep track of arrests for school-related activity specifically—they would need to look through each suspension to find out if it resulted in an arrest</p>
Bullying/harassment and other discipline data	<p>Many LEAs of all sizes reported difficulty in reporting discipline data for the CRDC. In most cases, there was no standard system for coding incidents that match the CRDC definitions. The process of reporting therefore required reviewing incident descriptions for each incident, and then coding them</p> <p><b>Large</b> LEA: Need to clearly identify who falls into which category LEA: Need more description for types of incidents (e.g., does firearm refer to hand gun, rifle, etc.?) School: Unsure whether there is a code for these details in the system</p> <p><b>Small</b> LEA: Collect data on the person who bullies or harasses, but do not collect data on the victim or basis for bullying; data collected is not as specific as CRDC, e.g., they track threats of physical attack but do not specify if threats include weapon/firearm/explosive device</p> <p><b>Mid-size</b> LEA: Must read through specific incident comments and then categorize incidents; no category for rape or sexual battery</p> <p><b>Small</b> LEA: Time consuming; level of detail needed for reporting for CRDC isn't in system; must obtain from the schools and create separate database for the data</p> <p><b>Mid-size</b> LEA: Collect the listed types of incidents but not in the same language LEA: Challenging to pull disciplinary data, because the disciplinary file is huge and requires a lot of data manipulation work; based on the ways schools enter infractions and comments</p>



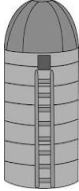






	<p><b>Mid-size</b> LEA: There are no business rules for standardizing entry and coding of data; LEA must disaggregate the reports that are sent to SEA in order to calculate incidents by student characteristics; CRDC report is much more specific than state reports; discipline data is primarily free entry, making it difficult to search</p> <p><b>Mid-size</b> LEA: Specificity is difficult to report because state does not require level of specificity that CRDC requires; must consult paper records</p> <p><b>Very large</b> LEA: Some schools do not want to collect victim data because they don't want the information following the student; CRDC wants number of victims and offenders to match, but that doesn't always happen</p> <p>LEA: Incidents may have more than one event and student involved; some of the specificity is included in narratives or not at all; this data requires mapping current codes to the CRDC categories</p> <p><b>Mid-size</b> LEA: Do not have codes that split documented incidents as specifically as CRDC requests</p> <p><b>Very large</b> LEA: Some incidents may be reported for a building as opposed to a school (some buildings house more than one school)</p> <p><b>Small</b> LEA: There is no reliable database for answers to CRDC data elements—students are usually reported for fighting or assault, not harassment/bullying</p> <p><b>Very large</b> LEA: Schools report incidents but not as specifically as is needed for CRDC and other collections</p> <p><b>Small</b> LEA: Allegations are not in student system; reporting of allegations depends on whether someone gave information to counselors/principals. At the LEA's charter school, the registrar pulls hard-copy data to report bullying and harassment</p>
Course information	<p><b>Mid-size</b> LEA: Combining students who have taken more than one of the requested courses is difficult</p> <p>LEA: There is no clear definition of failure, so providing information on students who passed/failed certain courses is difficult</p> <p><b>Large</b> LEA: Need to clearly identify who falls into which category</p> <p><b>Small</b> LEA: Want to submit all class info in a table and have CRDC manipulate it the way it is needed</p> <p><b>Mid-size</b> LEA: There is no universal definition of “passing”—it is defined at the</p>

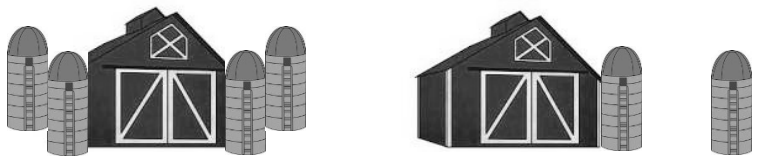
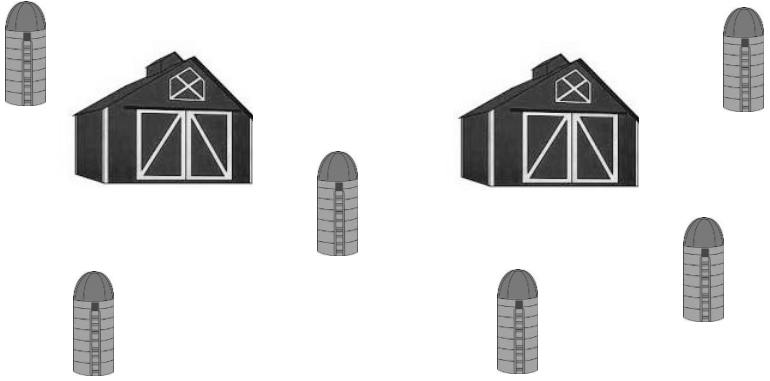
	<p>local level and data may not be comparable across LEAs; mobility can cause number of students passing a class to be greater than number of students enrolled; NCES needs to provide info on how to define classes (list of SCED codes would be helpful)</p> <p><b>Mid-size</b> LEA: Number of students taking a certain class/teacher caseload size is a moving target, so it is not accurate from point A to point B; need clear definition of classes that qualify as algebra I, etc.</p> <p><b>Very large</b> LEA: Schools have somewhat different names for courses and require mapping; number of students enrolled in a course and number of students passing doesn't always match</p> <p><b>Very large</b> LEA: Schools, LEAs, SEAs might have differing definitions for courses; in last CRDC, LEA evaluated which of its courses corresponded with given definitions; mobility causes beginning of year numbers to differ from end-of-year numbers</p> <p><b>Small</b> LEA: Course definitions need clarity—there could be a student in the justice facility who gets an algebra credit after only being in a class for 10 days because he had already done a year of algebra during the previous data collection time period</p> <p><b>Very large</b> LEA: Enrollment is reported at beginning of year, and number of students taking a course may be different when reporting numbers at end of year. CRDC compares the two numbers, and the LEA will get an error message even though the numbers are accurate</p>
Noninstructional support personnel	<p><b>Mid-size</b> LEA: Need clarity on definition LEA: Noninstructional personnel (security guards, custodians, etc.) are difficult to track because there is a lot of mobility based on needs of schools (need a specific date for the data point or else it will not make sense because of the mobility)</p> <p><b>Small</b> LEA: Psychologists hired at central office and shared between schools, so there would need to be something defining their FTE</p> <p><b>Very large</b> LEA: Need definitions for how to differentiate different types of support staff</p>
Limited English proficiency (LEP)	<p><b>Mid-size</b> LEA: Challenging because it is in a whole separate system (e.g., not in the SIS)</p> <p><b>Mid-size</b> LEA: LEP data is tracked outside of the SIS</p>
Justice facility	<p><b>Mid-size</b></p>

	<p>LEA: Does not collect data on when a student is sent to a justice facility outside the district aside from the information that they have left the district (enrollment record)</p> <p><b>Mid-size</b> SEA: Justice facilities are administered by Department of Justice, and SEA has no control over their actions and the information collected; all justice facility schools are part of a single LEA (state-operated agency) and are not “responsible” for high school completion/graduation</p> <p><b>Very large</b> LEA: Justice facility is a separate program within an alternative school and is not reported separately; participation in a program for less than 15 days would take a long time to gather because students go in and out of programs throughout the year</p>
--	--

Source: CRDC Improvement Project Site Visits, 2014.

**Exhibit 3 | Illustration of types of LEA data systems**

<p><b>Data barns:</b></p> <ul style="list-style-type: none"> <li>▪ Centralized LEA data</li> <li>▪ Stored in accessible electronic systems</li> <li>▪ Data linked by student/person ID</li> </ul>	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>Student Information System (SIS)</p> </div> <div style="text-align: center;">  <p>Staff /finance information system(s)</p> </div> </div>
<p><b>Data silos:</b></p> <ul style="list-style-type: none"> <li>▪ Decentralized LEA data—not in central LEA systems</li> <li>▪ Sometimes stored in hard copy</li> <li>▪ Sometimes stored in department-specific electronic systems</li> <li>▪ Sometimes not stored at all</li> <li>▪ No or unreliable student/person ID linking</li> <li>▪ Typically contain data not required for reporting to the SEA</li> <li>▪ Contain data on elements managed at the school level, not the LEA level, such as school athletics data</li> </ul>	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  <p>Guidance (e.g., AP, SAT, ACT, graduation)</p> </div> <div style="text-align: center;">  <p>Bullying &amp; harassment</p> </div> <div style="text-align: center;">  <p>Athletics</p> </div> <div style="text-align: center;">  <p>Distance education</p> </div> </div> <div style="display: flex; justify-content: center; align-items: center; margin-top: 20px;">  <p style="margin-left: 10px;">Security staff</p> </div>
<p><b>LEA data system type 1</b></p> <ul style="list-style-type: none"> <li>▪ Mostly centralized</li> <li>▪ Least difficult to report for CRDC</li> </ul>	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">  </div> <div style="text-align: center;">  </div> </div>

<p><b>LEA data system type 2</b></p> <ul style="list-style-type: none"> <li>▪ Decentralized but with electronic systems that can link to the SIS</li> <li>▪ Somewhat difficult to report for CRDC</li> </ul>	
<p><b>LEA data system type 3</b></p> <ul style="list-style-type: none"> <li>▪ Decentralized, with manual data storage and no linking to SIS</li> <li>▪ Most difficult to report for CRDC</li> </ul>	

# Achieve better respondent engagement through better communication

## Recommendations for communication

**Challenge:** Lack of clarity regarding the purpose of the CRDC

Recommendations:

Develop a mission statement and include on all CRDC materials

Develop an FAQ about the use of CRDC data

**Challenge:** Communications about new data elements or changes to the CRDC are often received too late to allow LEAs to adequately plan for reporting timely and accurate data

Recommendations:

Develop a standard communication timeline

Maintain respondent web space to access information about the CRDC

**Challenge:** Once communications are sent from OCR, they are not reaching their intended recipients within the LEA

**Challenge:** Lack of clarity regarding the purpose of the CRDC

In past CRDC communication efforts, the purpose of the CRDC has been conveyed through documents sent at the beginning of the collection or at other single points of communication. A review of the responses of site visit respondents shows that this does not appear to have been effective in communicating the purpose of the CRDC; many of the respondents said they did not understand the purpose of the CRDC or even know about it. Typically, only one person receives a communication regarding the CRDC, and it then falls on that person to relay the information in the communication to others, which does not appear to happen effectively.

At the LEAs we visited, only a few staff (aside from the POC) who collected data needed for the CRDC knew what the CRDC was. Sites reported that the initial letters to the LEA may have included statements about the purpose of the CRDC, but these communications were often sent directly to the superintendent and were not consistently shared with the primary POC(s) responsible for gathering and reporting the data.

**Recommendation:** Develop a mission statement and include it on all CRDC materials

We recommend that the CRDC **develop a mission statement** and FAQs for new respondents about the purpose and other key aspects of the data collection that can be used regularly in all forms of CRDC communication with respondents. For example, it can be included at the bottom of letters, in email signatures, on training materials, in reports, on the website, in the online tool, etc. Regular use of consistent statements will lead, over time, to greater saliency of the CRDC with stakeholders who experience any aspect of the collection. Survey research has consistently shown that topic saliency is a major predictor of survey engagement (Groves, Singer, and Corning 2000).

On the OCR webpage presenting information about CRDC, this



Recommendation:  
Engage SEAs in  
communication process

**Challenge: POCs did not have adequate information to share with other offices and staff to obtain accurate data**

Recommendations:

Strengthen the leadership role of the principal on what

Develop a welcome packet of key materials for the POC to use in communicating about the CRDC

**Challenge: LEAs were either unaware of training materials or did not find them helpful**

Recommendations:

Embed training materials within the data collection tool

Develop a CRDC “best practices” guide

Design a process to develop training materials in response to LEA needs, and track their use

**Challenge: The CRDC lacks a formal feedback mechanism for what it means**

What feedback does it provide?  
What is the product?  
You have to build

tagline appears: “Wide-ranging education access and equity data collected from our nation’s public schools.” This statement usefully describes CRDC data to potential users, but it does not emphasize its purpose for respondents.

The mission statement should **emphasize the impact of the CRDC on students’ educational success and opportunity** over compliance or data description. It was clear from the site visits that most CRDC respondents understood the importance of CRDC compliance and the kind of information that is collected in it. However, meaningful data for school and district administrators are data that inform their goals as administrators of learning institutions to help students succeed in school; it is recommended that communications about the purpose of the CRDC reflect these goals of schools and districts.

**Recommendation: Develop an FAQ about the use of CRDC data**

Compliance is an important part of the CRDC, and we believe this should be emphasized in other aspects of communications regarding the CRDC. Part of the new FAQs should explain the process of data analysis and how data are used after the data collection ends. The FAQ should **highlight compliance aspects and the critical need for high-quality data** to ensure that OCR can produce a true and accurate report for LEAs. It should state why it is important for LEAs to respond accurately and also indicate how individual districts and schools can benefit from quality CRDC data.

NCES/OCR should also **consider providing item-by-item justifications** for the data elements. These rationales could be linked from the online tool, similar to data element definitions.

**Challenge: Communications about new data elements or changes to the CRDC are often received too late to allow LEAs to adequately plan for reporting timely and accurate data.**

LEAs often reported receiving communications about the CRDC too late to allow LEAs to adequately plan to collect and report data in a timely and accurate manner. During the site visits, LEAs identified data elements for the 2013–14 school year where there was no collection process in place, significantly limiting their ability to respond accurately to the upcoming collection and potentially increasing the burden by creating an ad hoc data collection process not linked to the SIS system. LEAs expressed the desire to plan FTE allocations for responding to the survey and ensuring that the data are validated. As one LEA mentioned, “100% of the data was not accurate 100% of the time.”

## LEA user experience

### Recommendations:

Develop feedback mechanism for improving the current and future collections

Create and maintain a consistent email address to be used exclusively by CRDC respondents

### **Recommendation: Develop a standard communication timeline**

Among the SEAs and LEAs that provided feedback on the timing of communications, one message was universal—**communication about the CRDC to SEAs and LEAs should begin as early as possible.**

We agree. SEAs and LEAs benefit from as much lead time as possible to prepare for the CRDC data submission process. Specific suggestions include

- NCES/OCR sending a list of all data items/elements and definitions to SEAs and LEAs as early as possible prior to the start of data collection so that the reporting entities can begin to gather and prepare their data submissions.
- NCES/OCR providing early notification of record layouts in order to prepare data submissions to meet data collection deadlines.
- NCES/OCR providing early notification of changes to the data collection compared to the previous collection. One mid-size LEA suggested a 3-year minimum lead time for changes in data collection. The status of data elements, such as required and optional, should also be indicated.
- Creating a calendar for the CRDC reporting cycle: NCES/OCR should create a calendar that lists all key dates (i.e., the date the system opens, due dates for submitting data, requirements for each deadline/date, publication of results) and publicly distribute it.

This last bullet, creating a calendar, is important. There is currently no detailed schedule of activities to guide LEAs in their planning for the CRDC. This contributes to a lack of understanding of the process and inhibits appropriate planning. We recommend that NCES/OCR **agree to and approve for public distribution an official calendar for the 2013–14 and 2015–16 collections.** OCR and NCES should develop a communication timeline that is standard and consistent from year to year. This will allow LEAs and SEAs to better anticipate and prepare for key activities. We have provided a sample calendar in Exhibit 4 that includes data collection activities as well as recommended planning activities for POCs.

We also recommend that NCES/OCR **make public a timeline of the CRDC improvement activities,** particularly for changes to data elements. As referenced in Goal 1, consistent and strict development timelines should apply. A suggested timeline is proposed at the end of this report.

Principals, assistant principals, secretary, whomever the principal designates to provide the data – *Mid-size LEA on who in the school is responsible for reporting CRDC data*

**Recommendation: Maintain respondent web space for access to information about the CRDC**

CRDC respondents need a stable, consistent, and helpful web resource where they can find all of the information they need to help with their submissions and submission planning. All materials, timelines, communications, respondent FAQs, and other resources should be maintained in a user-friendly web space that is not dependent on contractor changes. **This web space should focus specifically on the needs of CRDC respondents.** It can be a page linked to the main CRDC website and accessible to anyone, but it should be designed and formatted to meet the needs of respondents.

The web address for the respondent web space should be used on all CRDC communications with respondents, and the web space should be updated regularly. The address should therefore be simple and easy to remember, such as [nces.ed.gov/crdc/participants](https://nces.ed.gov/crdc/participants).

**Challenge: Once communications are sent from OCR, they are not reaching their intended recipients within the LEA**

In many sites, the channels of communication from OCR to the key LEA staff were ineffective. First, communications shared with superintendents only were not consistently shared with those responsible for data entry or upload for the CRDC. Even getting a communication from a superintendent to a POC was challenging. One very large LEA reported there have been some instances where either the superintendent did not forward the CRDC communication to the correct person in the district or the communication was not handled in a timely manner. Second, communications that did reach the POC were not widely distributed to those in the LEA who routinely collected, validated, or reported the data. Many sites reported that school staff were largely unaware of the collection. In cases where school staff were engaged directly in reporting CRDC data, the POC often lacked the time necessary to adequately plan for the collection. In one site, schools were given a deadline of 10 days prior to the actual submission date to review and correct the data. This means communication is delayed or never gets to the correct CRDC-knowledgeable people at the LEA, which limits their time and ability to respond.

**Recommendation: Engage SEAs in communication process**

The site visits confirmed that there is a hierarchy of data priorities for school districts, and SEA data requests are at the top. The CRDC should develop a contact and **communication strategy that uses the**

**authority and fluency of SEA-to-LEA communication** to improve both the timeliness and quality of LEA responses.

The SEA needs to know when the survey is going to open, what has changed, and who is responsible at the LEAs. Similar to the recommendations for SEA assistance with data reporting, the strategy for involving SEAs in communication is something that will require time and outreach activities that will extend beyond the 2013–14 collection. These activities should happen in conjunction with efforts to engage SEAs in the data collection and reporting process, and we recommend that this be a goal for the 2015–16 data collection.

One SEA currently serves as a liaison to their LEAs by providing notifications in a weekly email (titled “The Tuesday Telegram”) containing information about the CRDC, including CRDC due dates, CRDC resources, and how to request data previously submitted to the state. The SEA also tracks which LEAs have submitted CRDC data, so that it can nudge LEAs who have not yet completed the CRDC. While the LEA POC for this SEA did not participate in any CRDC trainings or webinars, he did participate in a state-based PowerSchool Users Group and meets once per month with the SEA and vendor representative to discuss how the vendor can continuously update its system. Another SEA currently provides data support to their LEAs, responds to LEA questions via phone and email, and holds webinars about the CRDC submission process. These webinars are open to other SEAs and LEAs, as well as to other stakeholders. The SEA also corresponds with LEAs’ POCs, provides FAQs, and provides directions to “opt in” to receive data back from the state.

**Challenge: POCs did not have adequate information to share with other offices and other data collection POCs to obtain accurate data**

There is often a system of delegation for the CRDC within LEAs and schools, and the further removed any given respondent is from the initial CRDC POC, the harder it is to communicate the required information to the appropriate individuals. Information about the purpose and use of the CRDC is not shared widely among the offices responsible for gathering source data directly from schools and students. For example, programmers, HR specialists, guidance coordinators, school principals, registrars, and coaches, etc., may be asked to provide information about the CRDC, but these responders are not getting the information they need to provide accurate data. In most of the site visits, these secondary responders did not even know what they were providing data for. And when they did, it was because

the POC took the initiative to train responders (e.g., school registrars).

Below is an email from an LEA to staff that illustrates this type of request. The confusion over the year for which data are to be provided is evident. This request is for information about restraint and seclusion.

**From:** [REDACTED]  
**Sent:** [REDACTED] November [REDACTED] 2012 [REDACTED]  
**To:** Admin Principals All  
**Cc:** [REDACTED]  
**Subject:** OCR Report

I've been tasked with providing data to the Office of Civil Rights this year. Part of the information needed for FY12 is whether each school had any students who were restrained (mechanically or physically) and any students who were secluded.

So, from each of you, I need a response to this email that states which students fit any of the three above categories. I will need names and dates of each incident. There is no good way for me to get the information out of Infinite Campus because if it was there it would be imbedded in the comments and could be worded any number of ways, making it almost impossible for me to find.

If you could please get that information to me by November 20<sup>th</sup>, that would be great. Again, this information is for FY12. I'd like a response even if you had no students fitting those categories.

And as always, please call with questions.

Thank you.

And this was one response, which includes the LEA staff person's handwritten note about the year reference.

**Subject:** RE: OCR Report

[REDACTED], physically restrained on 11-16, 11-2, and 10-30

*these are FY13*

Leadership at the district level is important. Important to communicate the process to everyone who will be involved —everyone should see the big picture –  
*Mid-size LEA*

**Recommendation: Strengthen the leadership role of the POC**

NCES/OCR should reevaluate their communications strategy to ensure that POCs receive the necessary information for responding to the CRDC in a timely and accurate manner. NCES/OCR should also consider providing guidelines for LEAs on how to select a strong POC who has **direct oversight or direct reporting responsibilities** for the CRDC. When selecting the POC in advance of a CRDC administration, characteristics that make a good POC should be specified and requested. Primarily, a good POC is either someone who has direct oversight (meaning that he or she oversees others, but also has knowledge of CRDC requirements and definitions needed to troubleshoot problems); someone who directly responds to the CRDC data request by gathering and entering a large portion of the data; or someone who has long-standing experience with the CRDC. These POCs, as well as superintendents, should receive all CRDC communications, directly.

**Recommendation: Develop a welcome packet of key materials for the POC to use in communicating about the CRDC**

We also recommend that NCES/OCR develop a “welcome packet” of key materials to be offered to all POCs via the CRDC website as soon as they are identified. The welcome packet should ensure that the superintendent and POC know about timelines, recommended communications with other LEA staff who may be asked to respond to the CRDC, and training and user information tools.

**Recommendation: Consider sending CRDC “newsflashes” that could go to multiple CRDC respondents in an LEA**

We do not recommend sending official CRDC communications to multiple LEA respondents (aside from the POC and superintendent) because this could confuse LEA respondents about their role and accountability. However, to keep other users informed of key dates, changes, and other information, NCES/OCR could consider sending quarterly CRDC newsflashes either to all CRDC users for whom NCES/OCR have email addresses generated from the new online system or to all CRDC users who request to receive these newsflashes.

**Challenge: LEAs were either unaware of training materials or did not find them helpful**

Seven LEAs mentioned calling or emailing the PSC when encountering challenges or utilizing SEA or vendor forums to troubleshoot problems, but were either unaware of CRDC training materials or did not find them useful. Three LEAs interviewed mentioned that the webinars were helpful, eight indicated that the table layouts and definition documentation were helpful, and three indicated that no resources were useful...In cases where LEA respondents did not use the training materials provided, the reason most often seemed to be either they did not know these things existed or they felt that using the materials would make an already time-consuming task even more burdensome; the reason did not appear to have anything to do with the content of the training materials.

The CRDC training materials and videos were typically asynchronous and covered a wide range of topics, making it challenging for LEAs to quickly find the information needed. For example, one LEA explained that the user guide was too complex for them to use. However, resources that LEAs did suggest were helpful were succinct (e.g., list of CRDC definitions) and interactive (e.g., SEA webinar). For example, one LEA explained that the webinar and PowerPoints were helpful in providing basic overview information but the written documents were key points of reference; other LEAs emphasized the

table layout and definitions as being key documents.

**Recommendation: Embed training materials within the data collection tool**

One way to address this problem is to make training information available on the CRDC website, and publicize it as is currently done, but also make short videos and easily accessible help features that respondents can access directly from the online tool, in real time, as they discover that they need help. We recommend video clips that address only one or two concepts at a time.

**Recommendation: Develop a CRDC “best practices” guide to help LEAs efficiently collect, maintain, and report timely and high-quality CRDC data**

One suggestion, mentioned during at least one site visit, is for NCES/OCR to develop a CRDC “best practices” guide for data collection and submission that is similar to guides produced by the National Forum on Education Statistics. The Forum guides can be found online here:

[http://nces.ed.gov/forum/pubs\\_best\\_practices.asp](http://nces.ed.gov/forum/pubs_best_practices.asp).

Another example is the finance handbook found here:

[http://nces.ed.gov/pubs2009/fin\\_acct/preface.asp](http://nces.ed.gov/pubs2009/fin_acct/preface.asp)

**Recommendation: Design a process for developing training materials in response to LEA needs, and track their use**

Although webinars received positive feedback from those who participated, it is unknown how many respondents took advantage of these tools. In order to measure cost effectiveness, it will be useful to track the number of participants for any future interactive trainings.

We recommend that NCES/OCR **conduct a series of precollection Q&A sessions** that would take place between the release of precollection tools and the beginning of data collection.

**Challenge: The CRDC lacks a formal feedback mechanism for learning from and improving the LEA user experience**

Communication about the CRDC needs to be reciprocal. LEAs have vastly different capabilities and experiences in CRDC reporting, and it is important to maintain regular communication about problems and issues related to the CRDC administration. Particularly during the first two collections of using the new online tool, it will be important to



receive feedback in order to continue to improve the tool and fix any unanticipated problems.

**Recommendation: Develop feedback mechanism for improving the current and future collections**

We recommend that NCES/OCR **proactively request feedback** on the tool, flat file submission, and trainings and support a mechanism for receiving this feedback.

**Recommendation: Create and maintain a consistent email address to be used exclusively by CRDC respondents**

We also recommend that NCES/OCR create and maintain an email address to be used exclusively by CRDC respondents to keep collection-related inquiries separate from other inquiries. Currently there is an address of [CRDC2013-14@ed.gov](mailto:CRDC2013-14@ed.gov). However, this will change with every collection and, as a result, respondents will need to update their contacts to account for this, which is an additional burden.



## Exhibit 4 | Sample CRDC Calendar

Activity	2013-14 Collection Dates	2014-15 Collection Dates
<b>Data collection activities</b>		
Selection of points of contact (POCs)		
POC welcome packet explaining roles and responsibilities available		
Notification of reporting requirement changes	Completed	
Table layouts available	June 2014	April 2016
Flat file submission (FFS) tools available	June 2014	April 2016
Q&A sessions for table layouts and FFS tools		
Receive log-in information for system		
Validate contact information and school directory		
Reminder of upcoming data collection opening		
Data collection opens	October 2014	October 2016
Failure to initiate submission warning		
Failure to complete submission warning		
Data collection closes	January 14, 2015	January 16, 2017
OCR review and validation period		
Final data file available to LEAs		
Draft public reports to LEAs and SEAs		
Any comments on public reports due to OCR		
Final public reports released		

– continued

Activity (continued)	2013-14 Collection Dates	2014-15 Collection Dates
<b>Recommended planning activities</b>		
Review POC welcome packet		
Review table layouts	June 2014	April 2016
Review design and element changes		
Review training materials		
Contact SEA if your SEA provides data and agree on data upload content and deadlines		
Assign modules to any other LEA or school staff as desired for data entry or review, and provide information about roles, responsibilities, deadlines, and whom to contact in the LEA or SEA with questions		
Assign and distribute FFS precollection tools and materials		
Hold Q&A session with all LEA and school staff responders		
Begin gathering data for school year		
Remind involved staff of upcoming data collection opening	September 2014	September 2016
Notify staff when data collection is open	October 2014	October 2016
Follow up with staff to ensure timely response	November 2014	November 2016
Review presubmission QC reports		



## Recommendations for the Partner Support Center (PSC)

**Challenge:**  
**Responsiveness of the PSC to LEAs**

Recommendation: Setting expected response times by type of inquiry

**Challenge: LEAs relied on PSC to provide information that could be embedded in the tool**

Recommendations:

Streamline communications to answer common questions within the data collection tool

Set escalation levels to minimize delay in the PSC

Identify state liaisons to develop expert knowledge of the LEA and SEA context

During the 2011–12 CRDC, the Partner Support Center (PSC) received or made around 10,000 phone calls.<sup>3</sup> This is an enormous amount of communication. This amount of contact with the PSC suggests that other data collection processes that could have been used in order to minimize the need for these contacts were underutilized. Improved communication and data collection tools as described in this report should reduce reliance on the PSC. However, the PSC is still an important feature of CRDC communication support.

All communication between the LEAs and OCR flows through the PSC. Feedback from sites on the PSC varied. Some of those interviewed provided positive feedback on the PSC (e.g., some respondents gave the names of the PSC staff who were particularly helpful to them during the last submission), and others offered suggestions for streamlining, simplifying, and improving communications between LEAs, the PSC, SEAs, and OCR.

**Challenge: Responsiveness of the PSC to LEAs**

Some LEAs raised concerns about the responsiveness of the PSC, indicating that they had contacted their SEA for assistance when the PSC was not responding to their inquiry in a timely manner. For example, one mid-sized LEA explained that they had called their SEA after the PSC did not respond quickly enough when they needed an answer about a definition of an item or about whether to include or exclude certain data items in their submission.

**Recommendation: Setting expected response times by type of inquiry**

If not in place already, we recommend **setting expected response times by type of inquiry**. LEAs should be informed of the expected response times. For example, the PSC should assign a code to each inquiry— green, yellow, red, for example—based on predefined codes where possible. The LEA/SEA should be told

<sup>3</sup> The data for the Partner Support Center are derived from the ED-IES-13-R-0053 solicitation for Task Order 30 - 2013: Civil Rights Data Collection Support (CRDC) text: “During the 2011-12 CRDC collection an average of 1,720 calls per month were received lasting an average of 6.45 minutes each. An average of 1,800 emails was received per month. Additionally, a total of 12 faxes were received and 3,200 outbound support calls were made to respondents during the entire data collection field period.”

what code they were assigned. Each code would have a maximum number of turnaround days associated with it. For example, red = 1 working day, yellow = 2 working days, etc.

**Challenge: LEAs relied on the PSC to provide information that could be embedded in the tool**

LEAs reported calling the PSC to understand flat file upload error codes, to reset passwords, and to ask for definitions of data elements. The answers to these routine questions were not easily accessible, requiring LEAs to pause in their work, call the PSC, and await a response before proceeding to complete the CRDC submission.

**Recommendation: Streamline communication to allow common questions to be answered within the data collection tool**

The online tool can be designed to provide guidance about how to troubleshoot common problems and correct errors so that these can be more easily be fixed without contacting the PSC.

**Recommendation: Set escalation levels to minimize delay in the PSC**

We also recommend that NCES/OCR review levels of escalation assigned to categories of problems prior to data collection (which would be assigned appropriate expected response times as indicated above) and **set escalation levels to minimize delay in the PSC** for inquiries that require direct OCR response. Urgent or other legal matters requiring direct and timely OCR involvement should receive minimal handling by the PSC before moving directly to OCR, which would assume direct communication with the LEA.

**Recommendation: Use state liaisons to develop expert knowledge of the LEA and SEA context**

For inquiries that do not require OCR escalation, **additional help resources could be established**. One suggestion received at the site visits is to establish state liaisons with expert knowledge of individual LEAs' and SEAs' data and circumstances (e.g., unique language used to describe data, such as the statewide VA Teacher Credentialing system, New York state BOCES and BEDS) and have these state liaisons work with LEAs, instead of the PSC, on their submissions issues. Liaisons could also be federal liaisons.

It is frustrating to go through the PSC to reach the correct person at OCR who can override requests for data elements not gathered by the district. The PSC doesn't understand the specifics the way

Another suggestion was to create task groups of data content experts, especially in the field of school finance, to provide support on the reporting of data to CRDC.

### **Main challenges for implementing Goal 2**

LEAs give data requests from the SEAs a higher priority than federal data requests. However, the LEA is the legal unit of analysis and response for the CRDC. NCES/OCR must find a way to bridge this gap in priorities. While engaging the SEA in communication can help bridge the gap, it is unlikely that all SEAs will agree to help. Therefore, we recommend that NCES/OCR implement direct LEA communications improvement strategies while also attempting to engage SEAs.

# Achieve better data quality through better data collection tools

## Recommendations for the online data collection tool

### Challenge: Disparate data sources and submitters

#### Recommendations:

Allow for multiple users and permissions

Allow LEAs to grant permission to SEAs to directly upload data

Create topical data modules

### Challenges: Navigation of the submission system was not intuitive and posed a challenge to efficiently entering data

#### Recommendations:

Design the site navigation to align with the manner in which LEAs respond to the survey

Allow for increased and more detailed skip patterns

### Challenge: Methods of indicating missing data or nonresponse were time-consuming and unclear

Prior CRDC data collection tools have been maligned by users as lacking in usability and features that should be provided in any type of online data collection tool. This section describes recommendations grounded by best practices in web survey design for a powerful new system platform that allows for greater flexibility for data submitters and visibility of data entered. The new system should accommodate simultaneous large file uploads and provide immediate feedback on data quality and progress-towards-completion. It should also be designed with the need for future improvements in mind, as outlined in Goals 1 and 2.

### Challenge: Disparate data sources and submitters

As we reported under Goal 1, some data elements are maintained outside the SIS (e.g., athletics data; school finance data; and information about teachers, special education, gifted and talented, AP) but are available either in other data systems or in hard-copy by school, LEA, or SEA in different, decentralized department “silos.” All of the districts explained that to complete the CRDC they must gather data outside of their SIS from various data silos, and their methods for combining these other data sources varied. For example, one LEA explained that their athletic director compiles the data needed for CRDC by looking at team rosters and provides a hard copy list of names of students on teams for the LEA. This list often doesn’t include demographics of the students (e.g., race, ethnicity, gender) or student IDs, so the LEA must try to match the student names with the information in the SIS, which is a very time-consuming process.

### Recommendation: Allow for multiple users and permissions

**Allowing multiple respondents** for the CRDC data collection would spread the burden of reporting across those personnel with direct access to these data; this would reduce the need for the personnel in the department silos to send the data to the LEA POC, who would then have to enter the data. To illustrate, the screenshot below shows redacted athletic data provided to the LEA by the district’s athletic

Recommendation:  
Develop a standardized process to handle cases where LEAs are unable to report complete and accurate data

**Challenge: Submission system lacks usability features standard in survey design**

Recommendations:

Base the tool's usability features on the fundamentals of good web survey design

Ensure easy access to help from within the tool

**Challenge: Reporting and reviewing school-level data was not adequately supported**

Recommendations:

Tool should generate school-level summary reports of data entered into the system

Users should have the ability to export or download all of the data

director.

High Schools	Football	Soccer	x-Ctry	Bask.Ball	In.Trk	Hockey	Swim	Wrest.	Baseba ll	Out.Trk	Volley	Tennis	Soccer	Volley
	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Boys	Girls	Girls
	21	9	0	21	0	0	0	0	5	0	0	0	0	7
	0	0	0	0	0	0	0	0	6	0	0	0	0	0
	33	0	0	21	38	38	21	7	28	38	0	14	19	16
	19	20	0	20	8	0	0	0	9	5	9	0	13	6
	2	0	0	1	0	0	0	0	5	2	0	0	0	2
	3	0	0	3	0	0	0	0	7	3	0	0	18	4
	2	0	0	0	0	0	0	0	1	1	0	0	0	0
	16	13	0	26	11	0	0	0	13	6	0	0	15	9
	6	11	0	16	0	0	0	0		13	0	0	0	9
	11	18	0	16	0	0	0	0	10	5	0	0	13	16
	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	13	3	0	9	0	0	0	0	16	5	0	0	6	5

Each respondent designated by the LEA POC should be given **read and write permissions for specific data sets**. Giving schools or other departments the ability to enter their data directly would reduce the number of steps the LEA needs to take to complete the CRDC. For example, many LEAs receive data from human resources (HR) departments to complete the CRDC. These data might be in hard copy PDF files, in a spreadsheet, or in a database. If HR staff had the ability to enter these data directly into a system for the LEA, the LEA would not need to manipulate the data they receive from HR and then enter it into the CRDC system.

**Recommendation: Allow LEAs to grant permission to SEAs to directly upload data**

In terms of SEAs that wish to provide partial data, Wisconsin provides a good model of processes and procedures for SEA data providers. We recommend using Wisconsin as a model because this state's SEA is very involved in the CRDC, has provided webinars on how to provide SEA data, and generally has expertise in providing SEA data. By enabling multiple users and allowing multiple permissions, as proposed above, **LEAs can grant an SEA respondent permission to directly upload data**, serving as a confirmed "opt in" by LEAs to use SEA data. SEAs can then upload data directly to the online system (alternatively, SEAs can make data available for an LEA respondent to upload). LEAs will still be able to review and write over SEA data if desired, leaving the ultimate data reporting responsibility with the LEA. Timing for the SEA upload should be agreed upon between SEAs and their LEAs. A fixed cut-off period for SEA uploads is an option; however, we recommend that OCR/NCES **develop a suggested timeline for the data upload as well as a schedule of essential communications between the SEA and LEA**. These recommendations should be shared with the SEA and LEA contacts. We recommend that OCR/NCES base their recommendations on the Wisconsin model because of the extensive experience Wisconsin has

in providing partial CRDC data for its LEAs.

### **Recommendation: Create Topical Data Modules**

Enabling multiple users and allowing multiple permissions can potentially cause accidental data overwrites. To address this problem, we recommend **creating topical data modules** that are made up of groups of similar data elements likely to be reported together from the same data system or respondent. Each module can have a different set of users and permissions. This will greatly reduce the likelihood of accidental data overwrites that would be present if all users had full write permissions to all data elements, while also making it easier to assign data reporting responsibilities to various department staff. Recommendations for the topical module groupings will be presented in separate documentation.

It is still possible that data can be accidentally overwritten if modules are used. We recommend that an **overwrite notification message** be generated when a user is about to submit data; such a notification will remind the user to check the status of data already submitted in order to prevent data from being overwritten.

### **Challenge: Navigation of the submission system was not intuitive and posed an obstacle to entering data efficiently**

Data respondents expressed frustration with navigating through the online data collection tool. Many explained that “Refresh was your best friend,” and “too much time was spent waiting for the next page to load so that data entry could occur,” and that it was difficult to access the Part 2 items (e.g., you had to click out of the Part 1 items before accessing the Part 2 items).

Navigating the tool in order to edit data was cumbersome, and some LEAs expressed difficulty in navigating by selecting schools and survey section numbers from a list.

### **Recommendation: Design the site navigation to align with the manner in which LEAs respond to the survey**

Respondents want a tool that is flexible enough to allow them to **jump from section to section and from school to school within a section** of the survey (and bypass the need for sequential navigation) in order to accommodate the problem of data silos. LEAs need the ability to enter the data they have available from the different data silos at different points in time. For example, an LEA may have all of the course data for all of their schools available and would like to enter it



all into the collection tool without having to scroll through every section of the entire survey for each school just to enter this subset of data.

**Recommendation: Allow for increased and more detailed skip patterns**

The tool should also **allow for skip patterns** and auto-fills based on “gateway” or “guiding” questions when data elements are not applicable for LEAs or schools with certain characteristics. For example, elementary schools should not have to report any data for algebra courses. For most data elements, this will eliminate the often-heard complaint about having to enter “zeros” and thus reduce respondent burden. However, this presents challenges when there are multiple users.

While we hope to improve the ease of navigation tools, data entry still necessitates the use of drop-down menus for school selection. However, the system could provide LEAs with a drop-down list of only those schools for which they are to report data.

**Challenge: Methods of indicating missing data or nonresponse were time-consuming and unclear**

Some schools or districts are unable to report information for various reasons. Although it is mandatory for most data elements to be reported, respondents can indicate a problem of data availability by text comment or by calling the PSC. LEAs reported entering zeros when data was either not collected in the same disaggregation as required by the CRDC. This results in situations where OCR cannot distinguish, from the data received, between LEAs reporting zero occurrences and LEAs not having complete and accurate data to report. Additionally, while some LEAs mentioned working with the PSC when data could not be reported, not all LEAs were aware of this possible avenue for assistance.

**Recommendation: Develop a standardized process to handle cases where LEAs are unable to report complete and accurate data**

The new data collection tool should allow for an LEA to identify data that its district or schools do not currently collect. For example, some LEAs explained that they do not collect bullying and harassment data (because their state does not require it), or they do collect it, but not at the level needed to respond to the CRDC (e.g., no data on reason for bullying/harassment incident, no information on the victim of the bullying/harassment). Many LEAs are in the process of beginning to

collect these data, but they may not yet have the data available for the CRDC submission. Accounting for these types of scenarios could be done by having respondents select a reason for not providing data through an easy drop-down menu or similar feature. Submissions with nonresponse to mandatory items will still require approval from OCR, and a compliance plan will still need to be provided to the PSC before certification. This ease-of-use modification to the tool will simply make the initial response process less burdensome.

We also recommend that in the future NCES/OCR consider simplifying the compliance plan process by **providing preapproved plans linked directly to the reason for nonresponse codes** that respondents could agree to from within the survey tool. These recommended plans would be similar to mechanisms frequently used on websites by which customers agree to terms and conditions for the use of services or products.

### **Challenge: Submissions system lacks usability features standard in survey design**

The CRDC online tool is more than just a data repository. It is a self-administered survey instrument. Prior versions of the tool have not provided an appropriate interface for survey response, and users reported frustration with numerous aspects of the design. Examples include from not having enough character space to write a required comments, having to do unnecessary data entry for hundreds of data cells not applicable to a school or LEA, and experiencing to confusion created by a disjointed visual design. The CRDC is a self-administered online survey.

Users specifically requested the following features for inclusion in the online data collection tool, all of which would improve the survey design:

- Inclusion of definitions of data elements in the tool that are immediately accessible to users during the data entry process.
- Features to show submission status and table completeness. One LEA added, don't use term "complete"; it would be better to use "OK to submit" or "ready for validation."
- Ability to view all the data for a single school in a "school report" to verify and certify all the data for a single school.
- Consistency regarding which fields can be left blank and which require a zero. Previously, some fields required a zero and others did not, and there did not seem to be a logical explanation for the difference.
- Better and faster performance of the tool; tool should load each

Design is holistic. Respondents don't care that one person wrote the questions, but another programmed the instrument, and a third designed the interface. They don't care that...limitations require you to use illogical designs - Designing Effective Web Surveys (Couper 2008)

page quickly.

- Ensure that overall appearance, item formats, button functionality and placement, and edits/checks are consistent across entire instrument.

**Recommendation: Base the tool's usability features on the fundamentals of good web survey design**

The fundamentals of good web survey design (Couper 2008) include the following:

- Appropriate instructions and prompts for a self-administered survey; how to respond and what to do must be derived from the tool itself in both design and use of text
- Good customization of navigation and validations
- Use of dynamic features (e.g., to access on-screen help information)
- Responsiveness to real-time user interactions (e.g., warnings)
- Smart and consistent visual components (e.g., layout, graphics) that promote reporting accuracy and respondent engagement

**Recommendation: It is important to ensure easy access to help from within the tool**

Even with improved communication, many CRDC respondents' first introduction to the CRDC will be from the online tool. Data definitions and other help must be just a click away. Ideally, definitions and other help should be linked to each table so that users only need to click a button to get to the specific help they need. Users should not be required to then navigate additional interfaces like PDFs or websites to find information. However, creating this level of customization will take time in both software development and materials and content development. For 2013–14, we recommend that applicable definitions and existing help be linked from any given screen. We also recommend that NCES/OCR consider, for the future, linking to item-by-item justifications or rationales to improve respondents' understanding of the purpose of data requests, as described in Goal 2.

**Challenge: Reporting and reviewing school-level data was not adequately supported**

Sites have to depend on their own internal information and procedures for reviewing data prior to submission. This process is inconsistent and completely dependent on whether or not the LEA can or will

produce its own submission review output for schools or departments. LEAs that produce their own review files will produce submissions with fewer errors. (Examples of review files produced by an LEA are included in Attachment 2.) LEAs who cannot or do not have review files will produce submissions with more errors. One very large LEA explained that two staff compile the large data file that they upload to the CRDC website and each checks the other's work to ensure that the data they submit are accurate. Another very large LEA explained that they built their own system that consolidates, verifies, and validates the data needed for the CRDC from other systems into a single file for upload.

**Recommendation: Tool should generate school-level summary reports of data entered into the system**

The online tool should eliminate the need for the LEA to create its own presubmission review tools. The tool should generate school-level summary reports of data entered into the system that POCs can give to schools so that they can verify the data prior to further processing.

**Recommendation: Users should have the ability to export or download all of the data**

Additionally, users should have the ability to export or download all of the data entered into the system, as well as subsets of data that they can manipulate as needed for review. For example, LEAs reported the need to obtain athletics data from an athletics department with no way to verify the accuracy of such data. Viewing this type of data in relation to other schools and other data elements will be a useful data quality control tool for LEAs.

## Recommendations for data validation and error reports

**Challenge: Data validations and error reports happened entirely post hoc, making it difficult to find and correct errors**

**Recommendation:**  
Implement real-time data validation

**Challenge: Error messages lacked clarity and added confusion to the data validation process**

**Recommendations :**

Provide more informative error messages

Increase the visual distinction between error and warning messages

Provide edit checks to LEAs pre-collection

**Challenge: Inefficient process to resolve warnings through comments**

**Recommendation:**  
The system should implement an easier way to provide comments for like errors that does not require typing the same information repeatedly

**Challenge: Data validations and error reports happened entirely post hoc, making it difficult to find and correct errors**

The tedious process of managing and responding to data validations and error reports is unnecessarily confusing and time consuming. For example, one very large LEA explained that in order to validate their data in the last submission, they had to examine each screen for every school in the district—approximately 50 screens per school for each of the over 100 schools in the district (a total of over 5,000 screens to review). This LEA pointed out that there is a clear need for a more streamlined certification process. In this case the LEA should, at minimum, be able to view a school report that includes all the data for a single school, which could then be validated. Both a small and a mid-sized LEA echoed this suggestion and requested that school-level summary reports of the data entered would be useful, as they could share these reports with school principals who could review and verify their own data.

Many LEAs also expressed frustration about responding to errors and error reports in the previous CRDC submission. For example, one very large LEA explained that last time they had to enter multiple comments for the same issue (i.e., they did not have a specific data element and had to repeat the same comment for all 100 schools), they kept getting the same error message even though they had entered the comment in the correct field for addressing the missing data. Additionally, a small LEA explained that better error reports that clearly identify what has to be fixed are needed, because the previous tool’s error report didn’t provide enough information about what had to be fixed, causing them to have to call the PSC. Suggestions on streamlining the data validation process, improving the information in error reports, and improving the process of clearing errors are important for respondents that report their data via file upload, as well as for those who use the online data collection tool.

**Recommendation: Implement real-time data validation**

All of the error correction in the prior CRDC tool was done post hoc. One of the major strengths of an online survey is the ability to generate checks and notifications at the same time a user is inputting data. The tedious checking of Excel spreadsheets and screens to validate data is unnecessarily onerous and should be vastly reduced.

across schools and within an LEA

**Challenge: Data discrepancies due to high student mobility lead to high frequency of errors**

**Recommendation: Edit checks should be softened to accommodate LEAs with high student mobility**

The online tools should ensure that errors in data entry and uploads be available in real time, as data are entered or reviewed in the online tool. The online tool should have a series of edit checks to capture common problems such as logic checks, range checks, plausible values, etc., and be tailored by data element or group, as appropriate.

Users who want to submit their data via flat files should have two options for reviewing and mitigating data errors. The first option is running an edit report at the table, module, or submission level. These reports are interactive so that edits and comments can be made at the point of the error message. The second option for edits is viewing the uploaded data in the online entry screens after the flat file is uploaded. This method will produce the same user experience as if data were manually typed into the online interface. In that case, errors will be displayed as the user navigates through the tables.

**Challenge: Error messages lacked clarity and added confusion to the data validation process**

Besides the poor collating of error reports in the prior submission tool, many LEAs (from very large to small) explained that the information provided in them was also poor. For example, codes were used to indicate problem elements, which meant that respondents had to take additional steps to figure out what the problem was (e.g., what the code meant), then figure out where the problem was in the dataset, and then find a way to correct the error. Code numbers were not meaningful to the respondent and required contact with the PSC in order to understand the meaning of the error and obtain information on how to address the error, thus wasting time and resources.

It was also reported that distinctions between warnings and hard errors were not clear, again adding more unnecessary deciphering work for the respondent.

**Recommendation: Provide more informative error messages**

The online tool should provide more informative error messages when there is a problem with the data upload or entry. These messages should clearly explain and identify where the error exists in the data; the respondent shouldn't have to figure out what is going on or need to call the PSC to decipher a code. The error messages should include both a clearly labeled error element as well as a comparison element when there is a discrepancy; where appropriate, guidance should be provided on how to resolve discrepancies.

**Recommendation: Increase the visual distinction between error and warning messages**

The online tool should increase the visual distinction between error and warning messages in order to help LEAs prioritize the critical errors that need to be resolved prior to certification.

**Recommendation: Precollection provision of the edit checks should be provided to LEAs**

It would also be helpful if respondents knew in advance what validations are performed, so that they could plan better. Precollection provision of the edit checks should be provided to LEAs (and SEAs as needed) as part of the communications improvements. This information will create transparency in the edit check system and give LEAs crucial information they need to plan for, test, and check their own data for violations prior to submitting the data.

**Challenge: Inefficient process to resolve warnings through comments**

A common complaint from LEAs was that they received an error message when their beginning-of-school-year data (e.g., number of students in a specific course) did not match their end-of-school-year data (e.g., number of students passing the specific course). They explained that the discrepancy was due to high student mobility (e.g., students entered or left the school after the count date) and was not an actual error. The ability to explain with a specific code that such discrepancies are due to high student mobility and are not actual “errors” would help reduce the time needed to respond to each of these “errors” and thus reduce the burden on LEAs.

**Recommendation: Implement an easier way to provide comments**

The system should implement an easier way to provide comments for like errors that does not require typing the same information repeatedly across schools and within an LEA. For example, the ability to enter a code that explains the reason for an error, or the ability to select a reason from a drop-down menu of common reasons for errors (i.e., high mobility of students), would prevent LEAs from having to type the same explanation repeatedly to explain that their data are actually accurate.

Where text is needed, **provide enough character space** to allow for comments in response to error messages. One very large LEA

explained that the previous 256-character limit in the comment field limited their ability to explain why the data are the way they are; they requested more space to explain data issues to prevent political pushback from the press when the data are publicly released. This LEA explained that they wanted sufficient space to explain an “N/A” or “0” response when data elements are not relevant to them (i.e., corporal punishment), and space to explain that the results may not provide a full picture in situations where the LEA is not the primary agency to provide services (i.e., early childhood programs).

**Challenge: Data discrepancies due to high student mobility lead to high frequency of errors**

Enrollment data in Part I are based on the beginning of the school year, while data in Part II (e.g., number of students taking a course) are based on the cumulative/end of the school year. Due to student mobility, it is possible that the number of students who passed a course can exceed the number of students enrolled at the beginning of the year, but the matching of these data is a validation check. This validation leads to a very high frequency of errors for LEAs and juvenile justice facilities that experience high student mobility, and LEAs are forced to respond to each instance where there is a discrepancy. Where these discrepancies are legitimate (i.e., the result of student mobility) it creates unnecessary review burden for the LEAs.

**Recommendation: Change, or soften, edit checks to accommodate LEAs with high student mobility**

We recommend changing, or softening, these edit checks to accommodate LEAs with high student mobility, or, at a minimum, allowing for an easier way for LEAs to explain these discrepancies that do not require them to enter text in comment fields to explain each instance, as discussed above.

In addition, the reporting features suggested above will improve data validation by making it easy for LEAs, schools, and SEAs to look at and review data. These reporting features can be found in Goal 3, *Recommendations for the online data collection tool*.



## Recommendations for flat file submissions

**Challenge: The flat file submission process was cumbersome in all aspects – preparation, error review, and submission**

Recommendations:

Allow data to be uploaded in multiple smaller files and preserve the option for single file uploads

**Challenge: Uploading flat files was difficult because of the system functionality**

Recommendations:

Remove the 10-school limitation for the Excel templates to create flat files

Allow for smaller file uploads with fewer fields

Communicate character requirements for comment fields in the instructions

If templates are provided, make them accessible from the online tool

Allow blank data fields for non-applicable data elements

Populate online tool screens to aid in review

**Challenge: The flat file submission process was unwieldy**

Several SEAs and LEAs provided feedback to the effect that the flat file submission process seemed too unwieldy. Flat files were difficult for one large LEA to create because of the vast number of fields per school that were required in a specific order for the flat file submission format. However, this same LEA was able to routinely submit smaller files of the same content to its SEA without issue. One small LEA explained that it was preferable and more efficient to enter the data manually (even though they create a single database that contains all of the data for the CRDC) because the creation of a single file with over 1,000 fields per school was too cumbersome. One very large LEA explained that separating the submission into a Part 1 and a Part 2 is helpful, but it would be better to break out Part 2 into more sections because it is very hard to find a single specific element in such a large file, especially when they receive an error in their submission. Breaking out Part 2 into smaller sections that can be uploaded separately would make it easier for an LEA to validate their data within the CRDC system. On the other hand, another very large LEA said that it preferred to create a single flat file submission.

This section focuses on the technical aspects of the flat file upload process. For information about user access control, see “Users and permissions” under *Recommendations for the online data collection tool*, above.

**Recommendation: Allow for data to be uploaded in multiple smaller files**

The majority of LEAs and SEAs interviewed wanted to **be able to upload data in multiple files**, and gave the following reasons:

- Smaller files are easier to manage than larger files;
- Smaller files take less time to upload;
- Smaller files allow data to be submitted as they become available;
- Smaller files are easier to “troubleshoot” and correct if errors are detected in the file; and
- LEAs could create and upload data in groups that match data they already produce for the SEA.

**Challenge: The table layouts used by flat file submitters and the information collected by the online tool were different and this caused problems for flat file submitters**

Recommendations:

Ensure future table layouts mimic the overall look and content of the online tool

Conduct cognitive testing of the table layout materials and instructions used for flat file submitters

Suggestions for multiple file uploads were conceptualized in different ways by those interviewed. Some suggested that uploads could contain all the data elements for a single school, with multiple files containing all CRDC items for a subset of schools in the LEA. Others suggested creating “content” or “conceptual” files, with each file containing a subset of the CRDC data items for all schools within the LEA (similar to the ED<sup>Facts</sup> “file groups” submission). The suggestion for the content/conceptual files stemmed from concerns of the size of data files associated with “Part 2” of the CRDC submission process. The “Part 2” data submission is larger and more complex than “Part 1,” and the ability to break the “Part 2” submission into smaller files would be easier for most LEAs and SEAs interviewed. One small LEA indicated that they would use the file upload process if they were able to upload multiple flat files with a smaller number of fields—this would reduce their burden.

**Recommendation: Preserve the option for single-file uploads**

While some LEAs and SEAs expressed a preference for multiple-file uploads, two sites (both very large districts) expressed a strong preference for single-file uploads. These districts stated that having to create multiple files to upload from the current single file would increase their burden, as they have invested time and resources to develop systems to produce a single file for upload. Additionally, the ability to do multiple-file uploads would only be helpful when the districts have errors within a particular section of the data, in which case it would be easier to only correct and resubmit the section of the data containing errors. In all/most other cases, creating multiple files to upload increases these LEAs’ burdens, because these LEAs would need to cross-check the multiple files against each other.

**We recommend that both multiple- and single-file uploads be available**

**Challenge: Uploading flat files was difficult because of the system functionality**

SEAs and LEAs reported numerous **format and functionality issues arising from the use of flat file submissions that should be rectified** in the new online tool. These included the following:

- Uploads required a minimum of 10 schools per file. If an LEA had fewer than 10 schools within a file, the file would not upload, and this limitation prevented LEAs who wanted to do a

flat file upload of fewer than 10 schools from using the system. This limitation forced these LEAs to enter data via the online tool, which takes far longer.

- o Recommendation: **Remove the 10-school limitation for the Excel templates to create flat files**

- The CRDC template was too large (over 1,000 fields per school) and took too long to populate, because it required merging data from multiple “silos” and “barns” into a single file. Additionally, adding comment fields to the already large file often created more problems/issues because it was difficult to keep track of all of the information in such a large file. Multiple file uploads will address this problem, as it will allow users to upload smaller data files related to how they store their data (e.g., separate file uploads for each data silo), and smaller files that allow comments will make it easier to keep track of the comments.

- o Recommendation: **Allow for smaller file uploads with less fields**

- The file upload template’s comment fields themselves also created problems. Specific characters in the comment fields (e.g., legitimate commas) often triggered file upload errors related to file delimiters and field length.

- o Recommendation: **Communicate character requirements for comment fields in the instructions**

- Availability of the templates was also an issue. One very large LEA was not aware that a template existed and thus created one from scratch in Excel format for their file uploads.

- o Recommendation: **If templates are provided, make them accessible from the online tool**

- Almost all LEAs reported that it was exceptionally burdensome to have to “zero fill” data fields that did not have data in the very large file.

- o Recommendation: **Allow blank data fields for nonapplicable data elements**

- Some LEAs requested that the **system allow for tab-delimited and/or Excel-file uploads**, as these file formats are more frequently used by them. However, other LEAs did not express a file format preference.

- During the previous CRDC submission, one very large LEA thought that when it uploaded a data file it would prepopulate the manual screens, but that did not happen. As a result, the

LEA had to go through all the screens for all the schools to ensure that the file upload worked.

- Recommendation: **Populate online tool screens to aid in review**

**Challenge: The table layouts used by flat file submitters and the information collected by the online tool were different and this caused problems for flat file submitters**

The precollection tools for the flat file submission that include the table layouts are currently provided as Microsoft Word documents. These materials represent the first experience that users have of the CRDC for any given collection; thus, they should receive greater attention and review as a respondent tool. Respondents have expressed a desire to have the table layouts more closely reflect the online tool, and we agree.

**Recommendation: Future table layouts should mimic the overall look of the online tool**

Once the new online tool is approved, we recommend that future table layouts mimic the overall look of the online tool (e.g., in terms of colors, font, and other graphics) and that this new design be used in the cognitive testing.

Additionally, if table layouts are to be released early in the year prior to the opening of data collection, we recommend that all modifications and revisions be finalized in time to prepare the table layout for precollection design and publication. In the previous submission, the table layout on the OCR website included different wording of survey items and item numbers that differed from the wording and numbers of the actual items in the online system.

**Recommendation: Conduct cognitive testing of the table layout materials and instructions used for flat file submitters**

The research conducted on the table file layouts in 2014 focused on the text wording and definitions of data elements and table requests. In 2015, we recommend that NCES/OCR implement cognitive testing of the flat file submission instructions and the table layout documents that accompany the precollection flat file submission specifications. These tools are the first materials CRDC responders interact with, so these materials should receive greater care and attention to ensure that they are useful to respondents and make a good first impression for the

future collection.

This testing should focus on content and clarity of the instructions and usability, format, and presentation of the table file layouts as tools for assisting the development of flat file submissions.

**Main challenges for implementing Goal 3**

Giving LEAs the ability to assign multiple users, while also implementing edit checks and skip patterns, are conflicting goals. Having multiple users requires nonsequential input across the survey. However, it is sequential input that allows one to best customize the user experience as one moves linearly through the tool. We see the need for multiple users as more important for ease of administration. The testing of the new approach with real data in real scenarios will be important for the pilot test. NCES/OCR will also need to implement a feedback mechanism to capture comments about this change during the live 2013–14 administration.

# Make data more useful and accessible to CRDC stakeholders

## Recommendations for reporting back to stakeholders

**Challenge: LEAs and SEAs do not have their final data files**

**Recommendation:**  
Create SEA- and LEA-level data files (prerelease of public use data files)

**Challenge: LEAs are unaware of the release of data**

**Recommendation:**  
Provide LEAs and SEAs access to their final data and advance notice public reports

Provide custom reports for each LEA

**Challenge: Rounded data can create a false picture of a school or LEA**

**Recommendation:**  
Consider data swapping, data suppression, and presentation of data as ranges.

**Challenge: LEAs and SEAs do not have their final data files**

Earlier in this report, we provided feedback noting that most SEAs and LEAs would like more information on the rationale, goals, and purpose of the CRDC. One way to assist SEAs and LEAs in understanding the importance of the CRDC is to raise awareness of the creation and release of final CRDC data files and reports, and to “return” data back to SEAs and LEAs, either as data files or in personalized reports that highlight the LEAs’/SEAs’ data. Additionally, the “return” of data back to SEAs and LEAs can assist those entities in the review and quality assurance of their own data.

SEAs and LEAs would like to obtain a summary of the CRDC data at the end of the data collection cycle in a data file format (e.g., comma-separated value file).

**Recommendation: Create SEA- and LEA-level data files (prerelease of public use data files)**

Some LEAs and SEAs suggested creating a window that would allow them to review their data before it is released to the public to check the accuracy of totals calculated by the system after the data are uploaded. **There are precedents for this at NCES, and we recommend that NCES/OCR consider this for the CRDC.** For example, the data imputations for the National Public Education Financial Survey (NPEFS) part of the CCD are reviewed by SEAs prior to data publication.

**Challenge: LEAs are unaware of the release of data**

SEAs and LEAs reported that they never see their data unless there is a compliance problem. This exacerbates the perception of the CRDC as punitive. SEAs and LEAs need to be able to review and consider their data in order to take proactive steps toward making improvements.

### **Recommendation: Provide LEAs and SEAs access to their final data and advance notice public reports**

Even if LEAs' data meet requirements, such data can highlight areas LEAs might wish to target for improvement; in addition good reports can be shared with schools and staff as recognition for their contributions toward creating and maintaining equitable schools and programs. Provide LEAs and SEAs access to their final data and advance notice public reports.

### **Recommendation: Provide custom reports for each LEA**

We recommend creating customized reports for each LEA in a PDF format. The report for a particular LEA would contain three or four graphs or charts highlighting a few key statistics for that LEA, and would be sent to the LEA. The report should have color charts and be visually appealing, so that the LEA can use it as a communication or marketing tool.

Additionally, one OCR office suggested having the reports better organized so that it is easier for them to identify major problem LEAs. It would also be helpful if the OCR office could have more search fields (e.g., district size, demographic data such as race/ethnicity) when running reports.

### **Challenge: Rounded data can create a false picture of a school or LEA**

Although data files and reports contain rounded data as a method of disclosure risk mitigation, rounding of data can create a false picture of the data. We recommend that NCES/OCR review this problem and **consider different methods of presenting data that are affected by rounding as determined** by preset guidelines for reporting. Some strategies NCES uses to address these issues are

- Data swapping
- Suppression of small cells for public release files
- Showing data as ranges if numbers are small—these methods are being used elsewhere at NCES for release of CCD data related to high school dropouts

**Recommendation: Consider data swapping, data suppression, and presentation of data as ranges.**

Lots of time is spent doing this, but except for errors, we never see why we do it –  
*Mid-size LEA*

#### **Main challenges for implementing Goal 4**

Releasing advance public reports or offering custom reports to LEAs is likely to generate feedback that NCES/OCR will need to anticipate and plan for. We recommend that NCES/OCR implement a pilot program for reporting back to stakeholders. The pilot test could happen now using 2011–12 data if NCES/OCR wishes to undertake a full reporting program for 2013–14. Alternatively, NCES/OCR could push implementation of custom reporting to the 2015–16 collection and pilot the reports with 2013–14 data.



# Challenges for Implementation

Implementing the recommendations for the CRDC Improvement Project will require resolving a few key challenges. This section presents the main challenges for accomplishing each of the goals and provides recommendations on how to approach these challenges.

<b>Goal 1. Reduce reporting burden</b>	Maintaining consistency, while making changes to the content and design of the CRDC, are conflicting recommendations for reducing burden. To tackle this problem, we recommend that the pace of change for the CRDC be slow and deliberate, particularly regarding the content. We also recommend that NCES/OCR inform LEAs and SEAs of the forthcoming plans for change and when changes will happen.
<b>Goal 2. Achieve better respondent engagement through better communication</b>	LEAs give priority to SEA data requests over federal data requests like the CRDC. However, the LEA is the legal unit of analysis and response for the CRDC. NCES/OCR must find a way to bridge this gap in priorities. While engaging the SEA in communication can help bridge the gap, it is unlikely that all SEAs will agree to help. Therefore, we recommend that NCES/OCR implement direct LEA communications improvement strategies while also attempting to engage SEAs.
<b>Goal 3. Achieve better data quality through better data collection tools</b>	Giving LEAs the ability to assign multiple users, while also implementing edit checks and skip patterns, are conflicting goals. Having multiple users requires nonsequential input across the survey. However, it is sequential input that allows one to best customize the user experience as one moves linearly through the tool. We see the need for accommodating multiple users as more important for ease of administration. The testing of the new approach with real data in real scenarios will be important for the pilot test. NCES/OCR will also need to implement a feedback mechanism that captures comments regarding this change during the live 2013–14 administration.
<b>Goal 4. Make data more useful and accessible to CRDC stakeholders</b>	Releasing advance public reports or offering custom reports to LEAs is likely to generate feedback that NCES/OCR will need to anticipate and plan for. We recommend that NCES/OCR implement a pilot program for reporting back to stakeholders. The pilot test could happen now, using 2011–12 data, if NCES/OCR wishes to undertake a full reporting program for 2013–14. Alternatively, NCES/OCR could push implementation of custom reporting to the 2015–16 collection and pilot the reports with 2013–14 data.



# Proposed Timeline of Improvements

Goal and Summary of Task	2014						2015						2016						...and beyond												
	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov		Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
<b>Goal 1</b>																															
Solutions for difficult data																															
Core and consistent items																															
Engage SEAs in data reporting																															
Strengthen POC leadership																															
Use ED Facts data																															
Review of timing																															
<b>Goal 2</b>																															
Establish mission statement																															
Develop calendar																															
Engage SEAs in communication																															
Revamp training materials																															
Respondent website																															
Feedback mechanism																															
Improvements to PSC																															
<b>Goal 3</b>																															
Redesign data collection tool																															
Improve error reports																															
Improve flat file submission																															
Cognitive test flat file materials																															
<b>Goal 4</b>																															
Custom data files mechanism																															
Custom report template																															
Guidelines for rounded data																															

# References

Couper, M.P. (2008). *Designing effective web surveys*. New York: Cambridge University Press.

Groves, R.M., Singer, E., and Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, 64(3): 299–308.