Impacts of Participatory Safety Rules Revision in U.S. Railroad Industry

An Exploratory Assessment

Joyce Ranney and Christopher Nelson

Increasingly, the proliferation of safety rules is viewed not just as a nuisance but also as a threat to safety itself. Indeed, voluminous and overlapping rules might contribute to poor compliance because of confusion and disagreement about which rules are to be followed. When combined with the existing regime of fault-based injury liability laws governing the industry, rules often become the focus of worker–management conflict rather than tools for communication about safety hazards and solutions. An exploratory assessment of participatory rules revision as an instrument of safety improvement in the U.S. railroad industry is presented. A key premise of the intervention is that rules proliferation is symptomatic of deeper problems with the organizational safety culture that can be addressed through extensive stakeholder participation in the revision process. Data for the study came from three railroads and one in-land barge line. While interview data provided evidence for an impact on the safety culture, initial statistical analysis of the incident data failed to find decreases in incident rates associated with the advent of the formal rules revision process. Further exploration, however, suggested that the intervention actually began earlier with various preparatory activities. Statistical analysis of this revised hypothesis found a significant impact of the intervention on incident rates at one rail carrier, while incident rate declines at two other carriers could not be attributed to the intervention with confidence.

Several pressures in the railroad industry have caused a proliferation of safety rules. First, mergers during the 1980s and 1990s brought together companies, each with its own sets of rules. Without attempts to harmonize them, these companies often simply added the books together, resulting in overlapping, and sometimes conflicting, safety rules. In addition, the Federal Employer's Liability Act (FELA), the fault-based injury compensation system that governs the railroad industry, provides an incentive for employers to maintain a vast array of rules as a defense against claims of employer fault in worker injuries. More generally, organizations often write rules to ensure that accidents precipitated by a wide variety of idiosyncratic circumstances never happen again. As Kaufman notes, "compassion spawns red tape" (*1*).

Increasingly, the proliferation of safety rules is viewed as not just a nuisance but also as a threat to safety itself. Indeed, voluminous and overlapping rules might contribute to poor compliance because of confusion and disagreement about which rules are to be followed. When combined with the existing regime of fault-based injury liability laws governing the industry, rules often become the focus of worker–management conflict rather than tools for communication about safety hazards and solutions.

This paper reports on an exploratory assessment of participatory rules revision, an intervention developed by industry safety executives in collaboration with a rules revision expert, as an instrument of safety improvement in the U.S. railroad industry. The intervention is predicated on the notion that rules proliferation is symptomatic of deeper problems with the organizational safety culture that can be addressed through extensive stakeholder participation in the revision process. Data for the study came from three railroads and one in-land barge line. Interviews with safety executives, workers, and union officials sought to assess the degree of implementation and the impact of the intervention on the safety culture, while analysis of incidence data from three of the four carriers sought to assess its impact on safety outcomes.

PARTICIPATORY RULES REVISION

Under the conventional approach to safety rules revision, the safety department commissions a technical writer to rewrite the rules with direction from management. As Perrow (*2*) notes, however, "complaints about excessive rules or bad rules generally are symptomatic of more deep-seated problems that cannot be solved by changing rules." Accordingly, the approach to rules revision examined here also seeks to change the values, attitudes, competencies, and patterns of behavior [in short, the safety culture (*3*)] by shifting the primary responsibility for rule making from management to workers. Rather than a top–down process in which safety experts hand down a revised rule book, participatory rules revision emphasizes broad participation, which is thought to build trust between labor and management and thus increase the likelihood of information sharing and cooperation, which are key components of a healthy safety culture. This, in turn, is thought to increase compliance because of increased workforce ownership of the new rules. Indeed, when rules are conflicting and voluminous, one can easily defend noncompliance. When there are only a few rules that one's peers have drafted personally, common sense and peer pressure might encourage compliance.

Participation in the rules revision process focuses on such issues as which rules are worthy of being rules, which rules cover all employees and which ones are craft specific, and what language will be used in the new rules. To minimize confusion, rules that are applicable for all employees are defined as "core rules," whereas those appropriate for members of a certain craft only are defined as "craft-

J. Ranney, Operator Performance and Safety Analysis, Volpe National Transportation Systems Center, 55 Broadway, DTS-79, Cambridge, MA 02142. C. Nelson, The Evaluation Center, Western Michigan University, Kalamazoo, MI 49008.

Transportation Research Record: Journal of the Transportation Research Board, No. 1899, TRB, National Research Council, Washington, D.C., 2004, pp. 156–163.

specific rules." Core rules must be such that compliance is possible 100% of the time and describe the only proper way to perform a work activity.

In addition to accidents, the reduction in the number of rules might also reduce carrier liability by clarifying which rules can be complied with all the time. The cost of injuries in the railroad industry is substantial. For example, annual expenses for one railroad's personal injury-related events were more than $200 million in 2000 (*4*). In short, if the theory of change holds true, the benefits provided by safety rules revision could lead to a significant reduction in injury-related costs, resulting in a significant financial boost to the industry.

PREVIOUS RESEARCH

Participatory rules revision is based on a two-step causal claim: that rules revision can improve the safety culture, which, in turn, improves safety outcomes. A growing body of research supports the link between safety culture and safety outcomes. Bailey and Petersen, for instance, found a link between safety cultures and safety performance in business units (*5*). Moving beyond the railroad industry, survey-based studies have found links between the safety culture and accident rates in chemical plants (*6*) and in a nuclear laboratory (*7*). Similarly, a qualitative case analysis study found evidence of a safety culture–accident relationship on an aircraft carrier (*8*).

Other studies illuminate possible mechanisms linking safety culture and safety outcomes. A survey-based study of an aircraft carrier by Hofmann and Stetzer (9), for instance, found that the safety culture–outcome relationship was mediated by unsafe behavior, as measured by random safety audits. That study also identified specific patterns of behavior and values that encouraged a proactive, vigilant safety culture.

There is less evidence about the link between rules revision and the safety culture. A study by Simard and Marchand, while not addressing rules revision per se, found that supervisors' propensity to use participatory management was positively related to rule compliance (*10*) in a sample of 97 Canadian manufacturing firms. This lends credence to the premise that increased rule ownership increases compliance. By and large, however, little is known about how participatory rules revision processes might influence the safety culture and outcomes.

DATA AND METHODS

The purpose of the research described here was to retrospectively assess the impact of participatory safety rules revision on the safety culture, incident rates, and liability costs. The retrospective analysis consisted of two parts. First, the evaluators interviewed safety executives and employees representing unions and analyzed the data. Second, the evaluators analyzed company incident rate data to assess whether the intervention had a quantitatively measurable impact. The remainder of this section discusses site selection, interviews, and the statistical analysis of the incident rate data.

Site Selection and Interviews

Given that participatory rules revision is not widespread, the study sought to include the entire population of railroad carriers that had implemented the intervention at the time. The Hile Group, the only consultant currently facilitating participatory safety rules revisions, identified five carriers. The team sent the safety executives at each railroad a letter explaining the research project and requesting an interview to generate the lessons learned. The safety executive at one of the carriers could not be contacted, resulting in exclusion of that carrier from the analysis.

Interviews, which took place during fall 2001, were conducted over the phone and lasted approximately 1 h each. Typically, there were two interviewers (a lead interviewer and a note taker); the conversations were not recorded. Written notes from the interviews were transcribed by the note taker and reviewed by three members of the evaluation team to ensure accuracy.

During the interviews, the safety executives were asked to provide names and contact information for one to five individuals who could provide a union perspective on rules revision. Twelve were identified and six were interviewed; most were union workers, while some were local union representatives. This second round of interviews, which used the procedures described above, took place during winter 2002.

Analysis of Incident Rate Data

The demonstration of program impact requires comparison of the actual performance with the counterfactual performance: what would have happened in the absence of the program? Random assignment to treatment and control groups remains the best way to approximate the counterfactual performance but was impossible in this situation. As an alternative, a difference-in-difference design was used to assess whether observed changes in rates in the rules revision firms exceeded contemporaneous changes in similar firms. For the railroad carriers, data for the comparison series came from the sources noted below, while comparison series for the barge line were downloaded from the Bureau of Labor Statistics website (www.bls.gov). To form the railroad industry comparison series, the respective counts for Kansas City Southern (KCS) and CSX Transportation (CSXT) (the railroads in the analysis) were removed. For construction of the barge line comparison series, incident counts and hours worked from the treatment group firm were subtracted from two-digit standard industrial classifications (SICs) to construct a clean no-treatment comparison. Unfortunately, missing data precluded use of three- or four-digit SIC series (444 and 4449, respectively). Thus, inclusion of dissimilar firms in the comparison series represents an important limitation to the internal validity of the quasi experiment.

Incident rates were calculated from FRA reportable injuries, illnesses, and fatalities to employees on duty. Since the frequency of incidents naturally varies with exposure, incident counts were normalized by using the following algorithm:

—— ( )

where

R_{it} = incident rate for a given railroad *i* during month *t*,
C_{it} = number of reportable incidents in railroad *i* and month *t*, and
H_{it} = total hours worked in railroad *i* and year *t*.

Thus, rates may be interpreted as the number of incidents per 100 full-time equivalents per year (100 workers per year ≈200,000 h). Data on incident frequencies and employee hours worked were taken from *Railroad Safety Statistics* (*11*) and *Accident/Incident Bulletin* (*12*). The counts for 2002 were based on the period from January to August; subsequent data were not available at the time of the analysis. The company's vice president for safety and health provided data for the barge company. The convention in this industry is to use the same algorithm to calculate rates; that is, 100 workers per year is ≈200,000 h.

Statistical analysis of the difference in differences proceeded in three steps. First, interrupted times series models were estimated against the treatment series only, which assessed whether the rules revision was a sufficient condition of reductions in incident rates. The paucity of observations for KCS and American Commercial Barge Lines (ACBL) restricted the model to a simple spline regression [see, for example, work presented elsewhere (*13, 14*)]:

where

RATE_t = incident rate at time t;
TIME_t = preintervention time counter, postintervention time;
TPOST_t = difference between the pre- and postintervention slopes; and
ϵ_t = independently and identically Gaussian-distributed disturbance term.

Since the time variables were centered at the month just before the intervention, β_0 is the estimated incident rate immediately preceding the intervention. β_1, in turn, is the rate of change before the intervention, while β_2 is the difference in the rate of change before and after the intervention. The availability of more datum points for CSXT allowed a nonlinear specification in both the pre- and the postintervention series, a log-normal stochastic process, and a structural break in the series:

t

Again, β_0 is the estimated incident rate immediately preceding the intervention. Coefficients on TIME_t and TIME^2 are the rate of decline and acceleration–reversion before the intervention, respectively. INTERV_t, by contrast, is coded 0 for preintervention observations and 1 for postintervention observations and estimates the immediate change in rate just after the intervention. Finally, the coefficients on TPOST and TPOST^2 are the immediate change in the rate of decline in incident rate (slope shift parameter) and the rate at which incident rates reverted toward the preintervention mean, respectively. A fully linear model was also estimated. However, a Wald test rejected the null hypothesis of no difference between an unrestricted model with the squared terms and a restricted model without them ($p = .002$).

Causal attribution also requires demonstration that rules revision was a necessary condition of differences pre- and postintervention. Thus, the evaluation team also estimated the same interrupted time series models against the comparison group series (described above). Finally, a pooled regime-switching regression model was calculated to estimate statistically the differences in pre- and post-differences in the treatment series and the pre- and post-differences in the comparison series (*15, 16*). Specifically, Equations 2 and 3 were modified to include a dummy variable that distinguished treatment group data from comparison group data and interactions between the dummy variable and each of the remaining parameters. The coefficients on the dummy interaction terms represent the difference between the two groups and test statistically the null hypothesis of no difference between the coefficient estimates for each group. F-tests for the homogeneity of variance provided support for the assumption of common variance in the two sets of groups (*14*). (The p-values from the homogeneity of variance tests were 0.22 for CSXT, 0.39 for KCS, and 0.99 for ACBL.) All models were estimated by using robust regression (iteratively reweighted least squares), which downweights outlying observations.

Follow-Up Analysis

The most important decision in modeling the treatment series was placement of the point of intervention. Interviews with the Hile Group, which codeveloped the intervention with industry safety professionals, suggested placement of the intervention point coincident with the start of the formal rules revision process. As shown below, statistical analysis failed to reject the null hypothesis under this specification of the program effect. These findings were presented to safety executives during a round of follow-up interviews (during summer 2002) designed for member checking. Respondents consistently suggested placement of the intervention point 1 year earlier, during various participatory activities designed to prepare for the rules revision. Several respondents from each carrier were asked about the date of the rules revision intervention and the management preparatory activities to confirm the dates. Thus, the respondents, not the researchers, provided the decision about the date for the revised models. Given that the latter hypothesis was derived, in part, from analysis of the incident data, it is important that these data not be used to formally test the same hypothesis. However, the exploratory nature of this research, combined with the paucity of studies on participatory rules revision, justify this post hoc analysis.

KEY FINDINGS

Before the findings from the statistical and qualitative evidence can be discussed, it is necessary to confirm that the intervention was actually delivered as planned. Indeed, a program's failure to generate the intended results might be a sign of implementation failure rather than (or in addition to) faulty program design (*17, 18*). Also, program effects often take a considerable amount of time to work through organizational systems. Accordingly, this section also examines a number of intermediate indicators, since impacts might show up there before they manifest themselves in incident rates.

Reductions in Number of Rules

Perhaps the most obvious indicator of implementation is a reduction in the number of safety rules. Here, it is apparent that participatory rules revision led to fairly dramatic decreases in the volume of rules. Along with an overall reduction in rules, the intervention also made distinctions between core and craft-specific rules. Specifically,

- ACBL went from 400 safety and operating rules, job aids, and training information to 24 core safety rules (applied to everyone) and 101 craft-specific rules and recommended work practices.
- CSXT went from 900 safety and operating rules, job aids, and training information to 17 core safety rules and craft-specific rules for transportation ($n = 19$), mechanical ($n = 88$), track engineering

(the specific number was not available, but it was estimated to be similar to that for mechanical), and clerical (number not available).

- KCS went from 742 safety and operating rules, job aids, and training information to 17 core rules (applied to everyone), and craft-specific rules for transportation ($n = 93$), mechanical ($n = 242$), track engineering ($n = 227$), and clerical ($n = 98$).

At the time of the study, Canadian National/Illinois Central (CN/IC) was just initiating a rules revision. Thus, no comparable data were available. The Hile Group provided all the information on before–after comparisons (*19*).

It is also important to understand qualitative changes in rules. Below is an example of the type of confusion embedded in the old rules and how the old rule was changed to a rule with a simpler message and a more peer-to-peer writing style:

- Old Rules for Welding-Fire Protection:
  – Employees must not carry cigarette lighters or matches where they may be exposed to sparks or excessive heat (*20*).
  – Use only a friction lighter, stationary pilot light, or some other source of suitable ignition to light torch. Do not light torch with matches, cigarette, or cigarette lighter (*21*).
- New Rule for Welding-Fire Protection: Use only approved flint strikers to light welding torch (*22*).

A number of factors make straightforward interpretation of these differences pre- and postintervention difficult. First, preintervention rule counts often include operating rules, training material, and job aids that were interspersed throughout the old rule books. Second, reductions in the number of rules need not necessarily entail reductions in rule complexity and increases in rule intelligibility. For instance, it is possible that the scope of requirements that count as a "rule" might have increased, leading to reductions in the number of "rules" but no decrease in rule complexity. Finally, it was beyond the scope of the study to examine whether increases in the volume and the complexity of operating rules, training material, and job aids might have offset reductions in the number of safety rules. Despite these caveats, it is clear that the revision efforts succeeded in dramatically reducing the number of safety rules. Moreover, the interviews suggested a high degree of acceptance of the new rule books.

Impact on Safety Culture and Rule Compliance

As noted above, participatory rules revision seeks not only to decrease the number of safety rules but also to improve the safety culture: the values, attitudes, competencies, and patterns of behavior related to safety. Again, the evidence suggested at least a reasonable degree of success. Most respondents described a shift from a solely management-driven rule-making process to one that emphasized the role of the workforce in deciding what rules were worthy of being rules and each craft's responsibility for its own rule book. As one safety professional observed, "A main barrier at [our] carrier was defining the definition of a safety rule. All rules had been created by management and dictated to the workforce. Usually they had been created at the time of a catastrophe. The workforce was asked to comply with rules that didn't make sense." Interview data also suggested that the intervention did, in fact, precipitate a change in the way in which the workforce viewed rules, from "mostly not helpful" to "mostly helpful." The following quotation illustrates a typical change in the understanding of the safety rule's place in broader safety efforts: the core rules apply "to all departments; they're safety related, critical to safety, enforceable, compliable. If they're less than that, they might be 'departmental procedures,' which are departmental specific and also safety related. There are also 'suggested work procedures' which are not enforceable." Finally, examination of the rule books themselves revealed a change in rule writing language from a management-oriented "thou shalt" style to a workforce-oriented "we will" style.

Most respondents reported improvements in union–management relations, which can be a particularly challenging barrier to a collaborative style of rule making. Worker respondents were asked to rank union–management relations on a scale from 1 to 5 (with 1 being "very hostile" and 5 being "very cooperative") both before and after the rules revision. Of the six respondents who provided ratings, four reported ratings that were higher after the rules revision (the question was added during the study, precluding responses from all respondents). A fifth respondent reported no change in union–management relations (a rating of 3 both before and after). A sixth respondent said that while relations between management and rank-and-file union members had improved from a 2 to a 4, relations between management and union officials had deteriorated from a 2 to a 1.

Another intermediate indicator of rules revision is rule compliance. Revisions to safety rules cannot have an impact if the rules are disregarded. Respondents were asked to comment on the extent to which the rules revision process led to changes in compliance with the rules, the degree to which there was an increased focus on safety, and other factors related to the safety culture. Generally, there was agreement among the respondents that the rules revision process had been associated with improvements in rules compliance. While the interview data on rule compliance were encouraging, the evaluation team was not able to independently confirm the improvement in compliance. Indeed, data on compliance were requested, but the request was declined.

Impact on Incident Rates

Estimation of the statistical models with placement of the point of intervention at the beginning of the formal revision process failed in all cases to reject the null hypothesis. As noted above, however, further examination pointed strongly toward placement of the intervention point earlier in the time series, coincident with preparatory activities undertaken by management to address issues of distrust and suspicion on the part of labor and management. At two of the carriers, senior management became active sponsors of the project, and at another carrier, senior management was committed but less visible. (CN/IC data were insufficient to support any conclusions.) Senior management from one of the carriers wrote a value statement that embraced a more participative management style in all areas of the company, not just safety. This company held employee meetings and distributed materials to communicate the message. Rather than the traditional punitive approach, responses to safety rules violations became more developmental; employees were trained rather than disciplined. Another carrier had the rules revision consultant work with all of the safety committees before rules revision to help "till the soil" by discussing concerns with labor.

As noted above, respondents consistently mentioned these preparatory activities during the interviews. Moreover, it was clear that the respondents viewed them as part and parcel of the rules revision effort. Thus, the evaluation team reestimated the statistical models with the intervention point at the beginning of the management preparatory activities. The remainder of this section presents findings related to this alternative program hypothesis.

CSXT

The findings for the analysis of CSXT are provided in Table 1. The values in Column I provide estimates for the CSXT series only, while the values in Column II provide estimates for the industry comparison series. Estimates of the difference in difference are provided in Column III. Since incident rates were modeled in log units, the regression coefficients may be interpreted as the percent change in incident rate associated with a one-unit change in the independent variables.

The parameters of greatest interest are those for INTERV (the immediate drop in rates), TPOST (the change in slope after the intervention), and $TPOST^2$ (the change in the rate of acceleration–reversion after the intervention). Columns I and II show a small but nonsignificant immediate drop in rates coincident with management preparatory activities at CSXT but a significant increase in the comparison series. While the difference ($\beta = -0.37$) in immediate pre- and postintervention rates between CSXT and the comparison group reaches conventional standards of statistical significance for exploratory research ($p = .10$), examination of the raw data reveals that the comparison series experienced a similar, although delayed, improvement 3 months later. Thus, the immediate safety improvement at CSXT cannot be regarded as unique. Nor is there any evidence of a positive impact on the rate of improvement, as the coefficient on TPOST suggests that differences in the rate of decline in the two groups pre- and postintervention were nearly identical. Again, the absence of a clear difference is illustrated by the *p*-value in Column III (0.60). Finally, examination of the coefficients on the squared terms suggests that while the preintervention rates reverted to higher levels in both series, the reversion actually became stronger at CSXT after the intervention, while it became weaker at other firms [the difference is statistically discernible ($p = .006$)]. In sum, the only possible evidence of a positive impact at CSXT lies in an immediate postintervention drop in incident rates. However, there is no clear evidence that the decline was unique to CSXT. It is also clear that CSXT did not sustain its decline.

TABLE 1 Coefficient Estimates for Treatment and Industry Groups

Variable	Column I Treatment (p-values)	Column II Comparison (p-values)	Column III Treatment - Comparison (p-value)
CSXT[a]			
Intercept (β_0)	1.35[b] (0.09)	0.92 (0.19)	0.43 (0.39)
TIME (β_1)	-0.16 (0.27)	-0.24[b] (0.10)	0.08 (0.24)
$TIME^2$ (β_2)	0.01 (0.50)	0.02 (0.14)	-0.01 (0.76)
INTERV (β_3)	-0.05 (0.79)	0.32[c] (0.02)	-0.37[b] (0.10)
TPOST (β_4)	-0.29[c] (0.03)	-0.27[d] (0.005)	-0.02 (0.60)
$TPOST^2$ (β_5)	0.01 (0.67)	-0.02 (0.19)	0.03[d] (0.006)
KCS[e]			
Intercept (β_0)	3.47[d] (<0.001)	3.39[d] (<0.001)	0.08 (0.65)
TIME (β_1)	0.15[d] (0.001)	-0.21 (0.12)	0.36[c] (0.05)
TPOST (β_2)	-0.51[d] (<0.001)	0.25 (0.19)	-0.76[b] (0.07)
ACBL[f]			
Intercept (β_0)	4.87[d] (<0.001)	11.21[d] (<0.001)	-6.34[d] (0.005)
TIME (β_1)	-0.65[d] (0.01)	-0.59[d] (0.009)	-0.06 (0.23)
TPOST (β_2)	0.40 (0.22)	0.03 (0.93)	0.37 (0.13)

[a] Pre-intervention series runs from 1983 to 1989; post-intervention series runs from 1990 to 2002. Rates are logged.
[b] $p < 0.10$
[c] $p < 0.05$
[d] $p < 0.01$
[e] Pre-intervention series runs from 1995 to 1998; post-intervention series runs from 1999 to 2002.
[f] Pre-intervention series runs from 1990 to 1994; post-intervention series runs from 1995 to 2002.

KCS

The coefficient of primary interest in this analysis is the TPOST variable. The results reported in Table 1 show that KCS experienced a statistically significant decrease in its incidence rate ($\beta = -0.51$; $p < .001$) while the comparison series showed no discernible change ($p = .19$). The estimate in Column III suggests that the difference in the differences pre- and postintervention between the KCS and industry series ($\beta = -0.76$) is statistically significant ($p = .07$). In short, KCS posted a significant and unique increase in the rate of safety improvement.

An important caveat, however, lies in the fact that there was a statistically discernible difference between KCS and the rest of the industry in the rate of change before the intervention. On the one hand, this result questions the use of the remainder of the industry as a comparison group. On the other hand, the fact that the slope for KCS was less favorable before the intervention only strengthens the case for a positive impact. The findings are illustrated in Figure 1.

ACBL

The analysis performed on the ACBL data, and its matched industry dataset, is identical to the analysis of KCS data. The results are presented in the bottom panel of Table 1. Unlike the other two cases, there is no strong evidence of a difference pre- and postintervention either in ACBL ($p = .22$) or in the remainder of the industry ($p = .93$). Examination of the regression coefficients suggests that, if anything, ACBL experienced a slight relative decrease in the rate of improvement ($p = .13$). Thus, there is no indication that rules revision had a positive impact on incident rates and slight evidence of a negative impact in this case, which should be regarded with caution, as outlined in the following section.

Finally, robust regression downweights outliers and might lead to underestimates of program impact. Indeed, ordinary least-squares estimates produced slightly larger impact assessments. Thus, the estimates provided here are, if anything, conservative. All regressions were estimated by using STATA (version 7) software (robust regressions used Huber weights and biweights).

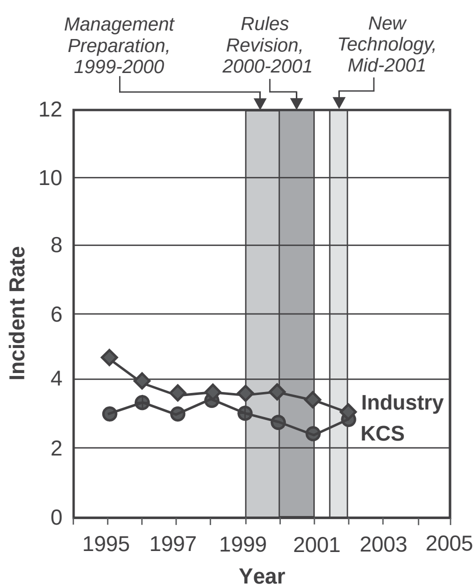

FIGURE 1 Comparison of KCS and industry incident rates, 1995–2002.

Impacts on Liability Costs

Respondents at two carriers noted that the rules revision decreased the number of claims and the claims payouts. One of the respondents reported being told by a financial professional in the company that "safety is contributing to the bottom line because of the drop." Another person mentioned that claims decreased both in the amount paid out and in number. However, the evaluation team was unable to independently verify these declines. Indeed, the same respondent noted that "claims data [are] very closely held, so there is no written documentation that can be given out." However, inclusion of the information seemed warranted, given the exploratory nature of the study.

In addition to the reduction in claims, another FELA-related impact concerns the reduction in the number of rules. One respondent pointed out a legal advantage to having fewer rules: "The reason it [the new rule book] is considered more defensible is because there are fewer rules. . . . With employees participating, there is more agreement on what is important to enforce in all situations on the side of labor and management."

DISCUSSION OF RESULTS

The analysis presented above suggests that participatory safety rules revision not only reduced the number of rules but also improved the usability of the rule books, compliance, and management–labor relationships. Indeed, interview data suggested that the intervention precipitated a change in the way in which the workforce viewed the rules from "mostly not helpful" to "mostly helpful." While the number of respondents on which these inferences are based is small, the interview data were remarkably consistent on these points. This is particularly important, because indigenous safety professionals with deep expertise about the industry and its challenges with safety rule books in a difficult labor environment developed the intervention.

Impacts on incident rates were less clear and less consistent. Statistical analysis failed to reject the original hypothesis that the advent of formal rules revision would be associated with a drop in incident rates. However, respondents in follow-up interviews consistently suggested that the intervention truly began with participatory activities undertaken by management as a prelude to formal rules revision. Taking this as a revised program hypothesis, the evaluation team found possible evidence of positive impacts at CSXT and KCS and no evidence of an impact at ACBL. The authors emphasize that the impacts must be described as "possible," given the paucity of observations and the fact that the control series included noncomparable railroads (because of missing values in three- and four-digit SIC data). In addition, comparisons between the treatment and the comparison series failed to yield unambiguous conclusions.

CSXT

At CSXT, for instance, it appeared initially that these preparatory program activities led to additional incidence rate declines, at least in the short run. Yet, this reduction in rates was mirrored—if belatedly—by the remainder of the industry, raising the possibility that it was caused by something other than rules revision. Moreover, the decrease in rates during the early 1990s was followed by a gradual leveling off, with rates on the rise again by the late 1990s. This raised the question of the sustainability of the impact of the rules revision. This might reflect the intervention's lack of resiliency to changes in the organizational and task environments. It could also indicate that management thought that safety was under control and became distracted by the leadership change at the company in 1995. As one respondent observed, "We lost a senior leader [in 1995], and it had tremendous impact. With new leadership, the rule book came under fire. We didn't have time to develop deep roots to the change effort and were in the middle of implementation when leadership changed. . . . There was a reversion back to the pre-[1989] status quo to a certain extent, but not completely."

Industry experts have stated that safety is a constant struggle. After the leadership change, for instance, CSXT acquired roughly half of Conrail in 1999. The addition of Conrail's rules to CSXT safety rule books and the resulting confusion in safety practices may have fueled an increase in incident rates. Other exogenous factors that might have contributed were technology, reduced employee expertise because of an influx of new hires, and a lack of training. Finally, the reverse in rates might reflect some sort of "floor effect." That is, the costs of efforts to reduce injuries might be much greater with lower incident rates and create barriers to further reductions. While the available data were not conclusive, it is clear that the resiliency of the rules revision intervention is an issue that bears further investigation.

KCS

KCS provides the strongest evidence that rules revision preceded by management preparatory activities had a positive impact. Similar to CSXT, positive findings at KCS were limited by the paucity of postintervention observations. Again, KCS's challenge was how to sustain safety improvements in the face of various exogenous factors; indeed, there was an increase in incident rates during 2002. One KCS safety executive explained that this increase might correspond with the introduction of a new car control system that led to delays and fatigue issues, as well as to other technological changes. Again, this raises questions about the resiliency of the intervention.

ACBL

Data from ACBL, by contrast, provide no evidence of an impact of the intervention. Here the null findings might be due, in part, to a number of confounding factors that might have increased incident rates. Indeed, the company experienced four mergers and acquisitions between 1996 and 2000 [the mergers and acquisitions involved ContiCarriers and Terminals, Inc., in 1996; National Marine, Inc., in 1998; Peavy Barge Line in 2000; and Ultrapetrol in 2000 (*23*)]. Therefore, the ACBL data are a combination of data for ACBL and four additional companies. While independent incidence data from before the merger and during acquisition were not available for the other carriers, the safety professional interviewed stated that the strong safety culture that had been achieved by ACBL, in part because of the rules revision, helped to smooth out the merger process and helped the other companies lower their incidence rates as they joined ACBL. For instance, "national Marine had an injury rate of 11 before they joined ACBL and dropped to 4 the year they merged [1998]." The respondent stated that if the mergers and acquisitions had not occurred, the slope change in ACBL might have had a more significant decline. Unfortunately, the available data did not allow the evaluation team to test these observations empirically.

Crosscutting Issues

For all three cases, it is worth reemphasizing that the alternative program hypothesis described above was derived from the data. Thus, these same data cannot provide a definitive test of the hypothesis. Having said that, the tentative finding about the link between participatory rules revision and the safety culture and outcomes is supported by other research on the relationship between employee involvement and organizational safety performance (*24–26*). This, however, leaves unanswered the important question of whether management participatory activities would have produced the observed effects without the subsequent rules revision. An adequate test of these rival program hypotheses would require comparisons among cases that had implemented (*a*) a rules revision only, (*b*) participatory management activities only, (*c*) both rules revision and participatory management activities, and (*d*) a no-treatment comparison group.

Moreover, several major industrywide improvements that occurred in recent years and that might confound estimates of the impacts of participatory rules revisions must be noted. Three such efforts are Switching Operations Fatality Analysis, FRA's Safety Assurance and Compliance Program, and FRA's Rail Advisory Committee. Unfortunately, it was beyond the scope of this study to examine further any differential effects of these industrywide improvement efforts.

Future research should seek data on better-matched comparison series than was possible here. Moreover, studies might also explore incident rate impacts in greater depth by analyzing departmental variations in exposure to the intervention and incident rate changes. One rail safety executive mentioned variations in implementation at the department level: "Mechanical had good sharing with one session a week. Not as good as [track] engineering. Clerical was okay. Transportation—it was hard to get them to spend time." Given the importance of intermediate outcomes, future studies should also seek to develop more rigorous measures of intermediate changes in safety culture and rule compliance through surveys, written records, and other large-sample instruments. For instance, it would be useful to examine whether rule violations decreased only for the rules that became core rules or whether there was a more generalized effect. Finally, future research must seek to determine the relative contributions of management preparatory activities compared with those of rules revision proper on the observed impacts.

CONCLUSION

More than a mechanism for reducing paperwork burdens and administrative complexity, participatory rules revision seeks to shift primary responsibility for safety rule making from management to the workforce and, in doing so, improve rules ownership, compliance and the safety culture, and incidence rates. This paper has provided an exploratory assessment of the effectiveness of participatory rules revisions as an instrument of safety improvement in the U.S. railroad industry. The findings suggest that participatory rules revision, when it includes preparatory activities, is a promising strategy for safety improvement. The findings reported here, however, are subject to considerable uncertainty. Given the potential for the intervention not only to decrease incident rates but also to improve management–worker relations, the safety culture, and liability costs, further research designed to reduce this uncertainty seems justified.

ACKNOWLEDGMENTS

The authors thank Lori Wingate of the Western Michigan University Evaluation Center, Mary Townsend of the Volpe Center, and Kate Chang of EG&G Technical Services, Inc., for invaluable insights and technical support. David Skinner of the Volpe Center provided specialized knowledge in obtaining the data sets and in the statistical data analysis. Michael Coplen of the FRA Office of Research and Development (RDV) provided valuable, ongoing guidance. Thomas Raslear, also of the FRA RDV, provided additional insights. The authors express their gratitude to the safety and unionized operating employees of CSXT, KCS, CN/IC, and ACBL who were interviewed in this study. The authors thank Jessie Sampson of the Teamsters from ACBL; Bob Keane of CN/IC; Sandy Hall and Paul Hillyard of CSXT; and Thomas Leopold and Fletcher Christian of KCS. Glenn Hotz, who led efforts as a safety executive at both ACBL and CSXT, provided perspectives from the leading rules revisions at two companies. The authors also acknowledge the insights gained from the Hile Group and Julie Hile regarding the technical principles underlying rules revision and obtaining access to the carriers involved in the study. Last, but not least, the authors acknowledge E. Donald Sussman and Jordan Multer of the Volpe Center for their valuable contributions.

REFERENCES

1. Kaufman, H. *Red Tape: Its Origins, Uses and Abuses.* Brookings Institution. Washington, D.C., 1977.
2. Perrow, C. *Complex Organizations.* McGraw-Hill, New York, 1986, p. 25.
3. Advisory Committee on Safety in Nuclear Installations. ACSNI Study Group on Human Factors. *Third Report, Organising for Safety.* Health and Safety Commission, London, 1993.
4. *Q 3 Filing—Third Quarter Filing: Securities and Exchange Commission.* Burlington Northern Santa Fe Corporation, Ft. Worth, Tex., 2001.
5. Bailey, C., and D. Petersen. Using Perception Surveys to Assess Safety System Effectiveness. *Professional Safety,* Vol. 34, No. 2, 1989, p. 22–26.
6. Carder, B., and P. W. Ragan. A Survey-Based System for Safety Measurement and Improvement. *Journal of Safety Research,* Vol. 34, No. 2, 2003, pp. 157–165.
7. Ostrom, L. C., C. Wilhelmsen, and B. L. Kaplan. Assessing Safety Culture. *Nuclear Safety,* Vol. 34, No. 2, 1993, pp. 163–172.
8. Rochlin, G., T. R. LaPorte, and K. H. Roberts. *The Self-Designing High-Reliability Organization: Aircraft Carrier Flight Operations at Sea.* Naval War College Review, Autumn, 1987, pp. 76–91.
9. Hofmann, D. A., and A. Stetzer. A Cross-Level Investigation of Factors Influencing Unsafe Behaviors and Accidents. *Personnel Psychology,* Vol. 49, No. 2, 1996, pp. 307–339.
10. Simard, M., and A. Marchand. Workgroups' Propensity to Comply with Safety Rules: The Influence of Macro–Micro Organizational Factors. *Ergonomics,* Vol. 40, No. 2, 1997, pp. 172–188.
11. *Railroad Safety Statistics: Annual Report.* FRA, U.S. Department of Transportation, 1997 to 2002.
12. *Accident/Incident Bulletin.* FRA, U.S. Department of Transportation, 1983 to 1996.
13. Greene, W. *Econometric Analysis.* Prentice Hall, Upper Saddle River, N.J., 2000.
14. Kmenta, J. *Elements of Econometrics,* 2nd ed. Macmillan, New York, 1986.
15. Card, D., and D. Sullivan. Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment. *Econometrica,* Vol. 56, No. 3, 1988, pp. 497–530.
16. Ashenfelter, O. Estimating the Effect of Training Programs on Earnings. *Review of Economics and Statistics,* Vol. 60, No. 1, 1978, pp. 47–57.
17. Stufflebeam, D. L. The CIPP Model for Evaluation. In *Evaluation Models* (D. L. Stufflebeam, G. F. Madaus, and T. Kellaghan, eds.), Kluwer Academic Publishers, Boston, 2000, pp. 279–317.
18. Pressman, J., and A. Wildavsky, *Implementation: How Great Expectations in Washington Are Dashed in Oakland; or, Why It's Amazing That Federal Programs Work at All.* University of California Press, Berkeley, 1973.
19. *Safety Rule Book Number of Rules Analysis.* Hile Group, Inc., Bloomington, Ind., 2002.
20. *Mechanical Department Safety Rulebook.* Grand Trunk Western Railroad, Chicago, Ill., Battle Creek, Mich., unknown date.
21. *Mechanical and Engineering Safety Rulebooks.* Illinois Central, 1994 (estimated).
22. *L.I.F.E.: Live Injury-Free Everyday.* Canadian National/Illinois Central Railroad, Montreal, Quebec, Canada, 2002.
23. Coleman, C. D. *W&H Main Yards: The Family Tree of North American Railroads.* 2001. webville@spikesys.com.
24. Erickson, J. A. The Relationship Between Corporate Culture and Safety Performance. *Professional Safety,* Vol. 42, 1997, pp. 29–33.
25. Millar, J. D. Valuing, Empowering Employees Vital to Quality Health and Safety Management. *Occupational Health & Safety,* Vol. 62, 1993, pp. 100–101.
26. Sarkus, D. J. Servant-Leadership in Safety: Advancing the Cause and Practice. *Professional Safety,* Vol. 41, 1996, pp. 26–32.

Publication of this paper sponsored by Vehicle User Characteristics Committee.

From Transportation Research Record: Journal of the Transportation Research Board, No. 1899, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 156-163.
Reproduced with permission of TRB.

The views and conclusions printed in this publication represent the opinions of the researchers who conducted the study and not necessarily the views of the department or funding agency, nor do they necessarily reflect official policy of the department or agency.

Safety performance indicators provide a means by which policymakers can ensure that their actions are as effective as possible and represent the best use of public resource. What are safety performance indicators? These are defined as any measurement that is causally related to crashes or injuries, used in addition to a count of crashes or injuries, in order to indicate safety performance or understand the process that leads to accidents. A large number of potential safety performance indicators exist.Â Among the road safety performance indicators most commonly used are those that relate to behavioural characteristics such as speed levels, the rate of drink driving and the use of seat belts. In addition, a number of infrastructure, vehicle or trauma-related indicators are relevant. Safety Risk Model data also underpins a number of other assessment tools used regularly in the GB rail industry. One such tool is the Signal Passed at Danger (SPAD) Risk Ranking Tool, which is used to assess the potential risk associated with each SPAD event that takes place on the GB mainline network in a simple, quick and consistent manner.Â So far as we know, it is the only industry-wide risk model of its type in the world of railways. Indeed there is considerable interest through the RSSB website from nearly 900 people in 43 countries around the world and it has already been licensed to one significant European infrastructure manager. References. Industrial safety procedures and rules date back to the 19th century when the United States enacted its first legislation to regulate the conditions of work in factories. They have since developed through federal and state legislation, voluntary organizational safety ...Â Industrial safety procedures and rules date back to the 19th century when the United States enacted its first legislation to regulate the conditions of work in factories. They have since developed through federal and state legislation, voluntary organizational safety programs, and investigations to determine the state of industrial safety. In the modern day, the Department of Labor regulates industrial safety.