U.S. Department of Transportation

Federal Railroad Administration

Confidential Close Call Reporting System (C^3RS) Lessons Learned Team Baseline Phase Report

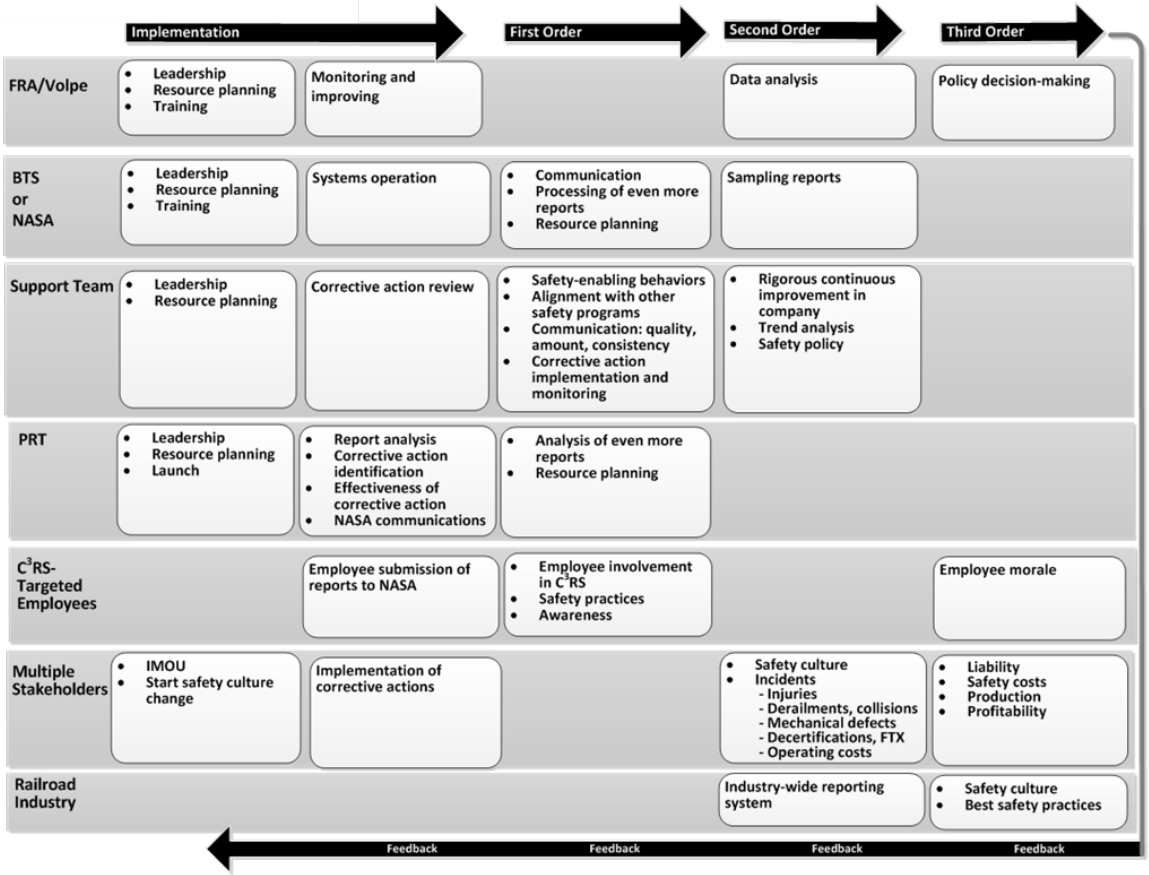
U.S. Department of Transportation

Federal Railroad Administration

OFFICE OF RESEARCH & DEVELOPMENT

C3RS

Process and Outcomes for C^3RS Stakeholders



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2015	Technical Report – December 2014

4. TITLE AND SUBTITLE
Confidential Close Call Reporting System (C3RS) Lessons Learned Baseline Phased Report

5. FUNDING NUMBERS
RR04A5-KMD57
RR04A5-MMD57
RR04A5-LMD57

6. AUTHOR(S)
Joyce M. Ranney (Volpe Center), Melinda Davey (Jacobs), Jonathan Morell (Fulcrum Corporation), Thomas Raslear (Federal Railroad Administration)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
John A. Volpe National Transportation Systems Center
55 Broadway
Cambridge, MA 02142-1093

8. PERFORMING ORGANIZATION REPORT NUMBER
DOT-VNTSC-FRA-14-01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research and Development
Washington, DC 20590

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
DOT/FRA/ORD-15/08

11. SUPPLEMENTARY NOTES
Program Manager: Thomas Raslear

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA Web site at http://www.fra.dot.gov.

12b. DISTRIBUTION CODE

13. ABSTRACT
The Federal Railroad Administration (FRA) has established a program called the Confidential Close Call Reporting System (C^3RS), which allows events to be reported anonymously and dealt with non-punitively and without fear or reprisal through structured collaboration between representatives of railroad management, labor unions, and FRA who work together in Peer Review Teams (PRTs). Close calls or near misses are occurrences that could have led to or resulted in an accident or casualty, but did not. Close call programs have been shown to contribute to improved safety in other industries, but their effectiveness in the railroad industry has not yet been systematically evaluated. To address this question, FRA implemented a rigorous evaluation that is taking place over the entire course of the C3RS life cycle. The C3RS evaluation is designed to answer three major questions: (1) What conditions are necessary to implement C3RS successfully? (2) What is the impact of C3RS on safety and safety culture? (3) What factors help to sustain C3RS over time? Data are being collected for three time periods: from the beginning of C3RS through an implementation period (baseline), at about the middle of the test period (midterm), and at the end of the test period (final). This report provides baseline findings from the evaluation at four demonstration sites.

14. SUBJECT TERMS
Close calls, continuous improvement, employee involvement, freight rail, human factors, passenger rail, risk reduction, safety culture

15. NUMBER OF PAGES
72

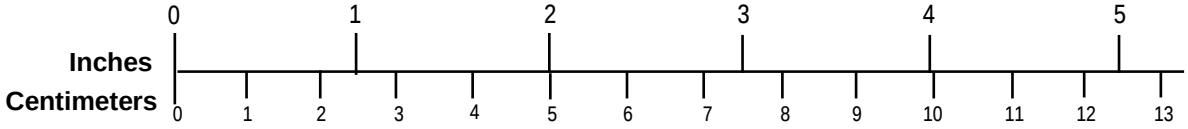
16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

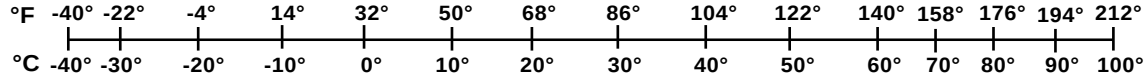
METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC	METRIC TO ENGLISH
LENGTH (APPROXIMATE)	**LENGTH (APPROXIMATE)**
1 inch (in) = 2.5 centimeters (cm)	1 millimeter (mm) = 0.04 inch (in)
1 foot (ft) = 30 centimeters (cm)	1 centimeter (cm) = 0.4 inch (in)
1 yard (yd) = 0.9 meter (m)	1 meter (m) = 3.3 feet (ft)
1 mile (mi) = 1.6 kilometers (km)	1 meter (m) = 1.1 yards (yd)
	1 kilometer (km) = 0.6 mile (mi)
AREA (APPROXIMATE)	**AREA (APPROXIMATE)**
1 square inch (sq in, in^2) = 6.5 square centimeters (cm^2)	1 square centimeter (cm^2) = 0.16 square inch (sq in, in^2)
1 square foot (sq ft, ft^2) = 0.09 square meter (m^2)	1 square meter (m^2) = 1.2 square yards (sq yd, yd^2)
1 square yard (sq yd, yd^2) = 0.8 square meter (m^2)	1 square kilometer (km^2) = 0.4 square mile (sq mi, mi^2)
1 square mile (sq mi, mi^2) = 2.6 square kilometers (km^2)	10,000 square meters (m^2) = 1 hectare (ha) = 2.5 acres
1 acre = 0.4 hectare (he) = 4,000 square meters (m^2)	
MASS - WEIGHT (APPROXIMATE)	**MASS - WEIGHT (APPROXIMATE)**
1 ounce (oz) = 28 grams (gm)	1 gram (gm) = 0.036 ounce (oz)
1 pound (lb) = 0.45 kilogram (kg)	1 kilogram (kg) = 2.2 pounds (lb)
1 short ton = 2,000 pounds (lb) = 0.9 tonne (t)	1 tonne (t) = 1,000 kilograms (kg) = 1.1 short tons
VOLUME (APPROXIMATE)	**VOLUME (APPROXIMATE)**
1 teaspoon (tsp) = 5 milliliters (ml)	1 milliliter (ml) = 0.03 fluid ounce (fl oz)
1 tablespoon (tbsp) = 15 milliliters (ml)	1 liter (l) = 2.1 pints (pt)
1 fluid ounce (fl oz) = 30 milliliters (ml)	1 liter (l) = 1.06 quarts (qt)
1 cup (c) = 0.24 liter (l)	1 liter (l) = 0.26 gallon (gal)
1 pint (pt) = 0.47 liter (l)	
1 quart (qt) = 0.96 liter (l)	
1 gallon (gal) = 3.8 liters (l)	
1 cubic foot (cu ft, ft^3) = 0.03 cubic meter (m^3)	1 cubic meter (m^3) = 36 cubic feet (cu ft, ft^3)
1 cubic yard (cu yd, yd^3) = 0.76 cubic meter (m^3)	1 cubic meter (m^3) = 1.3 cubic yards (cu yd, yd^3)
TEMPERATURE (EXACT)	**TEMPERATURE (EXACT)**
[(x-32)(5/9)] °F = y °C	[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO



For more exact and/or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Acknowledgments

Thanks to the following organizations and people:

- The National C^{3}RS Steering Committee and those who participated in the Federal Railroad Administration's (FRA) Research and Development-sponsored human factors workshop, Improving Railroad Safety Through Understanding Close Calls (in 2003), which led to the start of the initial C^{3}RS demonstration.[1]
- Cassandra Cantu (Volpe National Transportation Systems Center (Volpe Center)) for participating on the Evaluation Team and for her editorial support.
- Dr. Jordan Multer (Volpe Center) and Jane Saks (DigitaliBiz) both members of the Volpe Implementation Team, who provided expertise on C^{3}RS operations and reviewed all briefings and reports.
- Dr. Michael Zuschlag (Volpe Center) and Dr. Wayne Nelson for performing an external review of the statistical methods and results.
- Demetra Collia at the Bureau of Transportation Statistics (BTS) for providing redacted C^{3}RS corrective actions data, designing the Railroad Safety Culture Survey layout, and administering the survey.
- Those who participated in the Lessons Learned interviews and stakeholder reviews of findings and reports, including personnel from FRA (including the inspectors who sit on the C^{3}RS Peer Review Team, Deputy Regional Administrations, and FRA Office of Research and Development and Railroad Safety), labor and managers from the demonstration site railroads, the Volpe Implementation Team, BTS, and the National Aeronautics and Space Administration (NASA).

This work is being performed under an interagency agreement between FRA's Human Factors Research and Development (R&D) Program and Volpe's Surface Human Factors Division.

[1]Improving Railroad Safety Through Understanding Close Calls workshop. Proceedings available at http://www.closecallsrail.org/publications_workshops.aspx

Contents

Illustrations

Tables

Executive Summary

Historically, fear of reprisal, punitive disciplinary action, litigation, and a blame-based safety culture in the railroad industry have discouraged free and open discussions of hazardous workplace conditions and how they might be eliminated. Sometimes as a result, accidents occur that could otherwise have been prevented. In recognition of the problem, the Federal Railroad Administration (FRA) established a pilot project through which "close calls" can be reported anonymously and dealt with through open and honest communication among representatives of railroad management, labor, and FRA who work together on designated Peer Review Teams (PRTs). This innovative program, known as the Confidential Close Call Reporting System (C3RS), defines a close call or "near miss" as "an opportunity to improve safety practices in a situation or incident that has a potential for more serious consequences."

Close call programs have been successful and effective in other industries at bringing about safety performance improvements, but their effectiveness in the railroad industry has been an open question. To address this, FRA implemented a rigorous evaluation that is taking place over the course of the C3RS life cycle. Data are being collected for three time periods: from the beginning of C3RS through about one and a half years of C3RS reporting (baseline); at the middle of the test period after around three years of reporting (midterm); and at the end of the test period of five years (final). The evaluation process will examine how C3RS is being implemented, what impact it is having, and whether sustainability is being achieved.

This report summarizes the baseline findings from each demonstration site: Union Pacific (UP) North Platte Service Area, Canadian Pacific Railway (CP) Chicago Service Area Road Territory, New Jersey Transit (NJT), and Amtrak. The baseline phase for each of the four railroads took place between 2007 and 2012, depending on when the railroad joined the C3RS demonstration. Since the findings are from baseline reports, this document does not contain a great deal of information on impact or how the process of implementation developed over time; rather, this paper does the following:

- Explains why C3RS was needed.
- Describes the C3RS program—its development, structure, and process.
- Features a diagram that shows how the evaluation will make meaningful comparisons across sites and over time (Figure 7).
- Details the quantitative and qualitative measures used in the evaluation.
- Includes a diagram (logic model) that shows the relationship among the parts and expected outcomes of C3RS (Figure 6).
- Provides whatever baseline findings that can be made public without divulging the identities of particular participating railroads.

Even though this report covers only the baseline period, some important findings did emerge. These findings are based on multiple sources of data, interviews with program participants, field notes, a Railroad Safety Culture Survey, and reporting rates.

- The data indicated that C3RS start-ups were successful at all participating railroads though the process required cooperation and collaboration among multiple stakeholders, skill

acquisition in unfamiliar activities, and a major departure from traditional railroad industry behaviors and practices concerning the reporting of close call or near miss events.

- Even though the evaluation process added to the effort needed to implement and operate C3RS, both labor and management allowed the evaluation to accomplish its goals by cooperating with the Evaluation Team and providing extensive data and time for them.
- The inevitable disagreements about issues not covered by the Implementation Memoranda of Understanding (IMOU) were resolved to the point where C3RS was allowed to continue at the four demonstration pilot locations during the baseline phase of the evaluation.
- The survey scale scores suggested managers perceived safety culture to be more positive than labor (Figure 9 and Figure 10). Though some perception gaps between labor and management are expected, this data uses a large sample of labor and management input to provide deeper insight into their relationship
- Interview data revealed that, even at the baseline stage, there was a positive impact on safety culture and labor-management relations as a result of C3RS' start-up activities (negotiating the IMOU, PRT formation and training, labor-management interactions within the PRT, and program roll-out). These findings are not surprising because if a railroad is willing to negotiate an IMOU in the first place, that railroad is prepared to change in ways that would foster the cooperation needed to successfully operate C^3RS.

1. Introduction

This report provides baseline findings from the evaluation of the Federal Railroad Administration's (FRA) Confidential Close Call Reporting System (C^3RS).[2]

1.1 Background: The Need for a New Approach to Railroad Safety

In 2002, FRA noted that there had not been any major improvements in human-factors-caused accident levels in recent years. On the basis of this observation, FRA decided to implement and evaluate an approach to safety that has proven successful in the airline, chemical processing, nuclear, and transportation industries but had not yet been used with railroads.[3] That approach came to be known as C^3RS.

When the idea for C^3RS was first conceived there were good reasons to doubt its effectiveness in the railroad industry. This uncertainty sprang from the history of contentious labor-management relations in the industry and use of the fault-based Federal Employers Liability Act (FELA) system in railroads instead of Workers' Compensation.[4,5] Because FELA is fault-based, it inhibits the kind of free and open discussion that is needed to understand why an accident happened.

In brief, C^3RS developed as follows[6]:

- In 2002, FRA Office of Research and Development (ORD) decided to fund and conduct a workshop on close call reporting systems.
- To develop the workshop, a planning committee was formed. The committee included representatives of potential carriers from the National Transportation Safety Board (NTSB) and labor representatives from the United Transportation Union (UTU), Brotherhood of Locomotive Engineers and Trainmen (BLET), Brotherhood of Maintenance of Way Employees (BWME), Brotherhood of Railroad Signalmen (BRS), and American Train Dispatchers Association (ATDA). The Volpe National Transportation Systems Center (Volpe Center) was in charge of these activities under the sponsorship of FRA ORD.
- The workshop was also held under the auspices of the Volpe Center, and proceedings were published in 2004.[7] In addition to railroad and FRA personnel, representatives from

[2] http://www.closecallsrail.org

[3] A description of close call programs in other settings is presented in the literature review that was performed as the first task in the C^3RS evaluation. (See Morell, et al. Confidential Close Call Reporting in the Railroad Industry: A Literature Review to Inform Evaluation, April 2006.)

[4] 45 U.S.C. § 51 et seq. (1908).

[5] National Research Council. Compensating Injured Railroad Workers Under the Federal Employer's Liability Act: Special Report 241. Washington, DC: The National Academies Press, 1994.

[6] Not all of these steps were strictly sequential. The order presented here represents a sense of the logical order of the work and the points in the sequence where most of the activity took place.

[7] Saks, J., Multer, J., Blythe, K. (2004). Proceedings of the Human Factors Workshop: Improving Railroad Safety Through Understanding Close Calls. Department of Transportation: Federal Railroad Administration: Office of Research and Development. (DOT-VNTSC-FRA-04-01).

the Bureau of Transportation Statistics (BTS), the National Transportation Safety Board (NTSB), and the airline industry were present. One of the workshop's goals was to introduce railroad stakeholders to ways that other industries have handled close call reporting. A second purpose was to provide a setting in which a coalition of interests could form to advance close call reporting in the railroad industry.[8]

- In 2003, FRA ORD decided to establish the C^{3}RS program on the basis of stakeholder enthusiasm. Its reasoning was that because confidential close call reporting had worked in other settings, it was worth the risk of testing whether such a program would work for railroads. If evaluation proved the new approach successful, close call reporting could be scaled up throughout the railroad industry. ORD contracted with the Volpe Center to form an Implementation Team (VIT) to implement the C^{3}RS program.
- After the workshop, the primary task of the Planning Committee was to draft a Memorandum of Understanding (MOU) that was acceptable to management, labor, and FRA. Negotiating this MOU involved considerable work and good will among all participants because many thorny issues had to be resolved. These included:
    - What should the maximum time be between an employee's observation of a close call and when the report has to be submitted?
    - How many crew members are to be covered by a specific close call report? Only the reporter or the whole crew?
    - How indemnity and avoidance of discipline, which is the cornerstone of the reporting system, was to work? What events would qualify for providing indemnity and avoidance of discipline?
- Because of the many difficult issues involved, successfully negotiating the MOU took two years.[9]
- FRA and the Volpe Implementation Team worked to recruit railroads for the C^{3}RS pilot tests. Four railroads—Union Pacific (UP) North Platte Service Area, Canadian Pacific (CP) Chicago Service Area Road Territory, New Jersey Transit (NJT), and Amtrak—became pilot sites as a result of these efforts.
- After work at the first site began in 2007, the planning committee became the National Steering Committee and the membership was expanded. Steering Committee meetings were open for anyone to attend. The Volpe Implementation Team communicated progress to the Steering Committee and exchanged information.
- In late 2011 transition planning began to move the management of C^{3}RS from FRA's ORD to the Office of Railroad Safety (ORS). The transition is planned to be completed by the end of 2014. ORS's role is to fund, roll out, and manage programs that FRA determines are worthy of industry-wide application. FRA continues to sponsor the

[8]A full history of the Improving Railroad Safety Through Understanding Close Calls workshop can be found in its proceedings at http://www.closecallsrail.org/publications_workshops.aspx

[9]The full text of the MOU can be found at http://www.closecallsrail.org/publications_mou.aspx

lessons learned evaluation of the C^{3}RS program to provide data to improve implementation, outcome, and sustainability.

The close call approach tested had five essential characteristics:

1. A focus on precursor events that may lead to accidents, specifically on the operation of systems rather than on the fault of individuals.
2. Heavy reliance on close call data to determine root and contributing causes.
3. Cooperative labor-management collaboration to perform analysis and decide on corrective actions.
4. An organization capable of implementing corrective actions.
5. In exchange for providing information about close call the reporter would be provided with protection. The reports would be confidential and the reporter would be protected from discipline for the close call reported.

In addition to generating corrective actions that improve safety, close call reporting systems have been shown to contribute to safety by allowing labor and management to work together in a cooperative manner, thereby building trust and improving the organization's safety culture.[10]

Establishing a close call reporting system required a formal set of understandings among labor, management, and FRA. This, in turn, required an agreement that defines which events would technically be classified as close calls and described how management would handle the disciplinary implications of allowing close calls. FRA had to be willing to grant waivers to regulations that would allow exemption from discipline for reporting employees and protection from fines for railroads. All stakeholders had to be satisfied with provisions to protect confidentiality.

A considerable amount of negotiating was needed to articulate these understandings in a manner that would be acceptable to all parties. Because of the time and effort involved, C^{3}RS programs could not proliferate throughout the industry if each implementation required a separate negotiation that started from scratch. In recognition of this barrier, FRA, industry, and labor established a Steering Committee to draft a model Memorandum of Understanding (MOU) that could be used as a basis for any unique close call system implementation.[11]

After the MOU was completed, the Steering Committee remained to guide C^{3}RS as it was deployed and serve as a conduit of information for the industry. In addition to the practical tasks of drafting an MOU and advising, the success of the Steering Committee itself can be seen as a step in establishing the kind of cross-functional, cross-stakeholder cooperation that is needed to change safety culture in the railroad industry.

1.2 Confidential Close Call Reporting System

What makes C^{3}RS unique in the railroad industry is its use of a mechanism that allows people to report dangerous conditions without fear of retribution, thus providing input to an effective

---

[10]Morell, J.A., Hanssen, C., Thompson, D., Wallace, R., & Wygant, B. (2006). Confidential close call reporting in the railroad industry: A literature review to inform evaluation. Available at http://www.closecallsrail.org/pubs/litEval.pdf

[11]http://www.closecallsrail.org/pubs/Model_MOU.pdf

multiple-stakeholder problem-solving process. For the system to be effective within a given railroad, five conditions must be in place:

1. *Clear agreement among all parties:*

   A C^3RS Implementing Memorandum of Understanding (IMOU) must be negotiated among labor, management, and FRA at a given railroad. The IMOU must clearly articulate events that are and are not included within the program's scope and the obligations and commitments of all parties to the agreement. Examples of topics that need to be covered are conditions under which discipline can be avoided, and how disputes will be resolved.

2. *Frank discussion and assured confidentiality:*

   To facilitate free and open discussion of close calls, there must be a trusted mechanism through which workers can report close call events in such a manner that (a) their identity is protected and (b) enough detail is known for a problem-solving team to be able to investigate the situation and understand it well enough to formulate a corrective action.

3. *Team approach to problem solving:*

   The three key stakeholders must act as a team, entitled Peer Review Team (PRT), to perform multiple-cause-analysis problem solving. To partner effectively, labor, management, and FRA must be trained and adhere to the same problem-solving process. Team membership must comprise all stakeholders to the IMOU.

4. *Ability to implement change:*

   Participating companies must have the ability to act on recommendations made by their PRT. "Action" includes the ability to:

   - Assess and, as necessary, modify recommendations.
   - Implement change.
   - Communicate and coordinate with the problem-solving teams.
   - Track the progress of implementation efforts.
   - Assess the consequences of corrective actions.
   - Inform the workforce of changes that have been made.

5. *FRA agreement:*

   FRA must grant waivers that will allow participating railroads to refrain from taking disciplinary action in situations where such action would normally be required. The waiver originally dealt with Part 240, a provision that requires engineers to lose their certification if particular rules are violated. The waiver also applied to carriers to protect them from fines imposed by FRA in the event that an offending employee was not disciplined. Part 242, which required conductor certification, took effect several years after C^3RS began, at which point the waiver had to be extended.

The above five conditions focus on the implementation of C^3RS within specific railroads. To transcend the boundaries of any particular implementation, another requirement is an effective method for keeping the industry informed of C^3RS and its impacts. With these conditions in place, a process described in the next section can be implemented.

1.3 Description of the C^{3}RS Process

Information about a close call must flow from the observer (or observers) of the event to the PRT, whose members are drawn from each of the craft unions involved, management, and FRA. However, this information cannot be directly provided to the PRT because of the need to redact these reports. Confidentiality is critical. Therefore, a neutral third party is interposed between the reporter and the PRT.

At UP, CP, and NJT, that mechanism is the Bureau of Transportation Statistics (BTS). BTS is a legally designated "Federal statistical agency" and thus is able to operate under the provisions of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002. CIPSEA protection allows location information to be included in each close call report. However, the data is considered confidential and kept secure. The security includes a secure data storage room at each pilot site. In addition, agreements with each railroad included an additional security measure in which the PRT physically ships the C^{3}RS database on a laptop back and forth to BTS once a month. Originally BTS only accepted reports on paper, but as its systems were developed, electronic submission came into practice.

The National Aeronautics and Space Administration (NASA) is the second trusted third party and it handles reporting for Amtrak. NASA has extensive experience as a third party with close call reporting in the aviation industry through the Aviation Safety Reporting System (ASRS).[12] It uses a process similar to ASRS but modified to fit the unique needs of C^{3}RS, chiefly the requirement that reports are able to be directed to specific companies. NASA receives reports both electronically and through regular mail. To ensure confidentiality, all personal and organizational names are removed before entering them into the NASA database. Additionally, dates, times, and related information which could be used to infer identity are either generalized or eliminated.[13]

After receiving a report, the third party (BTS or NASA) acknowledges receipt to the person who submitted it and conducts a briefing with the submitter to collect additional information. The NASA interviews are less structured than those conducted by BTS. After the interview, any identifying information is removed. The third parties periodically send groups of cases to the PRTs to review.

A key difference between the BTS and NASA processes is that, because of CIPSEA, BTS is able to retain and protect data in their original form and thus does not have to purge any information from its records. As a result, BTS has the ability to go back over the detailed reports and analyze changes over time, or to make comparisons among types of reports. NASA must delete certain information to ensure confidentiality, such as location and name. NASA sends the cases to the PRT in batches through encrypted e-mail. Because all identifying information is removed by NASA, the PRT does not know the location of the close call and may be forced to think about system-level issues. A "system level issue" is a problem in need of corrective action in which the "spread' of the problem or its corrective action crosses significant organizational boundaries.

---

[12]NASA's Aviation Safety Reporting System (ASRS), http://asrs.arc.nasa.gov/

[13] NASA's Confidentially Close Call Reporting System: Confidentiality and Immunity Policy, http://c3rs.arc.nasa.gov/information/confidentiality.html

These may be geographical boundaries, e.g. service areas or regions, or functional boundaries, e.g. "Transportation" and "Mechanical".

Once the PRT is in possession of the information from BTS or NASA, it performs multiple cause incident analysis (MCIA) to determine the underlying causes of an event and to develop corrective actions to address it. The PRT then sends recommendations for changes to local and corporate management (Figure 1).

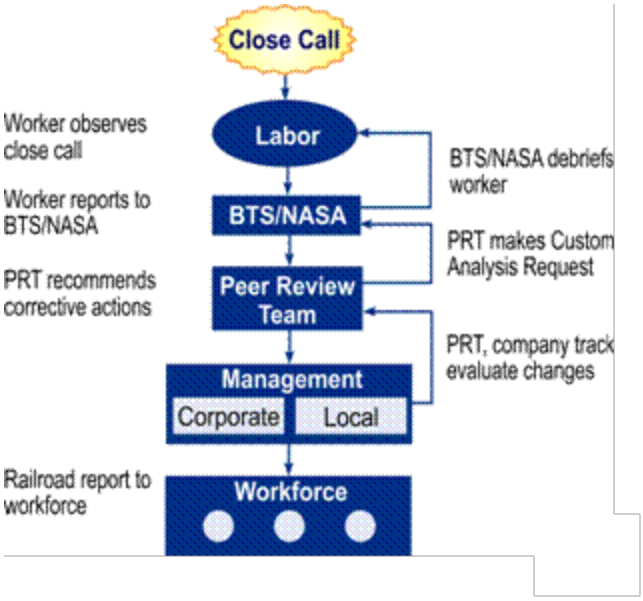


Figure 1. C^{3}RS Process

1.4 Railroad Demonstration Sites

With support from the Steering Committee, the Volpe Implementation Team sought to recruit two passenger and two freight railroads. Criteria for inclusion included a sufficient number of employees to obtain enough reports (goal: 1,200), high interest in participating, and a reasonable level of trust between labor and management. The recruiting of participants was a protracted process that lasted from 2007 to 2011. Table 1 shows which railroads joined and when.

Table 1. C^{3}RS Demonstration Sites

Railroad	C^{3}RS Start Date
Union Pacific (UP) North Platte Service Area	February 2007
Canadian Pacific (CP) Chicago Service Area Road Territory	April 2008
New Jersey Transit (NJT)	October 2009
Amtrak	February 2011

1.4.1 Union Pacific North Platte Service Area

Union Pacific (UP) is a freight railroad that operates in 23 states in the western two-thirds of the United States. UP has 46,500 employees and 8,400 locomotives. UP has more than 10,000 customers including companies from the agricultural, automotive, chemical, coal industries. UP connects West Coast and Gulf Coast ports to eastern gateways, Canadian railroads, and Mexico.[14]

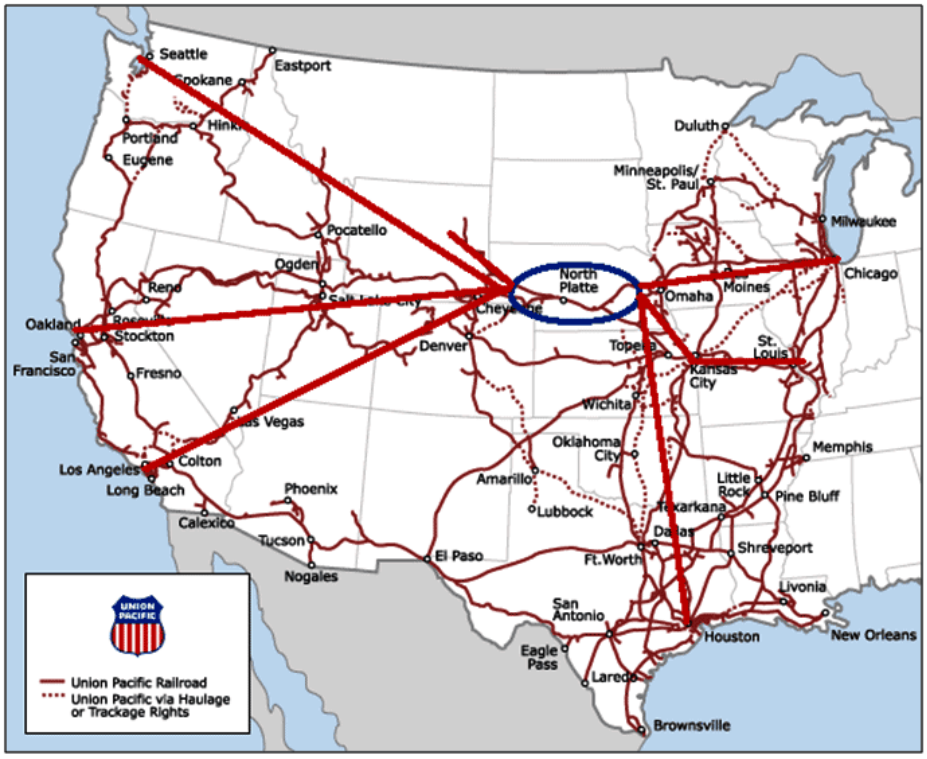


Figure 2. UP C[3]RS Boundaries

UP's North Platte Service Unit located in North Platte Nebraska began participating in the C[3]RS demonstration in February 2007 (See Figure 2). The service unit includes employees from BLET and UTU. Approximately 2,500 employees who work on yard and road crews were eligible to submit close call reports. The UP IMOU contained the milepost boundaries of the North Platte Service Unit and included road and yard employees from UTU and BLET unions working within the milepost boundaries of the service unit.[15]

UP's Bailey Yard in North Platte, Nebraska is the largest rail yard in the U.S. It covers 2,850 acres and eight miles. It handles 14,000 rail cars every 24 hours. 3,000 cars per day are sorted daily in the "hump yards." The hump yards allow four cars a minute to roll into any of 114

---

[14] Union Pacific (2014). "Company Overview." (Web site) Omaha, Nebraska. Accessed online: August 4, 2014. (http://www.up.com/aboutup/corporate_info/uprrover/index.htm)

[15] Confidential Close Call Reporting System Implementing Memorandum of Understanding (C3RS/IMOU) North Platte Service Unit of the Union Pacific Railroad Company. (http://www.closecallsrail.org/pubs/North_Platte_Signed_IMOU.pdf)

"bowl" tracks. These sorted cars become trains headed for locations in the East, West, and Gulf Coasts of North America as well as the Canadian and Mexican borders. The trains carry raw materials and finished products such as automobiles, coal, grain, corn, sugar, chemicals, and steel, and consumer goods such as electronics and apparel.[16]

1.4.2 Canadian Pacific Chicago Service Area Road Territory

Canadian Pacific (CP) Railway covers 14,000-miles from the Port of Vancouver in western Canada to the Port of Montreal in eastern Canada as well as to the U.S. cities of Chicago; Newark; Philadelphia; Washington, DC; New York City; and Buffalo.[17] CP has around 16,100 employees. [18]

In April 2008, employees of the BLET and UTU unions within the CP Chicago Service Area Road Territory began to participate in the C^{3}RS demonstration. Approximately 350 employees were eligible to submit close call reports. Unlike the UP site, which contains a large yard and road crews, most of the pilot employees in the CP demonstration work on road crews that are stationed on the main line. The boundaries of the C^{3}RS pilot are defined from east to west, as the territory operated by train crews going on or off duty between Newport, Minnesota and Tower A-20, near Chicago; and all track in between (See Figure 3). Crews operating in Portage and Milwaukee were covered during their entire tour of duty even while operating on foreign roads. Also included and covered during their entire tour of duty were all crews working between St. Paul and Portage, River subdivision road-switch crews, and crews operating to Waterloo Pit.

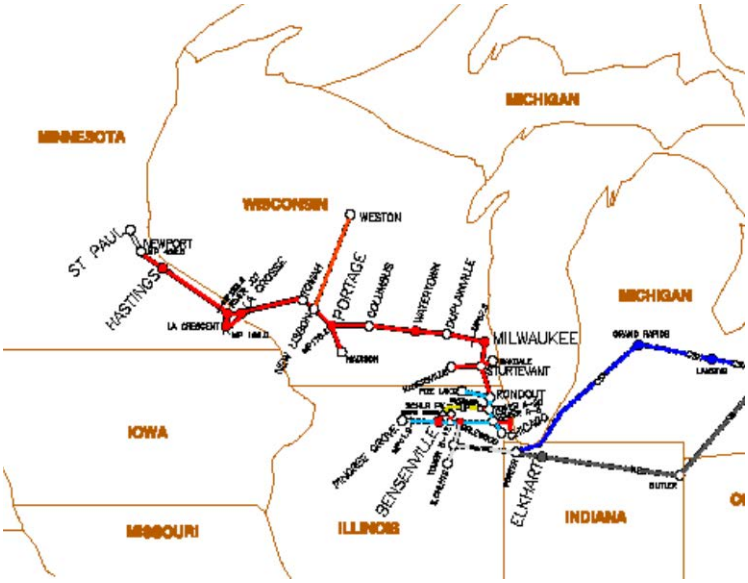


Figure 3. CPR Chicago Service Area Road Territory C^{3}RS Boundaries

[16] Union Pacific (2014). "Bailey Yard." (Web site) Omaha, Nebraska. Accessed online: August 4, 2014. (http://www.up.com/aboutup/facilities/bailey_yard/)

[17] Canadian Pacific (2014). "Our History." (Web site) Calgary, Alberta, Canada. Accessed online: August 4, 2014. (http://www.cpr.ca/en/about-cp/our-past-present-and-future/Pages/our-history.aspx)

[18] Canadian Pacific (2014). "About CP." (Web site) Calgary, Alberta, Canada. Accessed online: August 4, 2014. ,http://www.cpr.ca/en/about-cp/Pages/default.aspx

1.4.3 New Jersey Transit

New Jersey Transit (NJT) is New Jersey's public transportation corporation. NJT's service area is 5,325 square miles; NJT is the Nation's third-largest provider of bus, rail, and light-rail transit, linking major points in New Jersey, New York, and Philadelphia. NJT has 2,027 buses, 711 trains, and 45 light-rail vehicles on 236 bus routes and 11 rail lines. NJT provides nearly 223 million passenger trips each year.[19] See Figure 4.

NJT was the first passenger railroad to participate in the C^{3}RS demonstration, starting in October 2009. Approximately 1,700 employees from BLET, the American Train Dispatchers Association (ATDA), and UTU were eligible to submit close call reports when NJT joined C^{3}RS. All NJT-owned and/or operated territory was part of the initial pilot, including the Southern Tier and Pascack Valley Line but excluding Conrail and Amtrak territories not covered by C^{3}RS.[20]

Figure 4. NJT Map

[19]New Jersey Transit (2014). "About Us." (Web site) Newark, NJ. Accessed online: July 28, 2014. (http://www.njtransit.com/tm/tm_servlet.srv?hdnPageAction=CorpInfoTo)

[20] U.S. Federal Railroad Administration (2014). "Questions and Answers – New Jersey Transit." Accessed online August 4, 2014. (http://www.closecallsrail.org/faq_nj.aspx).

1.4.4 Amtrak

Amtrak is a nationwide passenger railroad which serves more than 500 destinations in 46 states and three Canadian provinces. It covers more than 21,200 miles of routes and has more than 20,000 employees. In 2011, an average of more than 831,000 passengers per weekday rode trains on Amtrak-owned infrastructure, dispatching, and/or shared operations, or rode commuter trains operated or maintained by Amtrak under contracts with local or regional agencies.

Amtrak began participating in the C^{3}RS demonstration in February 2011. Participants included employees from BLET and UTU working at a number of yards across the United States as defined in the original IMOU. No road crews were included during the baseline phase, although in later years their C^{3}RS program was expanded. Approximately 1,400 employees were eligible to submit close call reports. Amtrak employees in the Northeast Corridor (PRT East) yards and in the Chicago-, Miami-, Seattle-, and Los Angeles-area (PRT West) yards can report close call incidents to NASA (see Figure 5).

Figure 5. Amtrak C^{3}RS Locations

1.4.5 Implementation Differences among Sites at Baseline

There are three models of how C^{3}RS scope was distributed across railroad operations. In the first model, used by UP and CP, the scope of the PRT was limited to eligible crafts within a specific geographic or organizational area of railroad operations. The second model, used by NJT, was a PRT that covers all transportation-craft employees in the entire company. The third model, used by Amtrak, includes multiple geographic or organizational areas spread across the country. The transportation yard crafts were covered by the PRT, but the scope at baseline only covered crafts within the yards. UP, CP, and NJT all had a single PRT that analyzes the cases. Amtrak had two regional PRTs: PRT East and PRT West, which analyze cases falling within their specified geographic region.

1.5 Need for Rigorous Evaluation

FRA's decision to demonstrate close call reporting was ambitious and risky. Ambitious because the goal was to effect a major change in how the US railroad industry improves safety, the test

encompassed four different railroads, and they had planned for multi-year observations. It was risky because FRA was committing to spend a significant amount of money, time, and political capital on an experiment that could fail in numerous ways. Some ways the effort might not be succeed include: Would any railroad agree to participate? Would craft unions join? Would labor submit reports? Could confidentiality be preserved while providing enough information for a problem-solving process to occur? Would management and labor cooperate in problem-solving efforts? If they did, would solutions be generated that might improve safety? Could the railroads implement recommendations? Would safety and safety culture improve? Even if it was effective at the sites, would other railroads adopt it? If good things happened, would FRA understand enough about what it had done to replicate its work beyond the test sites? Could C^3RS be structured in such a way that it would become routine in the settings that adopted it?

These are important questions, and deserve answers through examination. Due to the importance of answering these questions, a systematic evaluation effort was planned using specialized expertise. It was important that the evaluation team was independent of implementation team, i.e. conducted by people who were not involved in implementing the program, to provide an unbiased assessment of C^3RS.[21] To conduct the evaluation FRA contracted with the Volpe Center to conduct the evaluation. To assist in the evaluation, the Volpe Center contracted with Jacobs Technology and its subcontractor partner, the Fulcrum Corporation.

1.6 Scope of Evaluation Presented in This Report

This report covers two broad topics: the methodology for the entire multiyear, multiple-railroad C^3RS evaluation, and baseline findings from each of the four participating railroads.

The demonstration pilot sites requested that their data be kept anonymous. Therefore, the findings presented in this report are constrained in a number of ways:

- Data and findings are not linked to specific railroads.
- The time periods of data collection are disguised because the timing of C^3RS at each location is publicly known.
- Only data that can be compared across at least three railroads are presented.

1.6.1 Evaluation Logic

Figure 6 is a "logic model" that identifies stakeholders, key activities, and impacts and shows the relationships between them. A logic model illustrates the relationships between a program's activities, its consequences, and its environment. This model provides clarity by listing the many different stakeholders and activities, what activities each stakeholder group was involved with, and what the outcomes and impacts were over time. stakeholders at each of the demonstration pilots, the FRA office of railroad safety personnel, and the Steering Committee were presented with this model. Their feedback was sought and incorporated where appropriate.

[21] FRA Office of Research and Development. (2013). Evaluation Implementation Plan: Office of Research & Development. DOT/FRA/ORD-13/47.

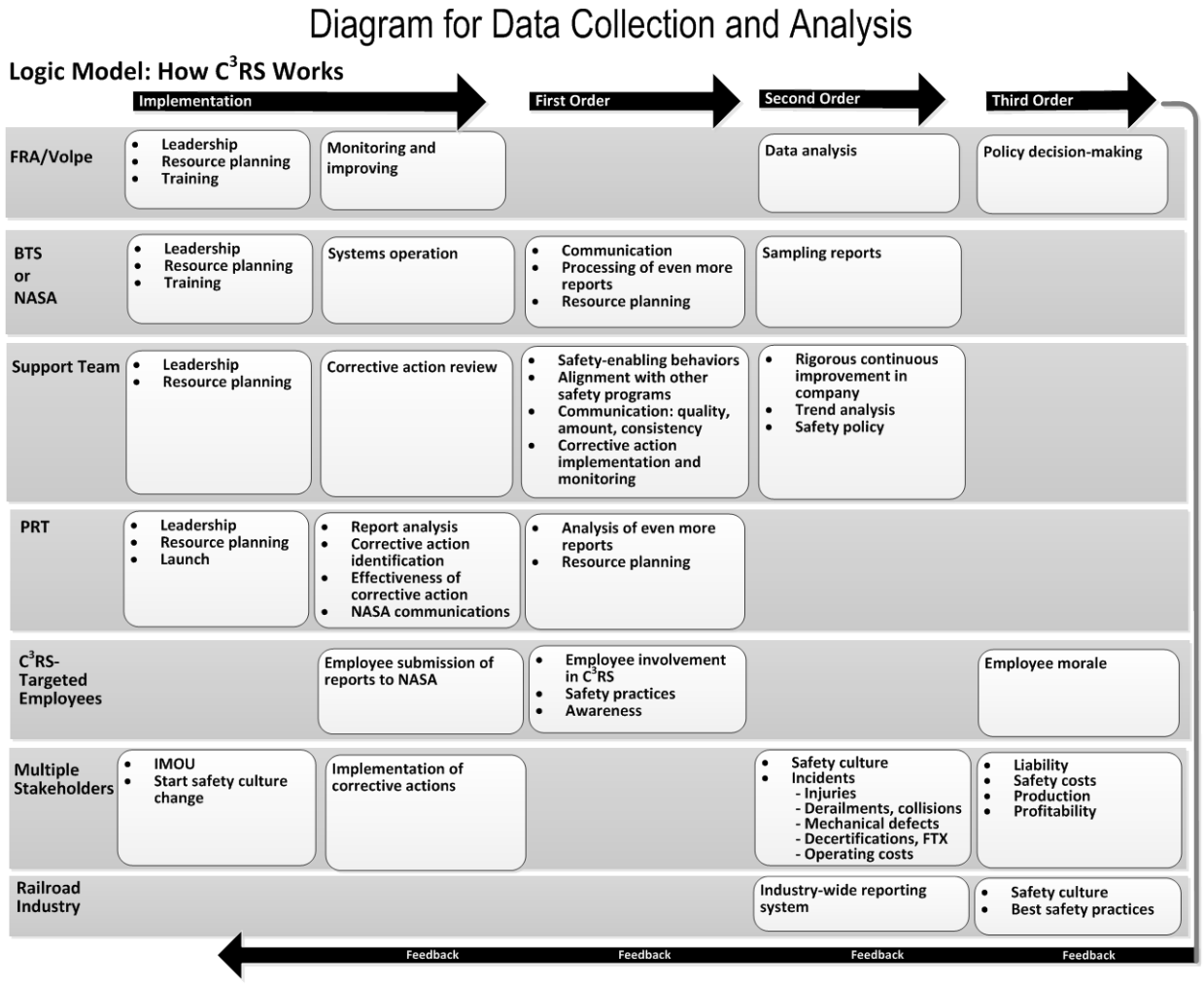


Figure 6. Stakeholder Version of C[3]RS Theory of Change Logic Model

Figure 6 was the result of an exercise conducted jointly by all stakeholders and the Evaluation Team. It displays a row for each involved group.[22] These rows highlight the role of each stakeholder group in the C[3]RS process and the outcomes that derive from each stakeholder group. In addition, the model recognizes that some outcomes will come from the collective actions of multiple stakeholder groups.

The four black arrows above the rows each signify a different phase of C[3]RS activity. The first phase, implementation, shows the activities that occur during the implementation of a C[3]RS demonstration. The three subsequent arrows indicate the expected first-, second-, and third-order impacts of the C[3]RS program. Feedback arrows appear at the bottom of the diagram, indicating that, at each stage, results may affect previous activity (which demonstrates that C[3]RS is an iterative process).

Each of the top five rows contains one stakeholder group: FRA/ Volpe, BTS or NASA, Support Team (consisting of railroad managers), the PRT, and C[3]RS targeted employees (who are

[22] Both NASA and the Bureau of Transportation Statistics (BTS) have served as C[3]RS third-party data collectors during the C[3]RS demonstration.

eligible to submit reports). The first two columns in each row contain activities each group does during the implementation of C3RS. The third through fifth column contain the first, second, and third level impacts from a successful implementation. The sixth row is concerned with multiple stakeholders

The overall logic of C^3RS is as follows: An early activity is negotiating the IMOU (row 6, column 1). During the IMOU process the local FRA inspector, participating unions, and railroad management negotiate an agreement that sets out the rights and obligations for the stakeholders involved. The process of negotiating the IMOU with the various parties requires trust and willingness to try out this new way of identifying and resolving safety issues. (Some railroads considered during the recruiting effort did not have sufficient trust and willingness for the IMOU process, so they were not selected as demonstration pilots.) As a result, the safety culture begins to change at this early step.[23] This initial shift in the safety culture was also seen in FRA's study on participative rules revision, where management became active and visible in its commitment to building a participative culture before the rules revision process began.[24] Each IMOU was based on a national-level MOU that was created by the National Steering Committee before the demonstration sites began.

As seen in column 1, a combination of FRA, the Volpe Implementation Team, and BTS/NASA supports implementation of C^3RS by providing leadership, resource planning, and training. Railroad management also provides leadership and allocates resources. The PRT then launches the C^3RS program.

Later in the implementation process, the Volpe Implementation Team, with advice from the Evaluation Team, takes actions to improve C^3RS (row 1 column 2). C^3RS-targeted employees submit close call reports (row 5, column 2). BTS or NASA processes the reports and communicates with other stakeholders, especially the PRT. The PRT analyzes the reports, determines root causes, and recommends corrective actions (row 4, column 2). Either local changes are made or action is referred for change at the corporate level, as appropriate (row 3, column 2).

The model also contains specific details of the long list of discrete changes that are hypothesized as first-, second-, and third-order impacts (see black arrows at the top of the diagram over columns 3-5). The first-order impacts in column 3 were hypothesized as follows: As employees use the system and there are no repercussions, trust builds, and the system is used more. As more people submit reports, BTS/NASA processes more reports. BTS/NASA therefore is required to plan for more resources to accommodate the increased reporting. BTS/NASA has the role of communicating high-exposure issues to the industry so that action can be taken across all railroads, thereby improving industry safety. The PRT also starts to process more reports and to communicate more with employees about safety issues and corrective actions. This increases the awareness of C^3RS-targeted employees and causes them to behave more safely. At the corporate level, practices that improve safety are endorsed and they may also think about how to align

---

[23] Lewin, K. (1947). Group decision and social change. In T. N. Newcomb and E. L. Hartley (eds.) *Readings in Social Psychology*. Troy, MO: Holt, Rinehart and Winston.

[24] Ranney, J., and Nelson, C. (2004). Impacts of participatory safety rules revision in U.S. railroad industry: An exploratory assessment. *Journal of the Transportation Research Board,* No. 1899, Transportation Research Board of the National Academies, Washington, DC, pp. 156–163.

C^3RS with other safety programs. In addition, corporate will promote and monitor communication about the C^3RS program.

Second-order impacts in column 4 that the Evaluation Team hypothesized were as follows: At the corporate level, C^3RS leads to rigorous, continuous improvement (CI) in the company. Managers can then start to look at overarching safety trends. Safety policies are also affected. BTS/NASA can start to sample reports, and the Volpe Implementation Team can perform data analysis on report trends and C^3RS impacts. An impact on safety outcomes, such as injuries, derailments, collisions, mechanical defects, and decertifications may be observed. This may lead to a decrease in operational costs. As the labor, management, and government members of the PRT work together, and as the cost of disciplinary hearings is avoided for reporters, the safety culture improves. C^3RS expands beyond a demonstration, and an industry-wide reporting system is implemented.

Third-order impacts in column 5 that the Evaluation Team hypothesized are as follows:

- FRA uses the knowledge about safety issues to affect decisions about policies.
- Railroad employees have increased morale after seeing all of the improvements.
- Multiple project participants contribute to the reduction in liability claims and safety costs.
- Improvements in production and profitability are thus observed.
- The railroad industry as a whole experiences an improvement in safety culture and learns about best safety practices.

2. The C^{3}RS Evaluation Plan

This section presents the evaluation plan in three parts:

- Evaluation questions
- Evaluation stakeholders
- Evaluation methodology

2.1 Evaluation Questions

The C^{3}RS evaluation is designed to answer three major questions:

- What conditions are necessary to implement C^{3}RS successfully? (Implementation Evaluation)
- What is the impact of C^{3}RS on safety and safety culture? (Outcome Evaluation)
- What factors help to sustain C^{3}RS over time? (Sustainability Evaluation)

2.1.1 Implementation Evaluation

The Implementation Evaluation serves two purposes:

- It provides rapid feedback to program implementers to help them determine midcourse corrections.
- It provides guidance to people contemplating similar programs in the future.

To supply this feedback and guidance, the Implementation Evaluation addresses three questions:

- Is the program being implemented according to plan?
- Is implementation effective?
- Are deviations from the plan articulated clearly and implemented effectively?

2.1.2 Outcome Evaluation

This study defines "outcome" as the intended or unintended long term effects of the program.[25] The terms "outcomes" and "impacts" are used interchangeably in this report. Successful implementation is only useful if the innovation being tested has value. Conversely, if the implementation isn't adequate, then the innovation cannot be expected to show value. One must differentiate between program and implementation failure. If the evaluation indicates that the implementation succeeded, the outcome evaluation should assess that value by asking three questions:

- What has happened as a result of the program having been implemented?
- What were the short-, intermediate-, and long-term impacts?
- Were there any unexpected or unintended impacts?

[25] Evaluation Implementation Plan: Office of Research & Development" FRA/ OSD (Nov 2014). DOT/FRA/ORD-13/47.

2.1.3 Sustainability Evaluation

Successful implementations and successful outcomes do not guarantee that the project has longevity. Specific evaluation activities are needed to answer questions about sustainability such as:

- Did C^{3}RS expand beyond its initial test locations at the four participating railroads?
- Did C^{3}RS continue at the participating railroads after the initial test period ended?
- How can continuation or extinction be explained?

2.2 Evaluation Stakeholders

The C^{3}RS evaluation includes many stakeholders including FRA; railroad senior management; the PRT; C^{3}RS-targeted (front-line) employees; the industry as a whole; organizations that provide services to C^{3}RS, such as an implementation team led by the Volpe Center; and trusted third parties (BTS, NASA) who receive close call reports.

- *FRA Office of Research and Development (ORD):* FRA has an obligation to improve safety in the railroad industry so they fund research and development aimed at improving safety. C^{3}RS is one of ORD's important efforts. Therefore, the agency needs evaluation information about C^{3}RS as to whether and how it should disseminate C^{3}RS or C^{3}RS-like programs within the industry. It needs to know the impacts at the demonstration sites to see if C^{3}RS was worth the cost. It also needs information to justify budget requests to Congress.
- *FRA Office of Railroad Safety's Risk Reduction Program*: The Office of Railroad Safety is responsible for rolling out the C^{3}RS program to the industry. The evaluation provides data to improve implementation, outcome, and sustainability.
- *Railroad senior management (PRT Support Team):* Senior managers need to know whether to support continued participation at their demonstration pilot site. They also need to know whether to support the program's dissemination more broadly across their railroad on the basis of the effectiveness and efficiency of the corrective action implementation process.
- *PRT:* The PRT needs feedback on the effectiveness of its activities to enable it to make improvements and to guide dissemination of successful activities across the railroad. It also needs the evaluation process as a mechanism to confidentially share its perspectives about the effectiveness of other C^{3}RS participants with the Volpe Implementation Team.
- *C^{3}RS-targeted employees*: Front-line employees who are eligible to report to C^{3}RS need to know whether to support continued participation and the program's dissemination across the railroad. They also need a venue in which they can confidentially share their views about the C^{3}RS demonstration.
- *Railroad industry:* Labor and management throughout the industry have interests similar to those of FRA. They need guidance as to whether they should act to implement C^{3}RS and how they should go about it. The Steering Committee is one venue for providing feedback and collecting input from the railroad industry.

- *Volpe Implementation Team:* Implementation Team members need feedback about the effectiveness of their efforts to ensure success at the four pilot sites as well as any sites that may join C^{3}RS in the future.
- *BTS and NASA:* The third-party data collectors need feedback about the effectiveness of their efforts to ensure success at the four pilot sites as well as any sites that may join C^{3}RS in the future.

2.3 Evaluation Methodology

This evaluation should provide important C^{3}RS stakeholders with guidance and answer questions about the program's implementation, outcome, and sustainability. To accomplish these objectives, the evaluation was developed with the research design logic and the types of information needed to answer the evaluation questions in mind.

2.3.1 Evaluation Design

Figure 7 is a schematic overview of the research design. It features the data and the potential comparisons between the different participants.

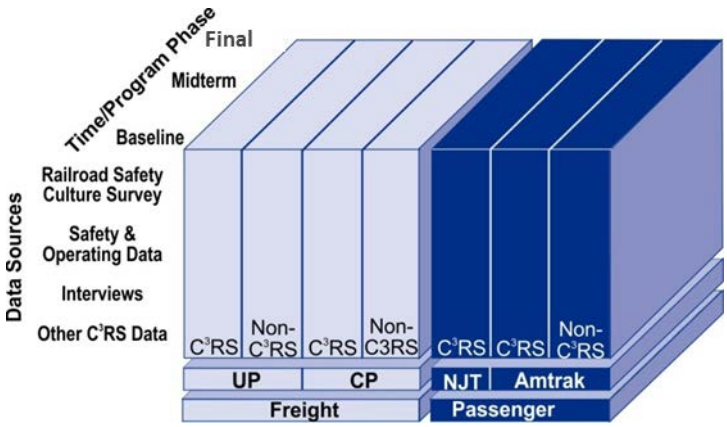


Figure 7. Schematic View of Evaluation Methodology

- The participating railroads are UP North Platte Service Unit, CP Chicago Service Area Road Territory, NJT (entire railroad), and Amtrak Yards. During the baseline period, only a subset of crafts chose to participate at each site.
- The sequence of entry into the program is displayed from left to right.
- Two participants are freight railroads and two are passenger railroads, because the C^{3}RS program designers wanted to account for the possibility that programs such as C^{3}RS might work differently in these different operating settings.
- One railroad, NJT, was a "whole company" intervention. For the other three railroads, comparisons between C^{3}RS and non-C^{3}RS parts of the company are possible.
- Multiple sources of qualitative and quantitative data are available. (The Evaluation Team always uses as many different sources of data as possible to estimate any given variable or to test any relationship.)
- Some data are available historically from the baseline through the final phases of the evaluation. For instance, corporate data can provide information on derailments for many years before the implementation of C^{3}RS. Other data allow comparisons only from the time of the introduction of C^{3}RS to several years after the program has begun. This is true, for instance, for culture survey information or interview data with PRT members.
- Reported findings will be defined by the three evaluation phases–baseline, midterm, and final. In the early stages, these reports will be unique to each participating railroad.

- Specific findings will not be shared because the agreement with the railroads states that their findings will be anonymous. As time goes on, collecting data across multiple railroads will allow aggregation of that data into cross-site findings, which can be shared.

Therefore, over time, this design will provide a rich set of cross-site and historical comparisons that employ both qualitative and quantitative data. Given the availability of data and the relevance of different data elements to different evaluation questions, Figure 7 does not imply that all data and all possible comparisons will be included in the analysis. However, the figure does convey which comparisons are possible and provides an overview of the various data sources that can be brought to bear on them.

2.3.2 Evaluation Phases

For each pilot, the evaluation process is broken into three phases: baseline, midterm, and final. Each phase consists of data-collection, analysis, and feedback to the demonstration pilot site.

The *baseline* phase starts when C^{3}RS is implemented at a given site. First, the railroad's initial safety culture and labor-management relationships are examined to learn the organization's initial values, which will be compared against the midterm and the final values to assess the changes over time. The initial views of the C^{3}RS program are also measured, which allows the effectiveness of the initial orientation and training to be assessed during the implementation evaluation. Some sustainability issues may have begun to appear as well. The baseline phase at each site covered about one and a half years after C^{3}RS reporting began and each of the four railroads started this phase between 2007 and 2012, depending on when the railroad joined the C^{3}RS demonstration.

The *midterm* phase continues with a second survey administration and second interview data collection that includes corporate and safety data, corrective action data, and summary data for the frequency and types of C^{3}RS reports. Safety data are collected at midterm but because they are archival, they include the baseline data from several years before C^{3}RS began and allow a comparison of baseline with midterm. Thus, at midterm, the first hints of outcome evaluation in terms of impact on both culture and safety are more likely to be seen. This phase helps to reveal implementation issues and provides an opportunity to adjust and improve the program. Sustainability issues are also becoming more apparent at midterm.

The *final* phase repeats the midterm data collection with additional information collected about C^{3}RS at the end of the evaluation, thus completing the impact evaluation and providing longer-term insight into sustainability.

Table 2. C^{3}RS Phase vs. Evaluation Type

	C^{3}RS Phase		
Evaluation	Baseline	Midterm	Final
Implementation	■	■	
Outcome		■	■
Sustainability	■	■	■

Each evaluation phase deals with a different type of information about the program. Data analysis and interpretation take place with the understanding that each type of evaluation can be relevant along many points in the project's life cycle, as depicted in Table 2. Information on how well a program is implemented and run (implementation evaluation) is a focus earlier in the evaluation. Outcome evaluation begins when outcomes may begin to appear, about the time of the midterm data collection phase. Sustainability evaluation is relevant throughout the project because factors related to a program's staying power can express themselves at any time.

2.3.3 C^3RS Logic Model: Technical View

Logic models are analogous to maps because, in both cases, they can employ different forms and scales for different purposes. Figure 6 is a useful logic model that shows the roles of various stakeholders in the overall C^3RS process flow. It maps activities and outcomes to specific stakeholder groups. To create a technical guide for evaluation, another logic model form is needed that has a detailed view of direct and indirect relationships, feedback loops, the time sequence, and also identifies what should be measured.

A bird's-eye version of the technical logic model is presented in Figure 8. The model is divided into three broad sections: green, blue, and yellow (or in black-and-white printing, light, medium, and dark gray). Green boxes (dark gray) are activities and relationships that deal primarily with implementation and successful operations, such as establishing input from BTS and effective joint labor-management problem-solving committees. Blue boxes (medium gray) are a set of intermediate outcomes that the Evaluation Team originally believed were precursors to the bottom-line effects that the team especially cared about. As a prime example, stakeholders who developed the model believed that C^3RS would have to prove itself by dealing with relatively minor issues before more serious close call reports were submitted. Yellow boxes (light gray) are the important impacts that stakeholders hoped would emanate from C^3RS in relation, for example, to accidents, liability costs, additional safety-culture improvement (beginning with the IMOU), and FRA policy. The long rectangles at the bottom are factors that pervade the entire C^3RS process and relate to program sustainability and changes in the social, economic, and policy environments that may affect the program.

Each element of the model is numbered to create an index for the Evaluation Team, which can assist the Team in sorting variables and organizing the analysis. Due to the complexities which accompany real-world data collection, not every element of the model is measured equally well or at all. However, even when an element is not measured, the existence of the element serves as a guide to where data are strong and weak with respect to testing the model. The advantage of this perspective is that it shows the whole scope of the evaluation buts a result, that individual elements are very small. Another version of this model enlarges each section to a scale that makes the details readily perceivable. These larger sections are presented in Appendix A. Logic Model.

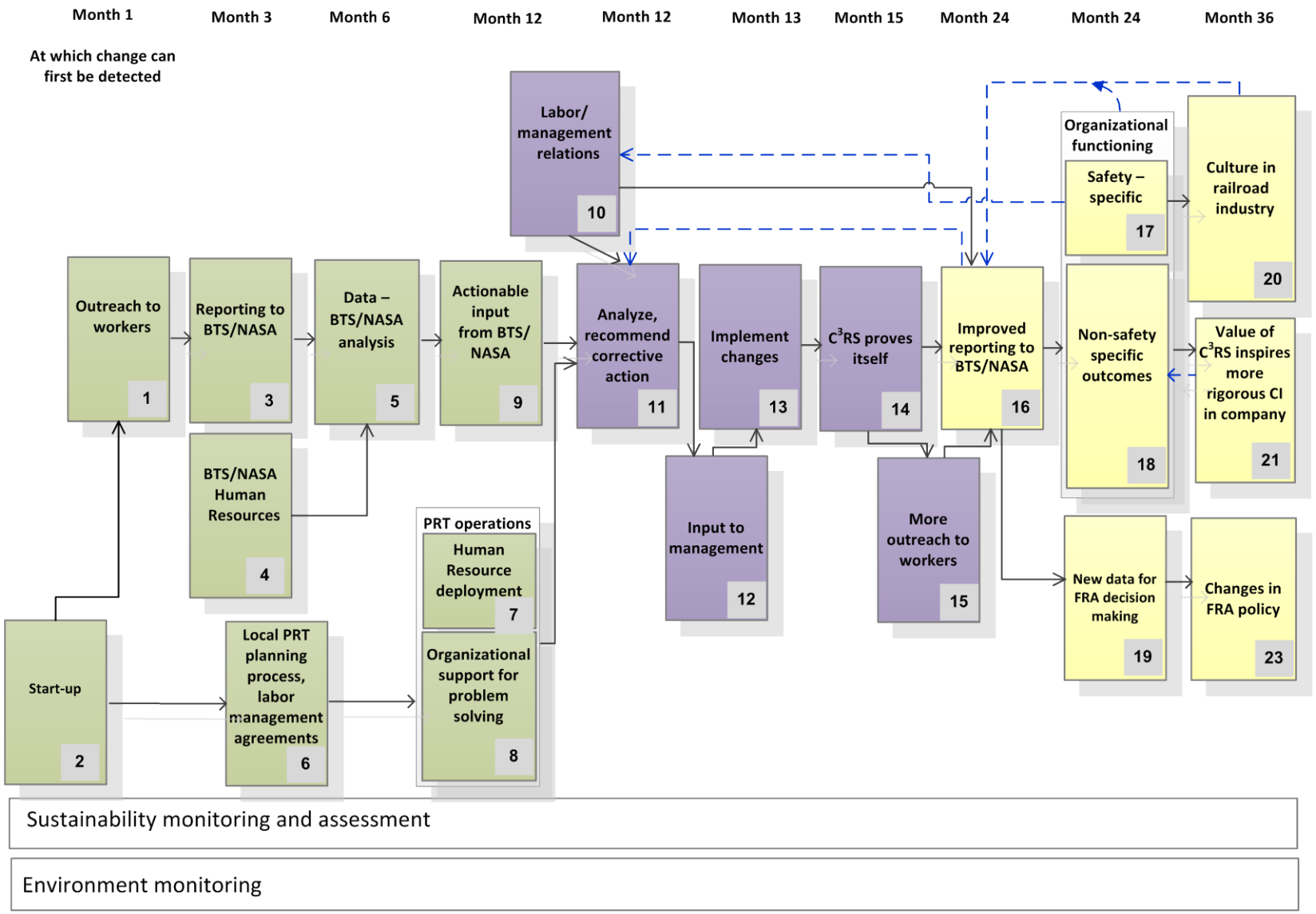


Figure 8. Technical Logic Model

2.3.4 Relationship among Design, Metrics, and Logic Model

The evaluation design identifies the quantitative and qualitative metrics that will be collected and the kinds of comparisons that can be made among the various treatment and control groups. Additionally, the technical logic model provides an index for organizing and interpreting data. For instance, all data related to safety culture can be indexed to those parts of the model dealing with that construct.

Table 3 summarizes the methods of data collection, types of data, and evaluation questions that each data type helps to answer.

Table 3. Methods of Evaluation and Their Relation to Logic Model Factors

Method	Type	Evaluation Questions
Field notes	Qualitative	Implementation, outcome, sustainability
Project records	Qualitative	Implementation, outcome, sustainability
C^3RS reporting rates	Quantitative	Implementation, sustainability
Interviews	Qualitative	Implementation, outcome, sustainability
Railroad Safety Culture Survey	Quantitative	Implementation, outcomes
Corporate safety data	Quantitative	Outcome
Corrective action data/summary multiple cause incident analysis (MCIA)[26] data	Qualitative and quantitative	Implementation, outcome

[26]The MCIA tool is used by the PRTs to analyze C^3RS cases. BTS provides a redacted summary of the results of the PRT's analysis with the MCIA tool to the Evaluation Team. More information on the type of data received and how the data are used will be included in the midterm report.

3. Data Collection and Analysis Methods

This section provides details on how the data were collected and analyzed at baseline.

3.1 Confidentiality

The railroad carriers considered any data that was provided to the evaluation team for evaluation purposes as company confidential. Therefore, the Volpe Center worked under a nondisclosure agreement (NDA) and each participating railroad limited the distribution of any data collected and findings that were summarized during the evaluation. While the Volpe Center had editorial control over reports that were made public, the demonstration pilots reviewed drafts of reports and offered perspectives in advance of publication. Their ideas were welcomed and considered in the final drafts of the reports.

In addition, the Evaluation Team contractor developed a data protection plan that was carefully reviewed and approved by Volpe. Highlights include:

- The Volpe Center operated under a NDA with each participating railroad.
- Volpe included similar data protection clauses in for all contractors working on the project.
- All project participants received CIPSEA training for access to BTS data.[27]
- Files were stored "off-network" in encrypted True Crypt drives, encrypted external hard drives, and a GSA approved safe.
- Draft reports were e-mailed to evaluation team members with password protection.
- Interview notes did not contain exact dates or personally identifying information.
- All project records used a notation marking them as company confidential, "For Official Use Only" (FOUO). This document may be exempt from mandatory disclosure under FOIA. Company Confidential."

3.2 Interviews and Other Qualitative Data

The evaluation included these types of qualitative data:

- *Phased interviews:* Semi-structured interviews with railroad employees and managers who were within and outside of the program. These interviews occurred once during each of the three phases of the evaluation (baseline, midterm, final). Interviewees were asked about safety, labor-management relations, how C^3RS program operations could be improved and observed impacts of C^3RS. These interviews took place on-site, and occurred face to face.
- *Implementation interviews:* Semi-structured interviews with key stakeholders, such as PRT members, senior managers, labor officials, FRA, the Volpe Implementation Team, BTS, and NASA. These interviews occur throughout the evaluation. Interviewees were

[27] Title V – Confidential Information Protection and Statistical Efficiency. 116 Stat. 2962. Public Law 107–347 – Dec. 17 2002. Available at http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/cipsea/cipsea_statute.pdf

asked about key events related to the implementation, functioning, and sustainability of C^3RS.

- *Project documents:* Documents created by C^3RS program participants; for example, newsletters, brochures, and presentations.
- *Field notes:* Notes taken by the evaluation team from project related meetings, conference calls, email communication, and informal discussions. Examples include monthly check-in meetings conducted by the Volpe Implementation team with each pilot site.

For all interviews, confidentiality was protected by excluding personally identifying information from the notes, marking documents with an FOUO notation, encrypting electronic files, and shredding paper notes. Interviews typically were done in pairs, with one primary interviewer and one note taker.

3.2.1 Data Collection

Phased Interviews

The semi-structured, phased interviews are performed once during each of the three phases of the evaluation—baseline, midterm, and final—at each demonstration pilot site. Interviews were conducted in person with pilot site workers who were eligible to report close calls and the managers who were affiliated with the project. A few of the workers were affiliated with the C^3RS PRT but most were not. The primary purpose of these interviews was to gain insight about safety culture, safety impacts, C^3RS acceptance, strengths, and weaknesses from workers eligible to report but not directly involved in the C^3RS program. The Evaluation Team requested that interviewees include a wide range of tenures and people who were supportive as well as skeptical of C^3RS. Railroad site managers sent employees who were available during the days of the site visit to participate in the interviews. The typical interview duration was 30 minutes.

The baseline-phase version of the interviews included three main topics of discussion:

- *Safety issues and ways to improve safety:* Questions about safety issues that C^3RS might address, and about safety programs at the pilot site that might be alternative explanations for any observed impacts
- *Labor-management relations:* Questions about the cooperation between labor and management.
- *C^3RS program behavior:* Questions about the understanding and views of C^3RS.

The complete list of questions is included in *Appendix B. List of Phased Site Interview Questions*.

At each site, 10 to 15 interviews were requested. When the evaluation team arrived at the railroad site, the person from the PRT assigned to help with scheduling did so. Sometimes more than one person was sent to an interview. A total of 52 people participated in the baseline interviews across the four sites.

Implementation Interviews

Implementation interviews were conducted with a variety of key stakeholders who were "close observers" of the C^3RS pilots. Interviewees included PRT members (labor, management, FRA),

the Volpe Implementation Team, senior railroad management, labor representatives, the third party (BTS and NASA, as applicable), and FRA ORD and Office of Railroad Safety. Essentially, interviews included all groups who were either involved in some aspect of C^3RS, or who could be considered "interested parties" to the C^3RS process. These interviews helped collect information for improving the implementation of C^3RS and sustainability was also discussed. The typical interview duration was 30 minutes.

The interviews were semi-structured and consisted of several primary questions. In some cases, certain questions were eliminated if they were not appropriate for the interviewee. Topics were as follows:

- Key events that affect the running of C^3RS
- How the PRT and Support Team were doing and how program participants could help improve the operation of C^3RS
- Important events outside program that might impact C^3RS
- Factors affecting the sustainability of C^3RS
- Recruiting of new railroads to C^3RS

A full list of questions can be found in *Appendix C. List of Implementation Interview Questions*.

At each site, 24 interviews per year were planned, including additional interviews with other stakeholders. E-mail and phone calls were used to request interviews. Interviews were most frequently conducted with individuals, but some were with small groups. Occasionally the PRT as a group was interviewed. About 110 people provided implementation interviews during the baseline phase of the evaluation.

Other Qualitative Data/Field Notes

The Evaluation Team also collected other types of qualitative data that were available, including: Meeting notes from the Volpe Implementation Team, which were taken during Steering Committee meetings and PRT monthly check-in conference calls; information about PRT activities from railroad newsletters; and e-mails concerning lessons learned and project status.

3.2.2 Qualitative Data Analysis

The Evaluation Team organized and analyzed the qualitative data[28] using Atlas.ti software.[29] A systematic process was employed:

1. Each document was added to the database and categorized in three ways: railroad (UP, CP, NJT, Amtrak); interviewee class (for example, labor, management, FRA); and data type (for example, baseline phased interview, implementation interview, field notes).

---

[28] In conducting this analysis we employed standard, accepted methods of qualitative data analysis. See

Patton, M.Q. (1987). How to use qualitative methods in evaluation. Newbury Park, CA: Sage.

Patton, M.Q. (2002), Qualitative research and evaluation methods (3rd ed.). Thousand Oaks, CA: Sage.

[29] Atlas.ti helps researchers to extract meaning from narrative data by allowing them to impose content categories, to rearrange and revise the categories over the course of an analysis, to nest and overlap groupings, and to sort out multiple categories. Available at http://www.atlasti.com/

2. Next, each interview or field-note text was read and broken into comments. The comments were assigned to codes derived from the technical logic model (Figure 8). As a quality control mechanism, coding was periodically reviewed by the Evaluation Team to ensure reliability.
3. Additional codes were created as needed, based on additional reading of the text and our developing understanding of the program. A full list of codes available at the time of this report is found in Appendix D. List of Qualitative Data Codes.
4. When reports were written, the Evaluation Team used the Atlas.ti tool to run queries for the most frequent codes. Interview comments under each frequent code were read again, summarized, and discussed among team members.

3.3 Railroad Safety Culture Survey Methods

3.3.1 Survey Items

The C^3RS Railroad Safety Culture survey included demographics, railroad safety culture scales from previous Volpe Center implementations and research, and specific C^3RS items. These survey sections are described below. Some questions had slight variations for labor and management. All questions were multiple choice, except for an open-ended item asking them to "Please add any other comments you have about the safety culture and the C^3RS program on the lines below," where respondents could write anything they chose at the end of the survey.

Demographics

The first part of the Railroad Safety Culture Survey asked each respondent to answer demographic items. The demographic information allowed the Evaluation Team to examine differences among respondents. The demographic questions focused on:

- Job type: labor or management.
- Job category: trainman/conductor, switchman, locomotive engineer, hostler, yardmaster, dispatcher.
- Location: road, yard, both.
- Shift: days, afternoons, nights, variable.
- Gender: male or female.
- Age: grouped to keep responses anonymous.
- Tenure: grouped into time-span categories in order to keep responses anonymous.
- Survey: respondents were asked if they had filled out this survey previously.

Railroad Safety Culture Scales

To examine safety culture scales and items, the Evaluation Team used FRA/Volpe's Railroad Safety Culture Survey as the base case and added or deleted scales based on the C^3RS situation. All scales that were used have been used previously and tested for validity and reliability. The safety culture scales and sources are summarized in Table 4. Not every scale listed was used at all three railroads. For example, some surveys were shortened on the basis of time constraints, with some scales eliminated or merged on the basis of lessons learned from past Volpe Center research. Also, the craft-specific scales were applicable only to the craft for which they were

written. For more detail on the source of the items, including additional references, see Section 7 References.

Table 4. Railroad Safety Culture Survey Scales and Definitions

Scale	Description	No. of Items	Item Source[30]
Organizational Concern for Employees	The extent to which employees believe that the organization is concerned about their needs	3	Eisenberger, R., Huntington, R., Hutchison, S., et al. (1986)
Labor-Management Relations	The extent to which employees believe that there is cooperation between labor and management in the organization	6	Dastmalchian, A., Blyton, P., & Adamson, R. (1989)
Organizational Fairness During Change	The extent to which employees believe that the organization was fair to them during past attempts to make changes	6	Colquitt, J.A. (2001)
Supervisor-Employee Relationships	The extent to which employees perceive the working relationship between themselves and supervisors to be strong	7	Wayne, S.J., Shore, L.M., & Liden, R.C. (1997)
Supervisor Fairness	The extent to which employees perceive that their direct supervisors treat them fairly	7	Niehoff, B.P., & Moorman, R.H. (1993)
Raising Concerns with Supervisors	The extent to which employees perceive that their supervisor is open to their raising safety concerns	7	Hofmann, D.A., & Morgeson, F.P. (1999)
Management Safety	Employees' perceptions of management's attitude toward safety	11	Mueller, L., Da Silva, N. Townsend, J.C., et al. (1999)
Work-Safety Priorities	Employees' perceptions of how committed the organization is to safety over productivity	5	Mueller, L., et al. (1999)

[30]This is the primary reference. For more detail and additional references, see Section 7 References.

Scale	Description	No. of Items	Item Source[30]
Safety Briefings	The extent to which safety briefings are part of doing business	1	Dedobbeleer, N., & Beland, F. (1991)
Respectful Workplace	The extent to which employees perceive that coworkers treat each other respectfully	4	Magley, V.J., Davies-Schrils, K., & Walsh, B. (2007)
Helping Behavior	The extent to which employees perceive that coworkers perform extra activities to help each other	4	Naumann, S.E., & Bennett, N. (2000) Organ, D., & Konovsky, M. (1989)
Propensity to Safe Behavior	The extent to which employees perceive that coworkers work safely	3	Simard, M., & Marchand, A. (1997)
Coworker Safety	The extent to which employees perceive that coworkers encourage each other to work safely	5	Mueller, L., Da Silva, N. Townsend, J.C., et al. (1999)
Safe Behaviors – Road Crews – Cab Red Zone	Self-reported extent to which road crews practice safe behaviors in the cab red zone	6	UP Code of Operating Rules, checked by UP subject-matter experts Hofmann, D.A, & Stetzer, A. (1996)
Safe Behaviors – Yard Crews – Switching	Self-reported extent to which yard crews practice safe behaviors during switching	6	UP Code of Operating Rules, checked by UP subject-matter experts Hofmann, D.A., & Stetzer, A. (1996)
Safe Behaviors – Dispatchers	Self-reported extent to which dispatchers practice safe behaviors during dispatching	6	Hofmann, D.A., & Stetzer, A. (1996)
Alertness	The extent to which dozing while doing a variety of tasks is probable	8	Johns, M.W. (1992)

C^{3}RS-Specific Questions

The survey also included some questions specially developed for the C^{3}RS evaluation, covering these topics:

- Safety changes at your railroad
- Awareness of conditions that could lead to accidents
- Belief that accidents can be prevented by changes to systems and procedures
- Belief that joint labor-management teams could work together
- C^{3}RS reporting
- Understanding of C^{3}RS
- Things keeping you from reporting to C^{3}RS
- Impacts from C^{3}RS
- Confidence in confidentiality

Comments

Respondents were allowed to provide a written answer to the question "Please provide comments about safety at your site and the C^{3}RS program."

3.3.2 Survey Administration

The survey was administered by BTS on multiple days and over a 24-hour period to cover multiple shifts. Each time the survey was administered, the Evaluation Team provided BTS with a codebook containing all items and multiple-choice-response options. For each survey administration, BTS prepared the fixed-choice survey forms using eListen software, planned times and locations with the PRT, traveled to the site to distribute and collect the surveys, processed the data, performed data quality checks, and provided the data to the Evaluation Team. The baseline survey administration occurred at UP, CP, and NJT (see below for information regarding Amtrak) close to the time that C^{3}RS was launched at each site. Each demonstration railroad will be administered the survey two more times (in the midterm and the final evaluation phases) with the same procedures.

BTS worked closely with the PRT to administer the survey and maximize the response rate. BTS first talked to the PRT and the manager of the location to schedule enough people to help proctor the survey. BTS sent two team members to administer the survey so multiple shifts could be accommodated at multiple locations over several days. (The exception to this process was the baseline data collection at UP. There, BTS worked with the site to prepare for data collection but did not have people present during data collection.) An example of the cover letter for the survey informing respondents about the purpose and confidentiality is included in *Appendix E. Example Survey Cover Letter*.

On the day of administration, survey proctors were stationed close to the locations where employees initiated their workday. At one site the survey was given in a conference room that the employees would walk by on their way to collect the paperwork for their shift, while at another site, the survey tables were set up near the locker room. Employees in some locations were called in early while in other locations they completed the survey after their shift. At some locations, the railroad offered employees a snack to show that they appreciated their

participation. The survey proctors included people from BTS and the PRT. Some proctors traveled to collect surveys from employees at more remote locations. In cases where these tactics did not yield a large enough sample, a second round of survey data collection took place.

After the survey was administered, BTS processed the surveys at their office. One person used a Scantron machine to scan each survey and in some cases the data was manually entered. Next, a second person validated a sample of the surveys by checking for a match between the paper survey and the scanned result. Handwritten comments to the open-ended question at the end were manually typed into the data set. The BTS confidentiality officer reviewed handwritten survey comments to remove any names or personally identifying information. Once all the data was entered in the database, BTS saved it in an Excel file and sent that file to the Evaluation Team.

One demonstration site, Amtrak, had a parallel safety culture effort underway and it was administering its own safety culture survey. Due to concerns about low response rates and survey fatigue, the C^3RS version was not administered to that site. The survey that Amtrak had administered was similar to the Railroad Safety Culture Survey and therefore could provide insight into the culture at baseline. However, since the scales were not identical to the C3RS survey scales, they could not be included in the formal cross-site comparison analysis.

3.3.3 Survey Analysis

By Railroad

To establish baseline scale scores for the first three pilot sites (UP, CP, and NJT), the baseline survey was analyzed for each individual railroad. The C^3RS-specific questions provided information about respondents' initial views on the value and practicality of C^3RS and the effectiveness of the initial training and publicity. Midterm and final surveys will be analyzed for changes in safety climate scales and changing views about C^3RS.

Even though the scales that were used are based on extensive testing, scale reliabilities were tested again to ensure that they would hold up in the current setting. First, all scales were transformed so that high scores represented the "positive safety culture." Cronbach's alpha was conducted to test the reliability of the scales. In general, alpha scores of over 0.7 were considered adequate.[31] Then, factor analysis was used to ensure that items were loaded on the same constructs as they had in previous research.

Scale scores were calculated separately for labor and management as past experience indicated that they tend to perceive safety culture differently, with management seeing it more positively than labor. A variety of multivariate and univariate statistical tests were used to check for overall and specific differences in factors such as tenure, location, and type of work.

For the C^3RS-specific questions, the Evaluation Team examined levels of support and knowledge of the C^3RS program by comparing demographics and labor-management differences.

---

[31]Rosenthal, R., & Rosnow, R.L. (1991). Essentials of behavioral research: Methods and data analysis (2nd ed.). New York: McGraw-Hill.

The open-ended survey comments were content-analyzed for trends with a method similar to that used for the interview data. After reading through the comments, the Evaluation Team determined the relevant set of codes/themes related to the logic model. Then, each comment was coded. The percentage of interviewees with comments related to each code was calculated, and the Evaluation Team determined summary results.

Comparisons across Railroads

Since three of the pilot sites used almost identical questions, a comparison of scale scores across railroads is possible and will be summarized here. There was some minor tailoring of the surveys to make them understandable at each railroad, but the constructs remained intact. In addition, some scales, such as Safe Behaviors for Dispatchers, were not applicable to all railroads.

The three sites have eight scales in common, which allowed the Evaluation Team to perform cross-site comparisons. Given the high number of comparisons, it was possible that many findings would be false-positives, that is, differences among groups were statistically significant but the groups did not actually differ because the probability of false-positive accumulates as many different comparisons are done. A two-stage process was used to address this threat to correct interpretation. First, an overall multivariate analysis was performed across all scales comparing the three railroads. Use of the multivariate test allowed a single significance test to provide confidence that some differences did in fact exist among the railroads with respect to the scale scores. Further tests were conducted to examine the differences. Significance across railroads for specific scales was determined by applying a Bonferonni correction to the tests that were used, thus adjusting each significance determination for the total number of tests performed.[32]

The comments that were handwritten in response to the open-ended question were also compared across pilot sites to identify similar and different themes. Content-analysis methods similar to those for interviews described in the previous section were used for this analysis.

Comparisons over Time

Because this report is concerned only with baseline findings, data on scale changes over time are not included here. Future technical reports will provide these data, using the same general approach described above but with added analysis to detect time-series changes.

3.4 C^{3}RS Reporting Rates Methods

3.4.1 C^{3}RS Reporting Rates Data Collection

BTS and NASA provided the Evaluation Team with the number of monthly close calls submitted by eligible C^{3}RS employees at the demonstration sites. This information shows labor's participation in C^{3}RS over time at each site.

[32]Wikipedia provides a good description of the Bonferonni correction. See http://en.wikipedia.org/wiki/Bonferroni_correction

3.4.2 C³RS Reporting Rates Data Analysis

As part of the effort to compare railroads, the first 12 months of the C³RS program at each railroad were charted. It is important to note that Month 1 did not occur at the same time for Site 1 and Site 2 but instead represents the first month of reporting for Site 1 and Site 2, respectively.

3.5 Corporate Safety Data

3.5.1 Corporate Safety Data Collection

Future technical reports will analyze and report on safety occurrence data. These data are not appropriate to include here because their role is to show changes in safety. Therefore, baseline data must be compared at the midterm and/or final phase when C³RS has been operating for a while. Examples of types of data that may be collected are:

- Corporate data
  - Incidents
  - Decertifications
- PRT data – list of cases
  - Month that they occurred
  - Category
  - Root cause
  - Contributing factor
  - Corrective action recommendation
  - Corrective action status

4. Summary of Baseline Data

Since the Evaluation Team has agreed to maintain railroad information confidentiality, the four demonstration sites are randomly assigned names (Site 1, Site 2, Site 3, and Site 4) in the Results section of this report. Since these designations are consistent throughout the report, each name always refers to the same demonstration pilot site.

4.1 Summary of Railroad Safety Culture Survey Data and Findings

This section contains an overview of the Railroad Safety Culture Survey data collected during this baseline phase and it also provides associated conclusions. A detailed analysis of results is available in *Appendix F. Railroad Safety Culture Survey Results*.

The Evaluation Team examined baseline safety culture results at the three demonstration sites where BTS administered the Railroad Safety Culture Survey. A count of responses to the baseline Railroad Safety Culture Survey is shown in Table 5 below.[33]

Table 5. Number of Baseline Survey Respondents

Railroad	No. of Labor Respondents	No. of Management Respondents
UP	427	72
CP	240	28
NJT	791	21
Total	1,458	121

4.1.1 Scale Scores

The results of the Railroad Safety Culture Survey are conveyed through a graph that has three scales (Organizational, Managerial, and Coworker) and represents managers with the letter M and labor with a capital L. Figure 9 contains results for each railroad. Figure 10 shows the average values for all survey respondents across the three sites.[34] Respondents could respond with a number from 1 through 5—a value of 1 indicates the most negative and 5 indicates the most positive perception of safety culture. The conclusions are as follows:

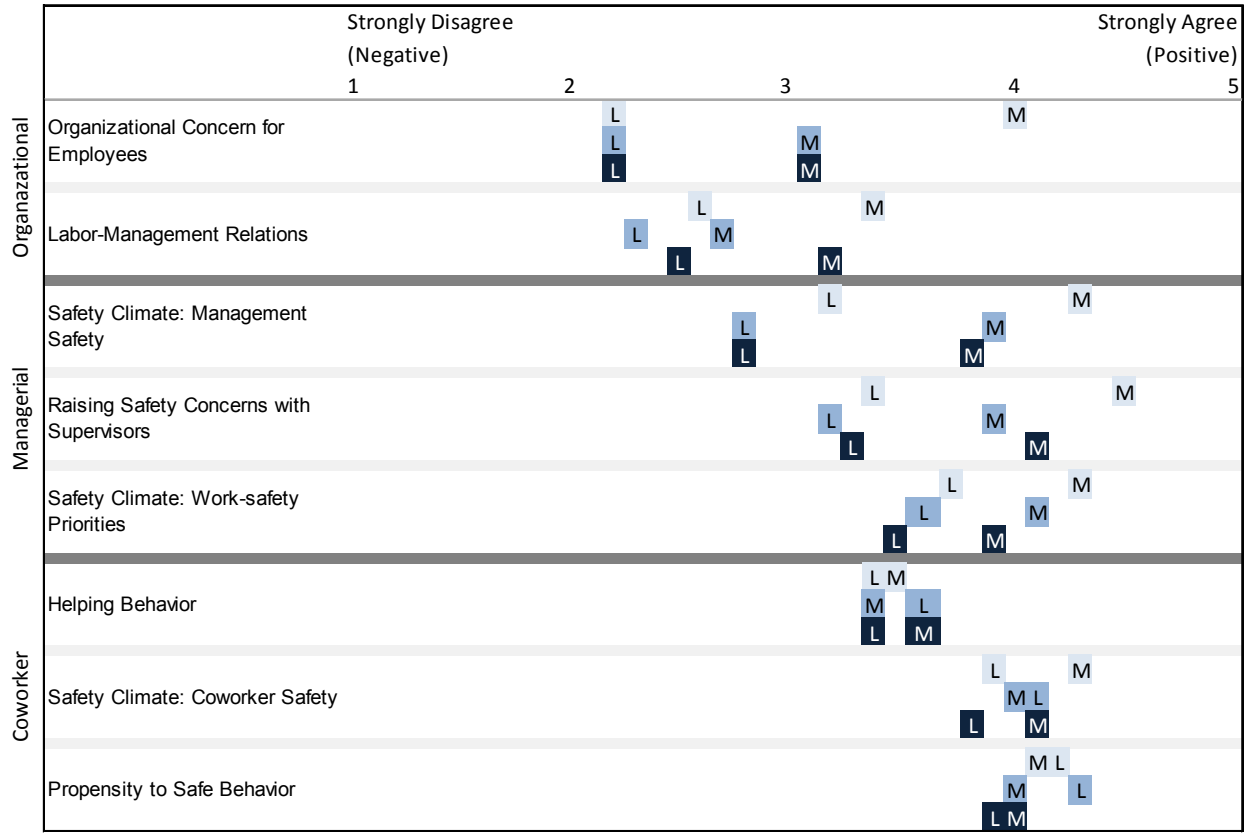
- Management tended to rate the organizational and managerial safety culture scales more positively than did labor.
- Labor was more positive about coworker safety culture scales than organizational or managerial scales.

[33]One railroad had a second round of baseline survey administration seven months after the first round for the purpose of increasing the response rate. The Evaluation Team found significant differences between the two survey administrations, so the second round cannot be considered truly baseline. Therefore, the responses from the second round are not included in the baseline results in this report. They will be discussed in the midterm report.

[34] We also calculated the average using the average of each of the three railroads and found the results to be similar.

- Management and labor were similar on the coworker scales, which all had scores close to 4 indicating, on average, some agreement that coworkers behaved safely.

More conclusions based on comparisons across sites are included in *Appendix F. Railroad Safety Culture Survey Results*.



Legend	
L	Site 1 Labor
M	Site 1 Manager
L	Site 2 Labor
M	Site 2 Manager
L	Site 3 Labor
M	Site 3 Manager

Figure 9 Safety Culture Survey Scale Results

Figure 10 Average Safety Culture Survey Scale Results Across all Surveyed Sites

4.1.2 Knowledge and Beliefs about C³RS

In addition to the safety culture data discussed above, the survey included questions about knowledge and beliefs concerning C³RS. Results included the following:

- *Awareness:* For all three railroads, labor reported a higher level of awareness of events or conditions that might lead to an accident than did managers.
- *Preventing accidents:* All groups agreed that changes in systems and individual behavior and knowledge of root causes can prevent accidents.
- *Ability of joint labor-management teams to function successfully to discover causes of unsafe conditions:* There were no significant differences at Sites 2 and 3 between labor and management. Site 1's railroad managers scored more positively than labor. Overall they were between neutral and slightly positive about it.
- *Ability to successfully implement changes:* For all three sites, managers were more positive than labor. Managers were positive about it, while labor was neutral.

4.1.3 Willingness to Report

Respondents were asked if they would report a close call. Only data from labor were considered because this group submitted the largest number of C^{3}RS reports. Approximately 60 percent of the respondents at Sites 1 and 3 said "yes," while 82 percent of the respondents at Site 2 said "yes."

For those who answered "no," a follow-up question asked why and provided five possible reasons (Table 6). The respondents at each railroad had same concerns that were ranked in the same order. The top concern was "being punished by management," followed closely by "not familiar enough with the reporting procedure." Approximately one-quarter to one-third of respondents "do not think [C^{3}RS] would result in any change" or "do not trust the BTS to maintain confidentiality." The least popular reason for not reporting was that "the reporting process is too much of a bother."

Table 6. Reasons to Not Report to C^{3}RS[35]

	Site 1 %	Site 2 %	Site 3 %
I worry about being punished by management	55	47	54
I am not familiar enough with the reporting procedure	53	47	40
I do not think it would result in any change	29	24	37
I do not trust the BTS to maintain confidentiality	20	21	39
The reporting process is too much of a bother	5	11	12

4.1.4 Survey Open-Ended Questions

The survey ended by giving respondents a space for writing comments about safety culture and the C^{3}RS program. Although these data are qualitative, they are included in the survey results section. The percentage of respondents who chose to provide comments is shown in Table 7. Labor's response rate ranged from 21 to 33 percent across sites. The high number of comments, especially at the end of a long survey, indicates that the employees were engaged.

Table 7. Survey Comments Response Rate

	Labor Comments		Manager Comments	
Railroad	Number	%	Number	%
UP	88	21%	20	28%
CP	73	30%	6	21%
NJT	260	33%	1	5%

At baseline, the labor comments were often focused on the perception that management valued money and schedules more than safety, e.g.

> "*Railroad X only does what is in its best financial interest. The company does not care about my personal wellbeing until after I am hurt and only then to protect its financial liability in court.*"

[35] Percent is based on the number of people who selected at least one option (not total number of survey respondents).

"You must leave on time - no matter what."

"Safety is lip service. Management will not spend $ on safety."

"To do the job safely, it takes time. I don't think the company cares if it gets done safely."

Each railroad also included comments about specific safety issues, with Site 3 going into more detail than the other two. There were also comments about the C^{3}RS program, some optimistic and others skeptical, e.g.

"The C^{3}RS is a good way to bring attention to matters that otherwise would not be reported and is an idea[l] way to create a safer work place."

"Railroading is a dangerous job, I'm glad a program like this is in place."

"I think the C^{3}RS system will be a welcome change and approach."

"I believe the program is well intentioned but I do not feel the culture of blame by management towards the employees can be overcome."

"I am apprehensive to embrace C^{3}RS due to my general distrust of the carrier."

4.2 Reporting Rates

All sites submitted C^{3}RS reports during their first year and each site's reporting rate for the first 12 months of the C^{3}RS program is shown in Figure 11 (see below). To disguise the sites, the graph starts all of the site's rates at the same time although the sites began the C3RS program at different times. Thus, Month 1 represents the first month of reporting for each railroad.[36] Each railroad consistently received reports every month and these numbers represent the number of reports received from employees, normalized by the number of employees covered by the program. (Those rates per covered employee were converted to the rate per 1000 covered employees, so the y-axis would not contain fractions.) Multiple reports on the same close call event may have been combined to create a single C^{3}RS case.

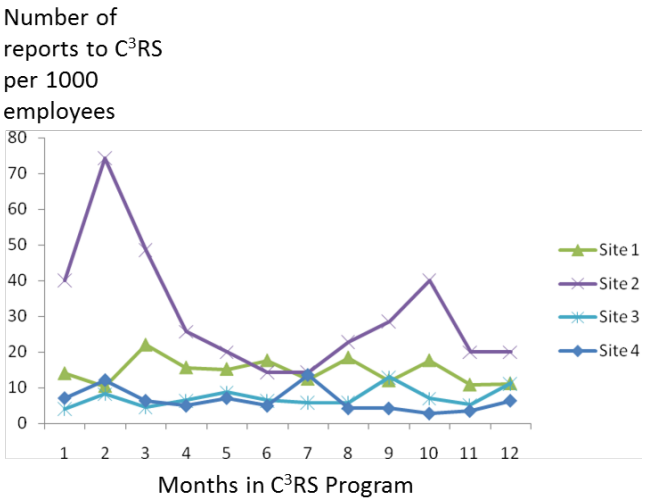


Figure 11. Number of C^{3}RS Reports per 1,000 Employees in Year 1 for Each Site

[36]Month 1 was not the same duration for each railroad, as some started reporting in the middle of the month.

4.3 Summary of Interviews and Field Notes

4.3.1 Differences in Demonstration Pilot Sites

The qualitative data obtained from interviews, review of project documents, newsletters, and observation of C^3RS operations, revealed key differences in the way that C^3RS was implemented at the different demonstration sites. These differences occurred because the Implementation Team learned from experience and made midcourse corrections as subsequent sites started up. The impact of these differences will be examined later in the evaluation.

PRT Support Team

The original designers of C^3RS assumed that recommendations from the PRT would be passed on to senior managers, who would in turn implement corrective actions. As experience with C^3RS accumulated, it became clear that a more formal interface was needed between the PRT and corporate management. That interface took the form of a PRT Support Team whose purpose was to review expensive or cross-functional corrective action recommendations and implement them. The Support Team comprised managers who could either act to implement a corrective action or work within the company to achieve implementation.

One railroad's Support Team configuration: At one site, the first iteration of the Support Team was an existing group of senior managers in a previously established safety board. The mission of this group was expanded to support the PRT. The group met quarterly. Infrequent meeting times and multiple purposes limited the ability of the team to assist with corrective actions. To correct this limitation, the team was reconstituted with different members and a more frequent meeting schedule. Over and above the difficulties presented by the team structure, communication between the PRT and the team proved problematic. The essence of the problem was that the PRT found it difficult and time-consuming to prepare information in a way that hid confidential data, per CIPSEA, while retaining the capability to explain why a corrective action was needed. This confusion affected the PRT's ability to share information with fellow employees. Furthermore, there was confusion about what could be revealed while adhering to the strictures of CIPSEA. This confusion was exaggerated because the IMOU had more restrictive requirements than CIPSEA.

Another railroad's Support Team configuration: At another site, the Support Team evolved into an effective group that included the transportation function as well as cross-functional involvement from engineering, mechanical, finance, and safety and training. The Support Team and the PRT met every two months and communicated well. For example, the team provided feedback to the PRT on the status and priority of corrective actions; called a PRT member to ask for more information when needed; performed cost-benefit analyses; and, due to its composition, provided the perspective of other functions outside of transportation.

Division of Labor to Accommodate Workload

To manage C[3]RS workloads properly, the Implementation Team developed mechanisms that deviated from the original model. At some sites, a PRT Leadership Team was formed, and it included people who handled three roles: process manager, facilitator, and scribe. Process managers handled administrative activities such as obtaining resources like meeting rooms and projectors; they were also responsible for coordinating staff transitions when existing members left and new members joined the PRT. Facilitators helped run PRT meetings. Scribes documented the results of the PRT's analyses and their corrective action recommendations with the automated MCIA tool. Members of this team received additional training from the Implementation Team on how to perform their duties.

At another site, a different mechanism (the PRT Sub-Committee) was installed to improve PRT efficiency by pre-processing cases before they were brought to the entire PRT for discussion and tracking corrective actions. This occurred later in the program, closer to midterm.

Changes in PRT Training

As the Implementation Team gained experience, the initial training for the PRTs evolved at each subsequent site. Three themes characterized this evolution:

- First, a major change was made when the Implementation Team provided more training to help PRT members minimize hindsight bias and confirmation bias as they assessed reports and attributed root causes. Hindsight bias is the tendency to think that a past event is more predictable than it was, given everything known today about it; in other words, the past event is viewed with today's insights. Confirmation bias is the tendency to favor information that strengthens one's own beliefs.[37]
- Second, training began to emphasize systemic fixes over local corrective actions. This change was needed because many of the early PRTs focused on effecting changes in individual behavior, providing formal training, and reminding people about things that they should know, such as rules content. Sites discovered significant system-level issues on only a few occasions; the Implementation Team sought to increase those discoveries.
- Finally, the Implementation Team moved from teaching PRTs about the MCIA tool to helping them understand the comprehensive corrective action process that starts with C[3]RS reporting and ends with implementation and monitoring of corrective actions.

[37] Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*. New Series, 185(4157), 1124-1131

4.3.2 Interview Results from Baseline Phase

As mentioned in Section 4, the evaluation included two types of interviews:

- Face-to-face "phased interviews," which were conducted with management and labor at sites where C^{3}RS was implemented to determine the program's impact and how well it was operating
- "Implementation interviews," which were conducted with a wide variety of stakeholders and close observers of C^{3}RS to identify key events affecting the ongoing functioning and sustainability of C^{3}RS

The most noteworthy themes that emerged from these interviews are summarized in two different ways. Table 8 abbreviates the themes and lists the number of railroads where each theme appeared, while Table 9 provides more detail on the themes.

These themes are only from the baseline phase, meaning that they are preliminary and may shift by the midterm and/or final evaluation phases. The problem "difficulty communicating with management and implementing corrective actions," may be experienced by more than just one pilot site at midterm or final.

Table 8. Number of Demonstration Pilot Sites Where Interview Themes Appear

	No. of Sites Where Theme Was Mentioned			
Baseline Interview Themes	**1**	**2**	**3**	**4**
Mixed levels of understanding about the C^3RS project, but still optimistic				x
There is room to improve the labor-management relationships				x
Discipline culture that included fear about discipline existed before the C^3RS project				x
Protection from discipline motivated employees to report				x
Success in keeping "hats off" — labor, management, and FRA members of PRT work well together				x
Reports that corrective actions are being implemented during baseline phase				x
Conflict over IMOU issues — which events would be covered under the discipline protection (IMOU Section 6.4)			x	
Formalizing management role on Support Team — PRT to send corrective actions to Support Team			x	
Innovation champions helped to successfully implement C^3RS		x		
Difficulty in implementing corrective actions		x		
Initial indications of improvement in labor-management relations at local level		x		
Optimism about C^3RS improving discipline culture		x		
Managers supporting 6.4 protection from discipline	x			
Worry that 6.4 may be a "get out of jail free card"	x			
Some skepticism about C^3RS safety improvements getting implemented	x			
PRT's initial recommendations seemed to be mostly local	x			
Return on investment is on managers' minds	x			
Complaints about the mechanical department	x			
Communication between PRT and management was insufficient	x			

Baseline Interview Themes	No. of Sites Where Theme Was Mentioned			
	1	2	3	4
Worried about budget to implement corrective actions	x			
Insufficient outreach to employees	x			
Interested in expanding C^{3}RS at baseline	x			

Table 9. Detailed Baseline Interview Themes

Detailed Comment Themes
Mixed levels of understanding about the C^{3}RS project, but still optimistic Some interviewees did not fully understand the C^{3}RS program, including how to report and what types of events qualified. However, they did generally seem interested in participating and reporting and had some optimism about the C^{3}RS program and its ability to improve safety.
Room to improve in labor-management relationships Similar to the survey responses, interviewees indicated that labor-management relationships were problematic. Sometimes these concerns were about discipline vs. a cooperative approach to solving problems. Considering the emphasis that C^{3}RS places on cooperative problem-solving, this concern might have an important bearing on the effectiveness of the program. These feelings seemed stronger at some railroads than others. Some interviewees felt that management and labor with lower tenure got along better, perhaps indicating that a negative organizational culture overwhelms new employees' initial optimism about the nature of their workplace. Labor sometimes had higher opinions of their supervisors than of corporate management.

Railroads have an existing discipline culture Employees felt a high level of fear about discipline before the C^{3}RS project. They felt there was a strong blame culture that led to individuals getting blamed for safety issues.
Protection from discipline motivated employees to report Employees were strong supporters of the part of the C^{3}RS program that protected them from discipline in situations where they would have otherwise been disciplined but for the fact that they submitted a report to C^{3}RS within 48 hours of the incident.[38] There is no doubt that, at baseline, the common belief was that 6.4 was a major motivator for employees to participate in C^{3}RS. 6.4 was not seen in such a positive light by managers, who had some concerns about labor abusing their protection. Some labor members at baseline at one site also were concerned about abuse. At one site, managers were positive that 6.4 would help to decrease the cost and time involved in investigations.
"Hats off" PRT members were very impressed at how well the labor, management, and FRA members of the PRT were able to work together. Time and again, they talked about how people were generally doing a good job at "taking their hats off" and working together as a team instead of engaging in the usual labor-management conflicts. Everyone acknowledged that this was a significant improvement in culture and a departure from the past. This culture change came earlier in the C^{3}RS life cycle than first expected. The idea that culture can change early, before an innovation is completed, matches the findings of an earlier study on safety rules revision.[39] Based on what the evaluation team has seen, culture change begins with the agreement to do the IMOU because some minimal level of trust is needed to begin such negotiations, and unless trust builds as a result of the discussions, successful agreements could not be reached. The existence of an initial trust level can be seen as part of an "unfreezing" process in which established patterns in an organization begin to loosen, thus providing an opportunity for change and a subsequent "refreezing" into a new order.[40]
Corrective actions implemented during baseline phase Railroads reported implementing corrective actions to address safety issues identified through C^{3}RS reports and recommended by the PRTs.

[38]More detail on C^{3}RS protection from discipline is available at http://www.closecallsrail.org/faq_protect.aspx

[39]Ranney, J., and Nelson, C. (2004). Impacts of participatory safety rules revision in U.S. railroad industry: An exploratory assessment. *Journal of the Transportation Research Board,* No. 1899, Transportation Research Board of the National Academies, Washington, DC, pp. 156–163.

[40]Lewin, K (1947). Group decision and social change. In T. N. Newcomb and E. L. Hartley (eds.) *Readings in Social Psychology*. Troy, MO: Hold, Rinehart, and Winston. Schein, E. H (1992) *Organizational Culture and Leadership* (2nd ed.) San Francisco: Jossey-Bass Publishers.

Dispute resolution issues At least three railroads encountered disputes concerning events that would be covered under the 6.4 protection. The involved parties spent considerable time negotiating the IMOU before reporting began. The PRTs then encountered unexpected disagreements about what events should be included as eligible C^{3}RS reports. The disagreements halted the PRT work for a while and caused some friction among labor and management members of the PRT and sometimes higher management. At baseline, the issues were eventually resolved to a degree that allowed the PRTs to resume work. Two related conclusions emerge from this finding. First, it may be impossible for an IMOU that will cover all contingencies to be prepared in advance; after all, the IMOUs were developed sequentially, with the second site benefitting from the experience of the first, yet very contentious, context-specific issues still arose. Second, because IMOUs seem unlikely to be able to cover all contingencies, it is important for participating railroads to have competent people and effective mechanisms that together can work through difficult issues. It is also possible that boundary-testing is a normal part of the project-acceptance process.
Formalizing management role on Support Team A Support Team was formalized at some sites to receive corrective actions from the PRT. At Site 3, there was frequent communication and feedback; the two groups seemed to be working together very well and getting actions implemented. At the other sites, there was more difficulty communicating and getting feedback from the Support Teams back to the PRT.
Innovation champion Site 1 had a very strong C^{3}RS supporter on the PRT. This person was a local manager who seemed to be key in getting the program implemented and the PRT working well with the capability to (1) take quick action to effect local change, (2) see systemic issues in the reports that the PRT analyzed and develop systemic corrective actions to address them, and (3) represent the needs of C^{3}RS to corporate management. Site 4 also had very strong management support in the baseline phase.
Difficult to implement corrective actions The corrective action process was not well defined, so it was difficult to prepare information for the next level of management and to track the status of corrective actions (especially those at the system level).
Initial indications of improvement in labor-management relations at local level As C^{3}RS got underway, there was less time spent in disciplinary hearings and drug testing. Labor and management seemed to be getting along better and were beginning to consider issues from each other's perspective. People seemed less worried about discussing safety issues with their managers.
Optimism about C^{3}RS improving discipline culture People believed that C^{3}RS might lessen management's discipline-oriented style.
Some skepticism about C^{3}RS safety improvements getting implemented This seemed to be based on past bad experiences with safety programs.
PRT's initial recommendations seemed to be mostly local There were few system-wide recommendations.

At baseline, return on investment was already on managers' minds From the beginning, managers were concerned about whether C^{3}RS was worth the investment.
Complaints about the mechanical department People at one site complained about mechanical department delays in implementing fixes to mechanical issues.
Communication between PRT and management was insufficient The PRT reported difficulty in communicating with upper management and getting feedback about the status of their recommended corrective actions.
Worried about budget to implement corrective actions There was some concern that future recommendations may be too expensive for the railroad to implement, given budget pressures.
Insufficient outreach to employees Rank and file had little understanding of the PRT's process; need to receive more communication, particularly on such issues as corrective actions and how their reports are used. Also, interviewees felt that the C^{3}RS program needed more local advocates.
Interested in expanding C^{3}RS at baseline One site was very quickly interested in expanding the scope of C^{3}RS after being involved for only a short time.

5. Discussion: Early Indications of C^{3}RS Performance

Even though this report covers only the baseline period, some important early indications of program performance emerged:

- Initial C^{3}RS start-up was accomplished at the four railroads participating in the demonstration. Considering the magnitude of the departure from long-standing practice (and tradition) represented by the C^{3}RS way of doing things, worker and management acceptance was high.
- Both labor and management were sufficiently committed to C^{3}RS that they were willing to commit considerable effort to provide the data needed for evaluation.
- The inevitable disagreements about issues outside the IMOU were resolved sufficiently to allow C^{3}RS to continue, though everyone was not completely satisfied.
- The survey indicated that there was a difference in cultural beliefs between management and labor at baseline (Figure 9 and Figure 10). For example, management was more positive about Organizational Concern for Employees, Labor-Management Relationships, and Raising Concerns with Supervisors. This gap is not surprising, but these data provide a depth of understanding that derives from three railroads, many different validated scales, and supporting interviews from a large sample of labor and management. (Improving safety culture and narrowing the gap between labor and management is one of the primary objectives of the C^{3}RS program.)
- Interview data revealed that, even at baseline, C^{3}RS had a positive impact on safety culture and on relationships between labor and management among members of the PRT. To progress in IMOU negotiations, trust had to grow among the parties. Cooperation between labor and management during PRT activities likely had a direct impact on the group's members.

6. Conclusion

Each demonstration site will be evaluated during the first four to five years of their participation in C^3RS. During that time, findings will shed light on the process by which C^3RS is being implemented and run, what impact it is having, and whether sustainability is being achieved. These findings will be presented in the following forms:

- Full reports containing methodological detail
- Research Results Reports on the FRA Web site
- Presentations at organizations such as the Transportation Review Board (TRB) and the American Evaluation Association (AEA)
- Articles for the professional journals and the popular press

Setting the stage for future reports, this report has done the following:

- Explained why a C^3RS approach was needed
- Provided details of the C^3RS program—its development, structure, and process
- Presented a schematic overview of the cross-site and longitudinal comparisons making up the methodology that will be applied to the C^3RS evaluation (Figure 7)
- Described the quantitative and qualitative measures that are being used
- Discussed the logic model that guides the evaluation process (Figure 8)
- Related the baseline observations that could be made public without divulging the identities of particular participating railroads

7. References

Aviation Safety Reporting System (ASRS), http://asrs.arc.nasa.gov/.

Canadian Pacific IMOU http://www.fta.dot.gov/documents/Canadian_Pacific_IMOU.pdf Improving railroad safety through understanding close calls. FRA-sponsored human factors workshop. Proceedings available at http://www.closecallsrail.org/publications_workshops.aspx.

Colquitt, J.A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology,* 86(3), 386–400.

Dastmalchian, A., Blyton, P., & Adamson, R. (1989). Industrial relations climate: Testing a construct. *Journal of Occupational Psychology,* 62, 21–32.

Dedobbeleer, N., & Beland, F. (1991). A safety climate measure for construction sites. *Journal of Safety Research,* 22, 97–103.

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology,* 71(3), 500–507.

Eisenberger, R., Stinglhamber, F., Vandenberghe, C., Sucharski, I.L., & Rhoades, L. (2002). Perceived supervisor support: Contributions to perceived organizational support and employee retention. *Journal of Applied Psychology,* 87(3), 565–573.

Hofmann, D.A. (2004). Safety climate: Management attitudes measure [personal communication].

Hofmann, D.A., & Stetzer, A. (1998). The role of safety climate and communication in accident interpretation: Implications for learning from negative events. *Academy of Management Journal,* 41, 644–657.

Hofmann, D.A., & Stetzer, A. (1996). A cross-level investigation of factors influencing unsafe behaviors and accidents. *Personnel Psychology,* 49, 307–339.

Hofmann, D.A., & Morgeson, F.P. (1999). Safety-related behavior as a social exchange: The role of perceived organizational support and leader-member exchange. *Journal of Applied Psychology,* 84, 286–296.

Johns, M.W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep: Journal of Sleep and Sleep Disorders Research,* 15(4), 376–381.

Lewin, K (1947). Group decision and social change. In T. N. Newcomb and E. L. Hartley (eds.) Readings in Social Psychology. Troy, MO: Hold, Rinehart, and Winston.

Magley, V.J., Davies-Schrils, K., & Walsh, B. (2007). Workplace climate for incivility: Development of a new measure. Unpublished manuscript, University of Connecticut.

Morell, J.A., Hanssen, C., Thompson, D., Wallace, R., & Wygant, B. (2006). Confidential close call reporting in the railroad industry: A literature review to inform evaluation. Available at http://www.closecallsrail.org/pubs/litEval.pdf.

Mueller, L., Da Silva, N., Townsend, J.C., & Tetrick, L. (1999). An empirical evaluation of competing safety climate measurement models. Paper presented at the Annual Meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.

National Research Council. (1994)Compensating Injured Railroad Workers Under the Federal Employer's Liability Act: Special Report 241. Washington, DC: The National Academies Press,.

Naumann, S.E., & Bennett, N. (2000). A case for procedural justice climate: Development and test of a multilevel model. *Academy of Management Journal,* 43(5), 881–889.

Niehoff, B.P., & Moorman, R.H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal,* 36, 527–556.

NJ Transit website, About Us, http://www.njtransit.com/tm/tm_servlet.srv?hdnPageAction=CorpInfoTo.

Organ, D., & Konovsky, M. (1989). Cognitive versus affective determinants of organizational citizenship behavior. *Journal of Applied Psychology,* 74, 157–164.

Patton, M.Q. (2002), Qualitative research and evaluation methods (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M.Q. (1987). How to use qualitative methods in evaluation. Newbury Park, CA: Sage.

Ranney, J., & Nelson, C. (2004). Impacts of participatory safety rules revision in U.S. railroad industry: An exploratory assessment. *Journal of the Transportation Research Board,* No. 1899, Transportation Research Board of the National Academies, Washington, DC, pp. 156–163.

Robinson, S.L. (1996). Trust and breach of the psychological contract. *Administrative Science Quarterly,* 41(4), 574–599.

Rosenthal, R., & Rosnow, R.L. (1991). Essentials of behavioral research: Methods and data analysis (2nd ed.). New York: McGraw-Hill.

Saks, J., Multer, J., Blythe, K. (2004). Proceedings of the Human Factors Workshop: Improving Railroad Safety Through Understanding Close Calls. Department of Transportation: Federal Railroad Administration: Office of Research and Development. (DOT-VNTSC-FRA-04-01).

Schein, E. H (1992) Organizational Culture and Leadership (2nd ed.) San Francisco: Jossey-Bass Publishers.

Simard, M., & Marchand, A. (1997). Workgroups' propensity to comply with safety rules: The influence of macro-micro organizational factors. *Ergonomics,* 40(2), 172–188.

Title V – Confidential Information Protection and Statistical Efficiency. 116 Stat. 2962. Public Law 107–347 –Dec. 17 2002. Available at http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/cipsea/cipsea_statute.pdf.

Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*. New Series, 185(4157), 1124-1131.

Union Pacific North Platte IMOU, August 25, 2006. Available at http://www.closecallsrail.org/pubs/North_Platte_Signed_IMOU.pdf.

Wayne, S.J., Shore, L.M., & Liden, R.C. (1997). Perceived organizational support and leader-member exchange: A social exchange perspective. *Academy of Management Journal,* 40, 82–111.

Zohar, D. (1980) Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology,* 65, 96–102.

45 U.S.C. § 51 et seq. (1908).

Appendix A. Logic Model

For reference, the model shown in Figure 8 is reproduced below. What follows are three enlargements for each part of the model.

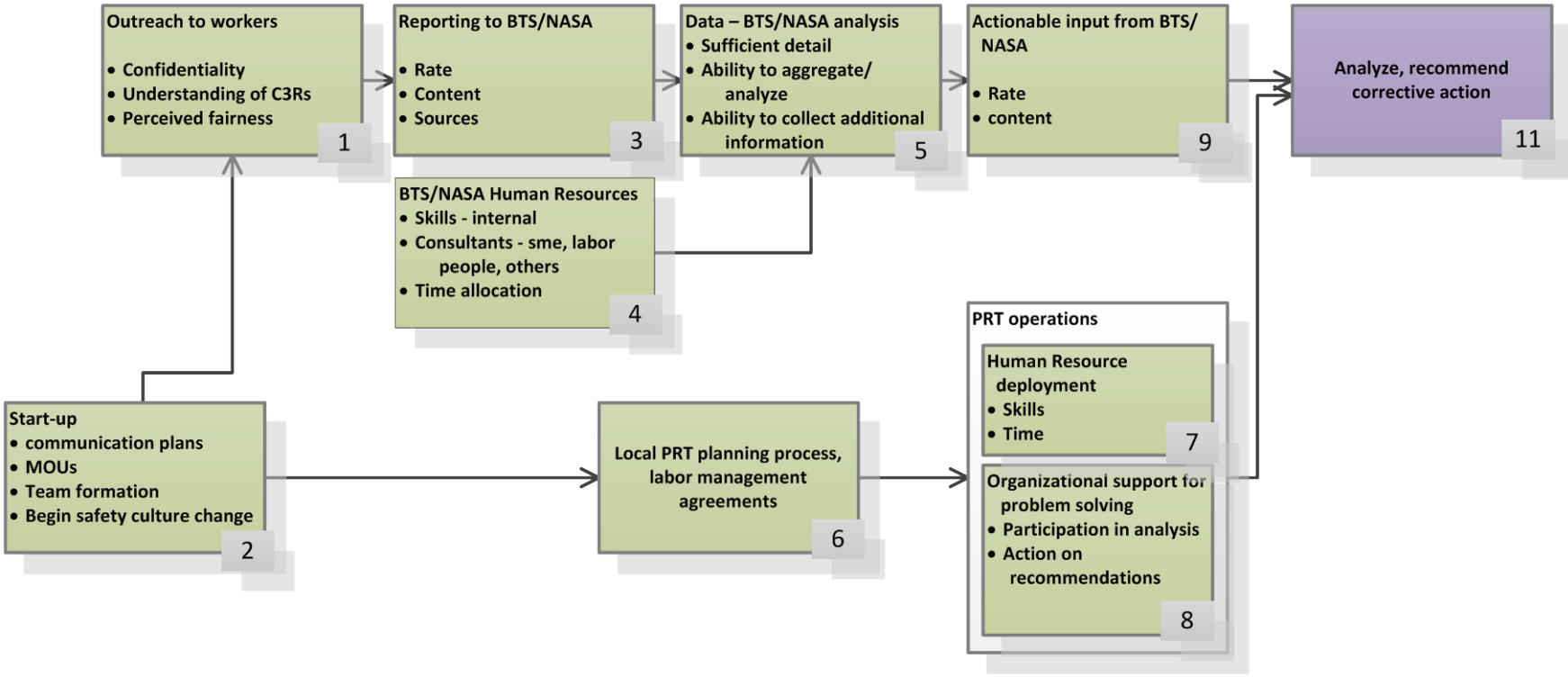


Figure 12. Logic Model: Detailed View of Implementation Portion

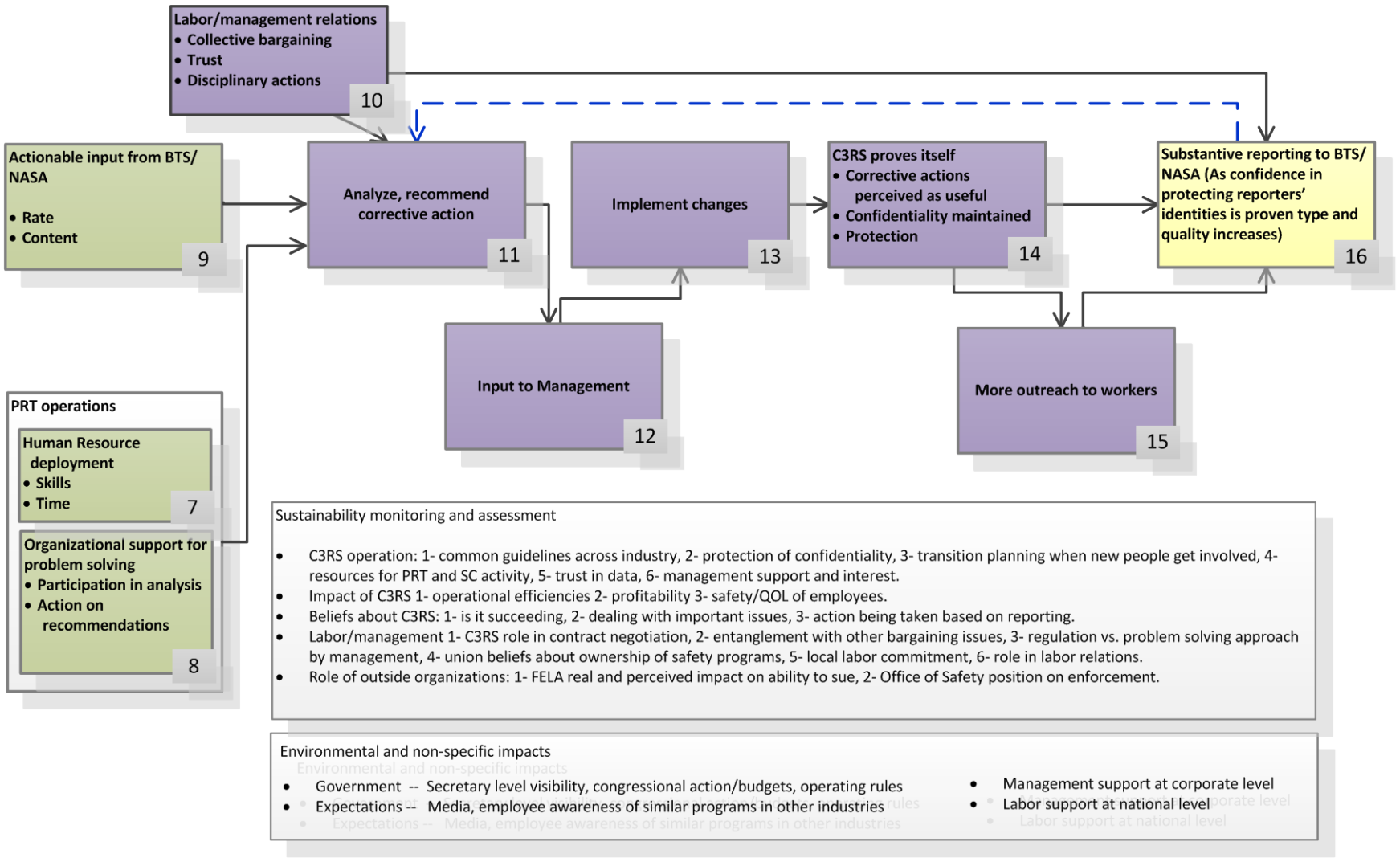


Figure 13. Logic Model: Detailed View of Operations Portion

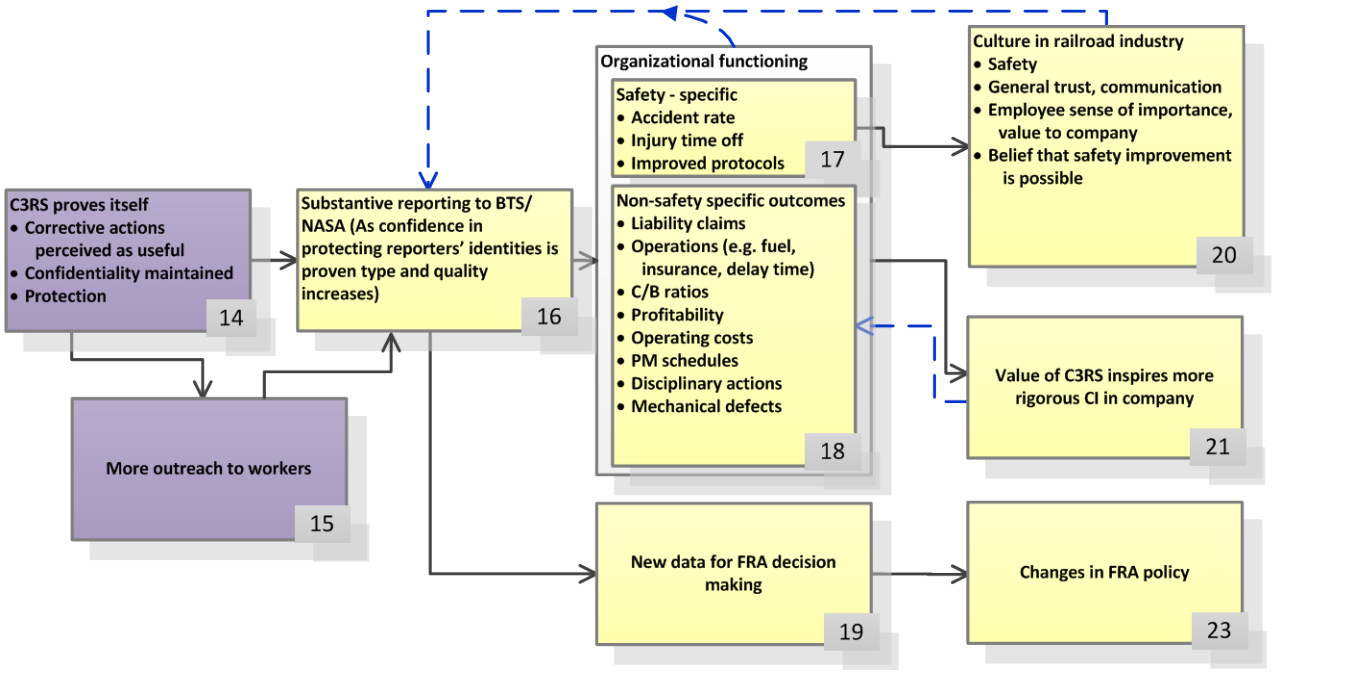


Figure 14. Logic Model: Detailed View of Impact Portion

Appendix B. List of Phased Site Interview Questions

OMB CONTROL NUMBER: 2130-0574

RAILROAD EMPLOYEES VIEWS OF C^{3}RS

Paperwork Reduction Act Burden Statement

A Federal agency may not conduct or sponsor, and a person is not required to respond to, nor shall a person be subject to a penalty for failure to comply with a collection of information subject to the requirements of the Paperwork Reduction Act unless that collection of information displays a currently valid OMB Control Number. The OMB Control Number for this information collection is **2130-0574**. Public reporting for this collection of information is estimated to be approximately 30 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, completing and reviewing the collection of information.

All responses to this collection of information are voluntary. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: Information Collection Clearance Officer, Federal Railroad Administration, 1200 New Jersey Avenue, SE, Washington, D.C. 20590.

Introduction

One of the objectives of the C^{3}RS lessons learned team is to determine what is required to improve the way C^{3}RS is implemented. We need this information to make recommendations for future implementations of the program. This interview is part of that effort. It will take about half an hour. I am only interested in the how C^{3}RS is going, not the substance of reports. To protect individual's privacy, we are not recording any names. All we need is a general description of respondents, e.g. "member of PRT; labor or management". Your participation in this interview is voluntary. If you want to skip any questions, please let us know. Thank you for meeting with us.

C^{3}RS (if labor)

L-1: Have you heard of C^{3}RS? (if No skip to S-1)

L-2: Do you think you understand the C^{3}RS well enough to know a reportable close call if you saw one?

Probe: What kinds of events have you been told can be reported?

L-3: Have you submitted a C^{3}RS report?

L-3a: (If they submitted a report) What did you think of your experience with the reporting system and BTS/NASA?

L-4: Do you know if C^{3}RS has resulted in any changes at your railroad?

L-5: Please tell us what changes you have seen.

L-6: How did you find about that these changes were made?

C^{3}RS (if manager)

M-1: Have you heard of C^{3}RS? (if No skip to S-1)

M-2: Do you think you understand the C^{3}RS well enough to give advice to your employees about what to report?

M-3: Do you know if C^{3}RS has resulted in any changes at your railroad?

Probes

- What are the changes?
- How did you find out about them? (formal vs. informal communication)
- Impact on
  - Safety culture: How management and labor interact.
  - Safety awareness
  - Safety (incidents, injuries, decertification)
  - Discipline
  - Cost

M-4: Have you personally been involved implementing any C^{3}RS corrective actions?

C^{3}RS (all)

A-1: From what you have seen of C^{3}RS, what changes would you suggest to make it work better or be more effective in improving safety?

A-2: To what extent do you think management is supportive of C^{3}RS?

A-3: To what extent do you think labor officials are supportive of C^{3}RS?

A-4: To what extent are your friends and colleague supportive of C^{3}RS?

A-5: If you had to bet $5.00, would you bet that C^{3}RS will be up and running at UP in five years? Why?

Safety in general (leave out if running out of time)

S-1: Over the past year or so have any safety initiatives taken place other than C^{3}RS?

S-1a:- Do they overlap or interact with C^{3}RS?

S-1b:- Do you think that C^{3}RS can improve safety in ways that other safety programs can't?

S-2: How would you describe the average worker's attitude about safety at your railroad?

S-3: How would you describe management's attitude about safety at your railroad?

S-4: How would you describe labor management relations regarding safety at your railroad?

S-5: How would you describe relations between labor and management regarding issues other than safety at your railroad?

S-6: Have relations between labor and management changed over the past year?

Appendix C. List of Implementation Interview Questions

OMB CONTROL NUMBER: 2130-0574

Paperwork Reduction Act Burden Statement

A Federal agency may not conduct or sponsor, and a person is not required to respond to, nor shall a person be subject to a penalty for failure to comply with a collection of information subject to the requirements of the Paperwork Reduction Act unless that collection of information displays a currently valid OMB Control Number. The OMB Control Number for this information collection is **2130-0574**. Public reporting for this collection of information is estimated to be approximately 30 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, completing and reviewing the collection of information.

All responses to this collection of information are voluntary. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: Information Collection Clearance Officer, Federal Railroad Administration, 1200 New Jersey Avenue, SE, Washington, D.C. 20590.

C^3RS Implementation Interview Protocol (OMB No. 2130-0574)

Introduction

One of the objectives of the C^3RS lessons learned team is to determine what is required to improve the way C^3RS is implemented. We need this information to make recommendations for future implementations of the program. This interview is part of that effort. It will take about half an hour. I am only interested in the how C^3RS is going, not the substance of reports. To protect individual's privacy, we are not recording any names. All we need is a general description of respondents, e.g. "railroad name, member of PRT; labor or management". Your participation in this interview is voluntary. If you want to skip any questions, please let us know. Thank you for meeting with us.

1- Thinking back over the past three months, what are the two or three most important positive or negative events that affected C^3RS?

Probes after description for each issue:

1- Why was this event so important?

2- Why do you think this event showed up when it did?

2- How satisfied are you with how C^3RS is currently working, and why do you feel that way?

2a. Probes:

- Peer Review Team
- Support Team Activities

2b. How could any of the groups involved in C^3RS change to improve C^3RS?

Listen, probe as necessary:

- Local management
- Local labor
- Railroad senior management
- BTS
- NASA
- FRA
- National labor

2c. Are there corrective actions that have been implemented that you think could have a big impact on safety? In addition to the ones you have mentioned, what are the kinds of corrective actions that are being implemented?

2d. Has C^3RS had any impact?

3- Over the past few months, are there any important events that took place outside normal C^3RS activities that affected the implementation or running of C^3RS?

4- Are there any issues effecting C^3RS' ability to maintain itself in the long run?

Appendix D. List of Qualitative Data Codes

This list contains codes used for analyzing qualitative data for the baseline report. They are organized by the five areas of the logic mode. Not all of these codes had significant frequency.

Code Family: Implementation

- 6.4 as implementation motivator
- Communication with all stakeholders
- Credibility of key members
- Groups opposing implementation
- Implementation start-up
- Innovation champion
- Key start-up meetings
- Local representation in early implementation
- Outreach to workers
- Past experience with change/collaboration
- PRT initial operations
- Signing IMOU

Code Family: Operations

- BTS activities
- C^3RS proves itself
- C^3RS reporting
- Confidentiality maintained
- Data quality
- FRA participation on PRT
- Implementation of corrective actions
- Irrelevant agendas
- Learning curve
- Poor participation
- PRT analysis
- PRT meetings
- PRT process experts

- PRT Support Team activities
- PRT tools
- PRTs sharing info
- Steering Committee/dispute resolution activities
- Tracking corrective actions

Code Family: Impact

- Culture change
- Impact
- Impact on FRA
- Impact on productivity
- Impact on safety
- Information-sharing among railroads
- Reporting impact

Code Family: Environment and Internal Climate/Culture

- Age differences
- Company attitude toward safety
- Competition
- Conflict among members
- Conflicted position of low-level managers
- Contradictory corporate policies
- Differences among railroads
- Discipline vs. cooperative approach to safety
- Groups vs. individuals
- Image of C^3RS among workforce
- Labor attitude toward safety
- New discipline policy
- Other accidents
- Other close call reporting programs outside railroads
- Passenger vs. freight
- Personal responsibility
- Safety problems
- Safety programs — other
- Safety vs. money
- Sharing track with other railroads
- Support from corporate
- Support from FRA
- Support from managers
- Support from NASA
- Support from Steering Committee
- Support from unions
- Weather interference

Code Family: Sustainability

- Confidentiality fears

- Cost and efficiency of running C^3RS
- Economy
- Funding
- Labor-management relationships
- Lack of understanding of C^3RS
- Maintaining interest
- National model
- Optimism about C^3RS
- Personnel turnover
- Post-pilot
- Public image of C^3RS
- Skepticism about C^3RS
- Stovepipes
- Sustainability

Appendix E. Example Survey Cover Letter

Confidential
Close Call
Reporting System

C^{3}RS Lessons Learned Survey

As you know, there is a joint effort by the FRA, UP labor (BLET, UTU), and UP management to test a safety improvement process known as the Confidential Close Call Reporting system (C^{3}RS) here at North Platte. If C^{3}RS works, the intention is to invest the resources needed to implement C^{3}RS across the railroad industry. But will it work? Will the investment be worth the effort? To find out, a Lesson Learned Team (LLT) was organized by the FRA to assess the impact of C^{3}RS on safety and safety culture. The assessment conducted by the LLT will provide both UP and the FRA with valuable information on C^{3}RS. The LLT is comprised of the Volpe Center, which is a US Department of Transportation (DOT) research center, Jacobs and Fulcrum Corporation, which are companies that support evaluation of safety initiatives, and the Bureau of Transportation Statistics (BTS), which is a statistical agency in DOT that supports data collection and data analysis.

You will see that the survey does NOT ask for your name. Your anonymity is important to the LLT. To further protect anonymity, the completed surveys will be sent directly to the BTS. Federal law 107-347 and the BTS Confidentiality Statute (49 U.S.C. 111(k)) gives the BTS the right and the obligation to protect data. By law, BTS will protect the identity of any survey respondent. BTS will not release any survey data collected from individual employees to FRA or any other public or private entity, including UP management. Any data and information collected through this survey will be use by the LLT for statistical purposes only and summary results will be published in a lessons learned report. The final lessons learned report will be available to all employees at this site. Further guidelines that will be used include:

- Summarized results will be given to the PRT, the C3RS steering committee and selected others.
- FRA will use the findings presented in the final report to deepen its understanding of lessons learned from the C^{3}RS project.
- The lessons learned will be shared with the railroad industry.

What we are asking you to do

- Complete the attached survey, seal it in the envelope provided, and give it to the person conducting your Safety Meeting.
- Use a pencil to mark the responses that best match your opinion.
- The survey looks long, but testing has shown that it takes only about twenty minutes to complete. Please give us those twenty minutes of your time.
- A 100% response rate is important to us. If you know someone who is absent, please encourage him or her to complete the survey. The person handing you this survey will have instructions as to how absent people can get a copy of the survey to complete.
- If you have already filled out this survey and you receive a second copy, please do not fill it out a second time.

Demetra Collia at the Bureau of Transportation Statistics is the survey coordinator for the Lesson Learned Team. If you have any questions about the survey, please call her at: XXX-XXX-XXXX, or send her email at: name @bts.gov

Thank you for your assistance.

Labor Representative Labor Representative Management Representative

Appendix F. Railroad Safety Culture Survey Results

F.1 Railroad Safety Culture Survey Results

The Evaluation Team examined baseline safety culture results for the three demonstration sites where BTS administered the Railroad Safety Culture Survey. A count of the responses to the baseline Railroad Safety Culture Survey is shown in Table 10 below.[41]

Table 10. Number of Baseline Survey Respondents

Railroad	No. of Labor Respondents	No. of Management Respondents
UP	427	72
CP	240	28
NJT	791	21
Total	1,458	121

F.2 Reliability

To determine the reliability of the Railroad Safety Culture Survey scales, the evaluators calculated Cronbach's Alpha for all scales, for labor and management separately and together at each site. As a result, 119 Cronbach's Alpha tests were conducted in total. For most scales, an acceptable score of 0.7 was achieved. Of the 119 scores, 18 were below the threshold of 0.7[42] and two were below 0.5 and were rated as "low." Both low alphas were found in the manager scale scores, which may have been due to the small number of respondents. The Propensity to Safe Behavior scale had the lowest alpha score. All scales were analyzed, even those with lower reliabilities, because this was only the baseline analysis.

F.3 Scale Scores

Multivariate Tests

Three multivariate analyses were performed on the eight scales that were common across the three railroads. These scales were Helping Behavior; Labor-Management Relations; Organizational Concern for Employees; Propensity to Safe Behavior; Raising Safety Concerns with Supervisors; Safety Climate: Coworker Safety; Safety Climate: Management Safety; and Safety Climate: Work-Safety Priorities. A multivariate test answers one question: Are there any

[41]One railroad had a second round of baseline survey administration seven months after the first round for the purpose of increasing the response rate. The Evaluation Team found significant differences between the two survey administrations, so the second round cannot be considered truly baseline. Therefore, the responses from the second round are not included in the baseline results in this report. They will be discussed in the midterm report.

[42]As a rule of thumb, alpha levels below 0.7 are commonly rated as too low to consider a list of items as a single scale of a single underlying construct (Rosenthal, R., & Rosnow, R.L. [1991]. Essentials of behavioral research: Methods and data analysis [2nd ed.]. New York: McGraw-Hill).

statistical differences between the means of the groups in the study?[43] The evaluators looked at three different groupings of respondents and used a multivariate test to discern differences.

Are there differences among the railroads? The first multivariate analysis assessed differences among railroads on the eight scales that were common to all sites, taking differences in labor (L) and management (M) proportions into account.[44] This analysis showed that the sites were significantly different from each other ($p = 0.000$) with respect to safety culture indicators. Post hoc Bonferonni tests were also performed to see where those differences occurred.

Are there differences between labor and management? The second multivariate analysis assessed differences between labor and management, taking into account the differences among railroads.[45] It showed that labor and management were significantly different from each other ($p = 0.000$). Again, post hoc Bonferonni tests were performed to see where those differences occurred.

Are there differences among the six groups of respondents? The final multivariate analysis compared six groups (Site 1 manager, Site 1 l labor, Site 2 manager, Site 2 labor, Site 3 manager, Site 3 labor) and was followed with a Bonferonni post hoc test to determine precisely where the differences lay.[46] Differences between the six groups existed ($p = 0.000$).

The statistical tests listed above revealed significant differences in the scale scores among the railroads and between labor and management. The next section will show the individual scale scores for each railroad's labor and management respondents.

Scale Scores

The results of the Railroad Safety Culture Survey are compared in Figure 15 for three categories of scales, Organizational, Managerial, and Coworker, with M standing for managers and L, for labor. All scale items are adjusted, with 1 indicating the most negative and 5 the most positive perception of safety culture. The conclusions are as follows:

- Management tended to rate the organizational and managerial safety culture scales more positively than did labor. Labor was the more positive about coworker safety culture.
- Labor scores for coworker safety culture were more positive than for the organizational or managerial scales.
- Management and labor were similar on the coworker scales, which all had scores close to 4 indicating, on average, some agreement that coworkers behaved safely.
- Labor perceived the organizational level of safety culture most negatively and saw the supervisory/managerial level slightly more positively.

---

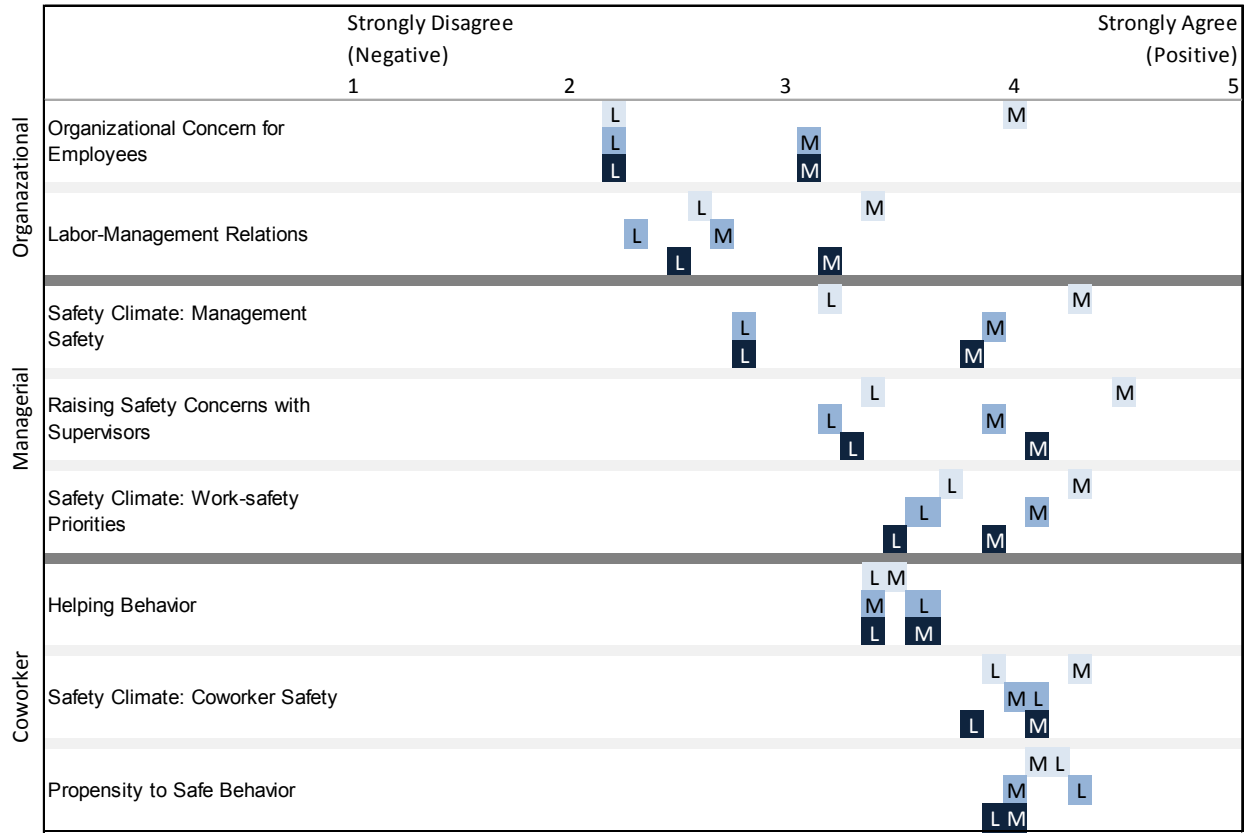
[43]Multivariate tests include multivariate analysis of variance (MANOVA) and multivariate analysis of covariance (MANACOVA).

[44]SPSS code: GLM help mlr oce psb rscs safem safes wst BY site WITH lvm /METHOD=SSTYPE(3) /INTERCEPT=INCLUDE /PRINT=DESCRIPTIVE /CRITERIA=ALPHA(.05) /DESIGN=lvm site.

[45]SPSS code: GLM help mlr oce psb rscs safem safes wst BY lvm WITH site /METHOD=SSTYPE(3) /INTERCEPT=INCLUDE /PRINT=DESCRIPTIVE /CRITERIA=ALPHA(.05) /DESIGN=site lvm.

[46]SPSS code: GLM help mlr oce psb rscs safem safes wst BY siteml /METHOD=SSTYPE(3) /INTERCEPT=INCLUDE /POSTHOC=siteml(BONFERRONI) /PRINT=DESCRIPTIVE /CRITERIA=ALPHA(.05) /DESIGN= siteml.

- Across sites, managers were most inconsistent in their responses for organizational scales and were mostly consistent in their responses for managerial and coworker scales, with Site 2 being the most positive.
  - Site 2 was higher than Site 1 and Site 3 on the Organizational Concern for Employees scale.
  - Site 2 was higher than Site 1 on the Labor-Management Relations scale.
  - Site 2 was higher than Site 1 on the Raising Safety Concerns with Supervisors scale.
  - Site 1 and Site 3 had no significant differences across any of the scales.
- Across sites, labor at Site 2 tended to be the most positive. Site 2 tended to be higher than at least one other railroad on most of the organizational and managerial scales. The exception was that Site 1 was high on the coworker scales.
  - All sites had similar scores on the Organizational Concern for Employees scale.
  - Site 2 and Site 3 were higher than Site 1 on the Labor-Management Relations scale.
  - Site 2 was higher than Site 1 and Site 3 on the Safety Climate: Management Safety scale.
  - Site 2 was higher than Site 1 on the Raising Safety Concerns with Supervisors scale.
  - Site 2 was higher than Site 3 on the Safety Climate: Work – Safety Priorities scale.
  - Site 1 was higher than both Site 2 and Site 3 on the Helping Behavior scale.
  - Site 1 and Site 2 were higher than Site 3 on the Safety Climate: Coworker Safety and Propensity to Safe Behavior scales.



Legend	
L	Site 1 Labor
M	Site 1 Manager
L	Site 2 Labor
M	Site 2 Manager
L	Site 3 Labor
M	Site 3 Manager

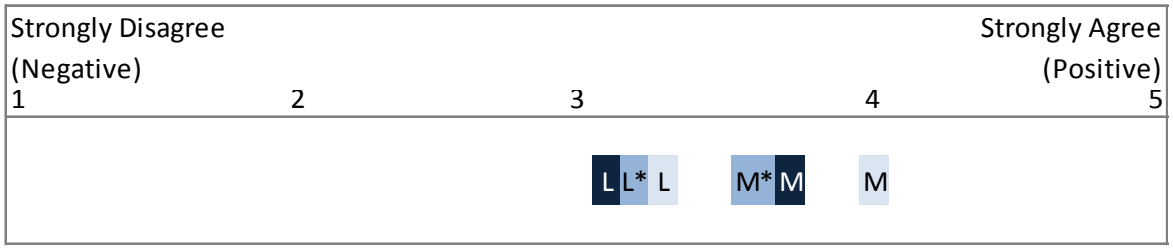
Figure 15. Safety Culture Survey Scale Results

Overall View of Safety Culture

The Evaluation Team also combined the scales to obtain a single summary view of the railroads' safety culture, using the steps that follow:

- *Step 1:* Scales that were common to all three railroads were identified, resulting in eight scales.
- *Step 2:* Six groups were defined: Site 1 labor, Site 1 manager, Site 2 labor, Site 2 manager, Site 3 labor, and Site 3 manager.
- *Step 3*: The average of all eight scales was calculated for each of the six groups, resulting in six overall safety culture scores.

- *Step 4*: ANOVA was performed to determine if differences existed across the six groups. Differences were found ($p = 0.000$).
- *Step 5:* A post hoc test was performed to determine where those differences occurred. The values are shown in Figure 16.



Legend	
L	Site 1 labor
M	Site 1 manager
L	Site 2 labor
M	Site 2 manager
L	Site 3 labor
M	Site 3 manager
*	Indicates that railroads on either side are not statistically different from each other.

Figure 16. Average Safety Culture Survey Scale Score Across the Three Railroads

At baseline, each site's average score for management was significantly higher than its average score for labor. Furthermore, every average management score was higher than any average labor score for any of the railroads.

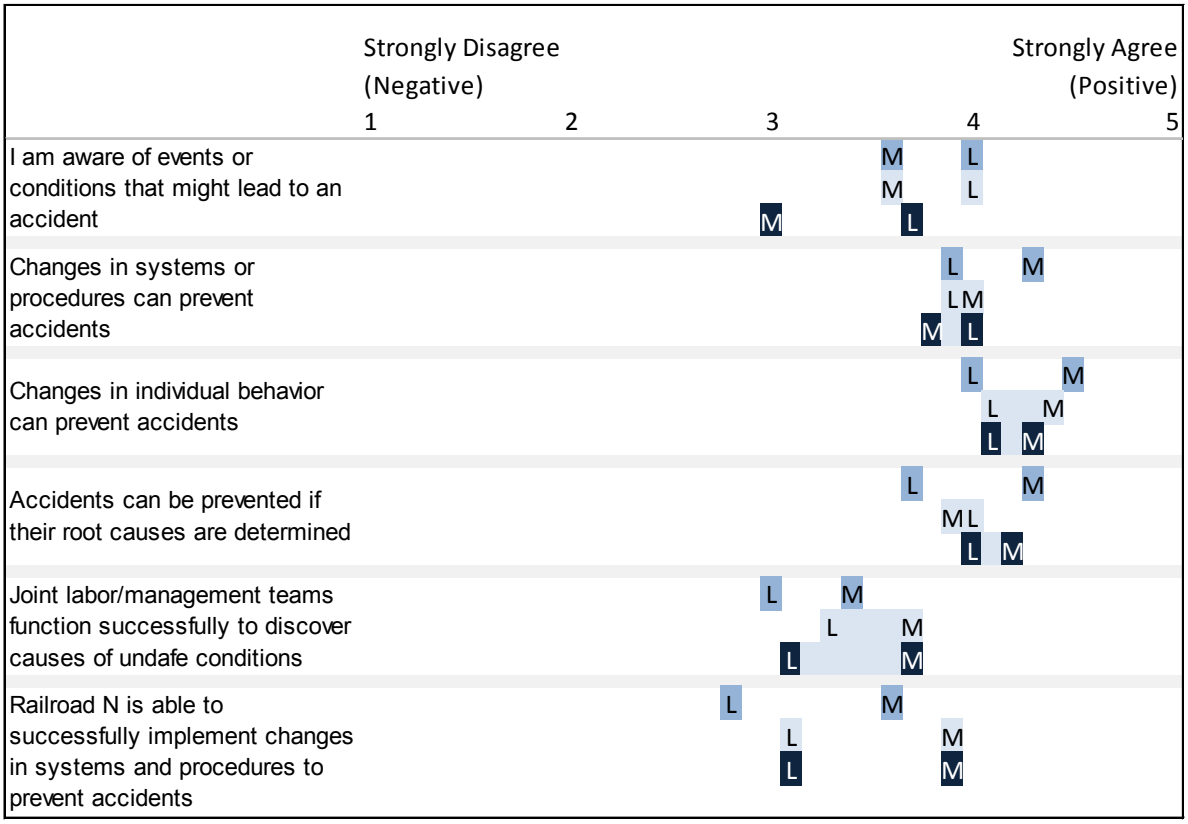
There were some differences between corresponding groups at each site:

- Managers at Site 2 (light blue) had significantly higher average scores than did managers at Sites 1 and 3. Manager scores for Sites 1 and 3, however, did not differ from each other.
- The three sites had very similar average labor scores despite so many differences across scales within each railroad.
    - Site 3 (dark blue) had slightly lower but statistically significant labor scores compared with Sites 1 and 2.
    - Labor scores for Sites 1 and 2 were not statistically different from each other.

F.4 Knowledge and Beliefs about C^{3}RS

In addition to the safety culture scales, the survey included questions about knowledge and beliefs concerning C^{3}RS. Two multivariate analyses were performed to look for differences between railroads. The first test examined whether there was a significant difference in responses according to the six categories used above when looking at all included C^{3}RS-specific questions.

This analysis showed that the groups were different ($p = 0.000$).[47] The second multivariate analysis examined the difference between labor and management when the data were aggregated across the three sites ($p = 0.000$). See Figure 17.



Legend	
L	Site 1 labor
M	Site 1 manager
L	Site 2 labor
M	Site 2 manager
L	Site 3 labor
M	Site 3 manager

Figure 17. C^{3}RS-Specific Survey Results

Results included the following:

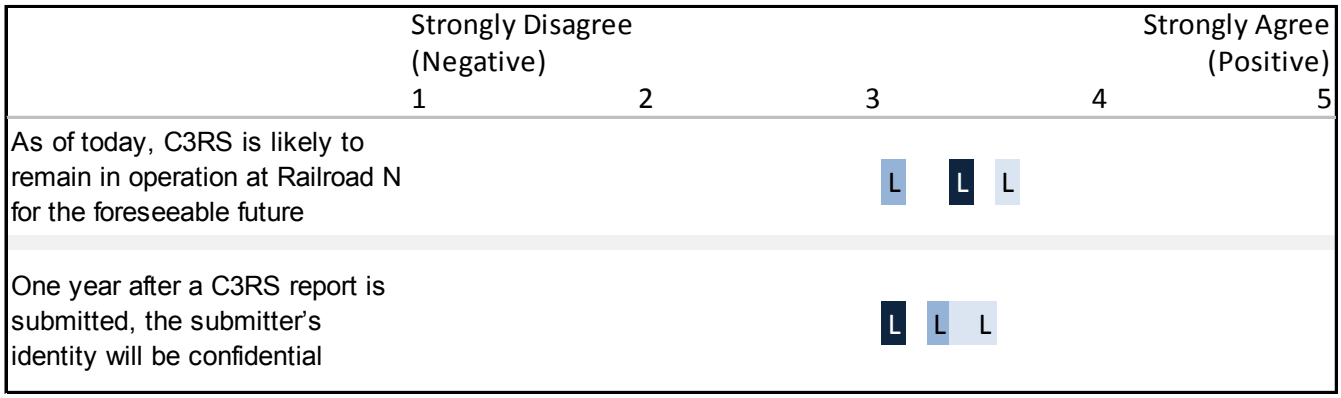
- *Awareness:* For all three railroads, labor reported a higher level of awareness of events or conditions that might lead to an accident than did managers. Differences among managers at the three sites were not statistically significant. Site 3 labor reported lower awareness than did Sites 1 and 2 labor.

[47] From a statistical standpoint, all p values are technically less than or equal to their value.

- *Preventing accidents:* All groups agreed understanding the root causes of problems as well as changes in systems and individual behavior can prevent accidents. There were no differences between Sites 2 and 3 labor and management for these three questions. At Site 1, management scored higher than labor on these questions.
- *Ability of joint labor-management teams to function successfully to discover causes of unsafe conditions:* There were no significant differences at Sites 2 and 3 between labor and management. Site 1's railroad managers scored more positively than labor.
- *Ability to successfully implement changes:* For all three sites, managers were always more positive than labor. Site 1 labor scored lower than Sites 2 and 3 labor.

F.5 C^{3}RS Implementation

Two questions on the survey addressed the quality of the implementation of C^{3}RS (Figure 18). Only data for labor were considered because that group submitted the most C^{3}RS reports. The Evaluation Team did a multivariate analysis comparing labor at the three pilot sites for these two questions listed in Figure 7 and found significant differences among them ($p = 0.000$)[48]. A post hoc test revealed that the three pilot sites had different views on the likelihood of C^{3}RS remaining in operation for the foreseeable future, ranging from 3.1 to 3.6; Site 2 was the most positive, followed by Site 3 and Site 1. Site 1 and Site 2 had similar views (no significant difference) about trusting confidentiality, with Site 3 being a little lower.



Legend	
L	labor
M	management
Light blue/grey	Site 2
Medium blue/grey	Site 1
Dark blue/black	Site 3

Figure 18. Confidence about C^{3}RS

[48]ONEWAY C9 C11 BY sitel /STATISTICS DESCRIPTIVES /MISSING ANALYSIS /POSTHOC=BONFERRONI ALPHA(0.05)

F.6 Willingness to Report

Respondents were asked if they would be willing to report a close call. Only data from labor were considered because this group submitted the most C^{3}RS reports. Approximately 60 percent of the respondents at Sites 1 and 3 said "yes," compared with 82 percent at Site 2. For those who answered "no," a follow-up question asked why and provided five possible reasons. (This question used a "check all that apply" format.) Table 11 lists the reasons offered. At all three sites, approximately half of the participants answered this question. The percentage that answered was calculated for each railroad (Table 11). As shown in the table, the "no" respondents at each railroad had the same concerns that were ranked in the same order. The top concern was "being punished by management," followed closely by "not familiar enough with the reporting procedure." Approximately one-quarter to one-third of the respondents "do not think [C^{3}RS] would result in any change" or "do not trust the BTS to maintain confidentiality." By far the least important reason for not reporting was that "the reporting process is too much of a bother." Site 3 respondents marked more options than did respondents at the other two sites.

Table 11. Reasons to Not Report to C^{3}RS

	Site 1 %	Site 2 %	Site 3 %
I worry about being punished by management	55	47	54
I am not familiar enough with the reporting procedure	53	47	40
I do not think it would result in any change	29	24	37
I do not trust the BTS to maintain confidentiality	20	21	39
The reporting process is too much of a bother	5	11	12

F.7 Survey Open-Ended Questions

The survey concluded with a blank space that allowed respondents to write their own comments about safety culture and the C^{3}RS program. Although these data are qualitative, they are included in the survey results section. The percentage of respondents who chose to provide comments is shown in Table 12. Labor's response rate ranged from 21 to 33 percent across sites. The high number of comments, especially at the end of a long survey, indicates that the employees were engaged. At baseline, the labor comments were often focused on the perception that management valued money more than safety. Each railroad included comments about specific safety issues, with Site 3 going into much more detail than the other two. There were also comments about the C^{3}RS program, some optimistic and others skeptical.

Table 12. Survey Comments Response Rate

	Labor Comments		Manager Comments	
Railroad	Number	%	Number	%
UP	88	21%	20	28%
CP	73	30%	6	21%
NJT	260	33%	1	5%

Abbreviations and Acronyms

AEA	American Evaluation Association
ASRS	Aviation Safety Reporting System
ATDA	American Train Dispatchers Association
BLET	Brotherhood of Locomotive Engineers and Trainmen
BTS	Bureau of Transportation Statistics
C^3RS	Confidential Close Call Reporting System
CFR	Code of Federal Regulations
CI	continuous improvement
CIPSEA	Confidential Information Protection and Statistical Efficiency Act
CP	Canadian Pacific
FELA	Federal Employers Liability Act
FOUO	for official use only
FRA	Federal Railroad Administration
IMOU	Implementation Memoranda of Understanding
MCIA	multiple cause incident analysis
MOU	Memorandum of Understanding
MANACOVA	multivariate analysis of covariance
MANOVA	multivariate analysis of variance
NDA	nondisclosure agreement
NASA	National Aeronautics and Space Administration
NJT	New Jersey Transit
NTSB	National Transportation Safety Board
ORD	Office of Research and Development
ORS	Office of Railroad Safety
PRT	Peer Review Team
TRB	Transportation Review Board
UP	Union Pacific
UTU	United Transportation Union
Volpe Center	Volpe National Transportation Systems Center