EFFECTS OF BEHAVIOR-BASED DRIVER FEEDBACK SYSTEMS ON COMMERCIAL LONG HAUL OPERATOR SAFETY

Anuj K. Pradhan
University of Michigan, Ann Arbor, MI, USA
anujkp@gmail.com

Brian T. Lin
University of Michigan, Ann Arbor, MI, USA
btwlin@umich.edu

Claudia Wege
Volvo Group Trucks Technology, Gothenburg, Sweden
claudia.wege@volvo.com

Franziska Babel
Ulm University, Ulm, Germany
anujkp@gmail.com

Summary:

There are large economic and societal costs to commercial motor vehicle crashes. A majority of crashes are precipitated due to driver-related factors. Behavior-based systems that influence drivers with feedback from safety managers can help reduce driver-related risk factors. These systems harness the experience and knowledge of managers along with advanced driver telematics that monitor and record driver behaviors to positively influence driver safety. Safety solutions that focus on modifying driver behaviors thus hold promise for improving the safety record of commercial trucking. In this study, one such feedback system was examined by analyzing data from a commercial trucking fleet, treating the system deployment as a natural experiment. This made it possible, without experimental intervention, to compare drivers before and after system introduction, and to compare drivers that were subject to this system with those that drove with no supervisor feedback. Adverse event data were obtained for drivers in the fleet and weekly event rates were calculated taking into account driving exposure (in miles). Results show that drivers improved after receiving safety feedback and significantly more so than drivers that did not receive feedback.

BACKGROUND

Commercial motor vehicle safety has a broad societal impact and is an important public health issue. In 2014, there were 111,000 injuries and 3,903 fatalities from crashes involving large trucks and buses in the US (NHTSA, 2016). The estimated cost in 2014 of all large truck and bus crashes was $112 billion (FMCSA, 2014). Moreover, a crash causation study (FMCSA, 2007) indicated that over 87 percent of the critical reasons behind crashes were attributed to the driver.

Given the high impact of this particular factor, there is significant motivation to improve safety and facilitate cost-savings for fleets by targeting driver performance.

One approach to achieving this is through on-board monitoring systems. These are in-vehicle data recording devices that collect objective data by constantly or selectively recording video and kinematic data. This data can identify causes and outcomes of risky events, which in turn can be used to provide feedback on driving behaviors to help reduce crash risk (Horrey, Lesch, Dainoff, Robertson, & Noy, 2012). Such systems that collect objective data about vehicle and driver behaviors can be used by safety managers of commercial vehicle fleets to detect risky behaviors and to provide targeted and usable feedback to coach drivers and reduce crash risks.

Providing feedback to drivers may assist with driver self-regulation since it can provide objective information about one's performance and how that performance compares to normative expectations or other standards (Wrapson et al, 2016). Targeted and objective feedback can help provide information about performance, set goals for safety behaviors, help identify unsafe behaviors, and serve as a way to provide comparisons with peers. Such approaches have shown evidence of reduction in risky behaviors in other driver groups including teenage drivers (Toledo & Lotan, 2006; Simons-Morton et al., 2013).

OBJECTIVE

The study objective was to examine the effects of a behavior-based driver safety feedback program on driver risk. This was done by leveraging information from a deployment of such a system in a trucking fleet, and by analyzing aggregate fleet data from the deployment. This data offered a unique opportunity to consider it a natural experiment that allowed for comparison of drivers who had differential exposure to the system. Unlike a true experimental approach (Freedman et al, 1997), in a natural experiment a treatment variable is not manipulated by an experimenter and is thus more like an observational study. However, in specific cases such as this, the assignment of the observations can be treated "as if" random, allowing outcomes to be compared across 'control' and 'experimental' groups as if in an experimental design (Dunning, 2008). Accordingly, this design made it possible to compare the safety-feedback system between drivers who were exposed to the system and those who were not.

METHODS

The study was conducted as a secondary analyses of data available from a long-haul trucking fleet in the United States that had deployed the behavior-based driver safety system. The system was deployed by the fleet with approximately 500 power units in 2014. The system was video-based and included access to coachable driving events for fleet safety managers who then advised drivers with safety relevant events by phone, text, or in-person. The fleet conducted a phased deployment of the system with trucks being instrumented on a rolling basis. This included a period of time when a truck was instrumented and hence "live" but was not providing any feedback to the driver or managers. This was followed by a period of regular system activation when the drivers were alert to the system's intent and received feedback as prescribed by the program. The rollout also included a group of drivers that had an extended period when the system was 'live' and recording data, but the data were not viewed by the fleet safety

managers, and the drivers did not receive any feedback. (In both these cases the data were being received by the system vendor).

These circumstances allowed the treatment of the data as being from a natural experiment and analyzed accordingly. The rollout of the system allowed the definition of a "baseline" period during the initial rollout when the drivers were not being provided feedback. Subsequently, a set of drivers transitioned from this baseline period to the program period (or treatment period). The characteristics of the data, the grouping of the data sources, and the types of feedback that the drivers received (or did not receive), provided unique information that could be leveraged to conduct analyses on objective, exposure-controlled, event data for a natural experiment.

In-vehicle Data Recorder & Driver Safety Feedback Program

The commercially available program deployed by the fleet used a windshield mounted kinematics and video event recorder with one video lens toward the forward roadway and the other aimed inward at the driver. The system used GPS to track location and to calculate vehicle speed which was compared in real-time to posted speed limits. Safety-relevant and safety-critical events were captured when thresholds were exceeded (i.e., lateral/longitudinal forces in excess of .35G or vehicle traveling greater than 10 mph over posted speed for 20 seconds or more). Events were 12 seconds long with 8 seconds prior to trigger captured from a buffer plus 4 seconds post trigger. Each event contained video of the forward roadway, video and audio of the cab interior, speed, heading, location, and lateral and longitudinal G-forces. Two feedback configurations were used: 1) *No-feedback mode*: where all in-vehicle visual feedback were disabled and events were uploaded but only available to the vendor's review team; 2) *Feedback mode*: where an in-vehicle LED (visible to the driver) blinked when the system was triggered, and events were uploaded to the vendor's review team and then made available for safety managers. The system vendor also tracked safety manager performance and reported weekly to the lead safety manager. Reports called attention, for example, to backlogs of coachable events, drivers failing to reduce distractions, excessive speeding events, and near collisions. Drivers were coached via text message on an in-cab dispatch system, by phone, or in-person.

Event data

The event data was provided for analyses as reduced and pre-processed data from the fleet. No video data was obtained. All data was fully anonymized with no identifiable information about the fleet operators made available. The data only included events per driver ID, driver grouping, driver age, and daily driver mileage. The event data consisted of the date and time of event, the driver ID associated with the event, and other information such as event type (e.g. speeding, cell phone use, seat belt use, violation of traffic rules), event outcomes, and vehicle speed. Daily driver mileage information was also available for all driver IDs for the study period. Data were available for 37 weeks for 245 drivers after excluding drivers that were not active during the full duration of the analysis period, such as recent hires and terminations, or those out on illness or vacation. Of these, further 15 drivers were excluded due to data issues, including missing data and technical issues. Data from weeks 5-7 were also excluded due to issues with data integrity given the changeover of modes from 'no-feedback' to 'feedback' drivers. For the final analyses 7,123 events were generated by 230 drivers from all their trips during the 34-week period.

Design

The 230 drivers could be divided into two categories, the no-feedback group and the feedback group, and the time period under study could be divided into two sub-time periods: pre-feedback & post-feedback periods for the drivers in the feedback group, and a similar pre-feedback and post-feedback period for the no-feedback group drivers. (Despite the latter group not receiving feedback, it is useful to define similar time periods in order to make comparisons with the feedback group). For clarity, these time periods will be referred to as the *baseline* and *treatment* periods. The feedback group comprised 116 drivers. During the 4-week baseline period these drivers received no feedback from the system or from their managers. During the 30-week treatment period these drivers received in-vehicle feedback in the form of the blinking LED, and also received safety coaching and feedback from the fleet safety managers. On the other hand, 114 drivers that comprised the no-feedback group received no feedback at all, neither in-vehicle nor from safety managers, during both the baseline period and the treatment period (Table 1).

Table 1 – Feedback types for groups and time periods

Group	Number of drivers	Feedback type during Baseline period (Week 1 – 4)	Feedback type during Treatment period (Weeks 8– 37)
No-feedback	114	No feedback	No feedback
Feedback	116	No feedback	Flashing LED + coaching

Data analysis

The main dependent variable used for the analyses was event rate, which was calculated as the number of all safety critical events per week per 1,000 miles for each driver. These events included all event types including speeding violations, secondary task, etc. This was analyzed to compare the event rates between (1) baseline & treatment periods, and (2) between no-feedback & feedback groups for the treatment period. Wilcoxon paired test (for non-parametric distributed data) was used to compare the number of events per 1,000 miles between baseline and treatment periods of no-feedback and feedback groups. Wilcoxon Mann-Whitney U test was used to compare the event rates of non-feedback and feedback groups during treatment weeks (Figure 1).

RESULTS

Table 2 contains the summary statistics for the baseline and the treatment periods.

Table 2 – Mean (standard deviation) event rates per 1000 miles per week

|	Baseline period	Treatment period
No-feedback	0.743 (1.241)	0.514 (0.885)
Feedback	0.919 (2.233)	0.276 (0.499)

1. Differences in event rates between the baseline period and the treatment period:
    This analysis was conducted to explore differences in event rates between the 4-week baseline period and the 30-week treatment period for both the feedback and the no-feedback groups. For the feedback group, there was a significant difference in event rates between the baseline and treatment periods (V=4416, $p<0.001$). The event rates were also examined for the

no-feedback group and a significant difference was also found between baseline and treatment periods (V=3340, $p<0.05$) although the level of significance was much lower.

2. Differences between the feedback and the no-feedback groups during the treatment period:
This analysis examined event rate as a dependent variable while treating group (categorical) as an independent variable. A significant main effect of event rate was found for the group (no-feedback group: $M=0.5$, $SD=0.9$; feedback group: $M=0.3$, $SD=0.5$. U=5354, $p=0.01$). The event rate for the feedback group was significantly lower than that of the no-feedback group. The analyses also tested the effect of age for the two groups. It was found that age was a significant predictor for event rate in the no-feedback group ($F(1, 112)=5.56$, $p<0.05$). Essentially, a one-year increase in driver's age was associated with a decrease in the event rate by 0.02. However, in the feedback group, the effect of age was not significant ($F(1,114)=0.034$, $p=0.86$), potentially indicating that the effect of the feedback provided may have played a stronger role in event rate reduction than just age alone.
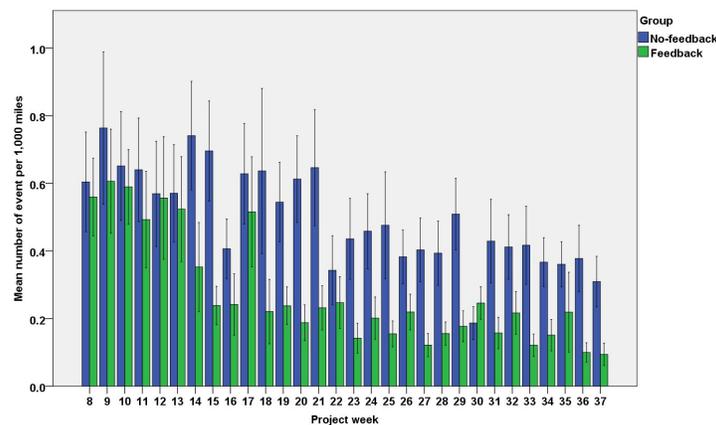
Figure 1 – Number of events per 1,000 miles for no-feedback and feedback group during treatment period (Error bars represent standard deviation)

CONCLUSION

In conclusion, this study shows a reduced rate of risky events for drivers in the monitoring and feedback program. These is also some evidence of age being significantly associated with lower event rates for older drivers when no feedback program is provided. An important caveat, however, is that no information was available about how long drivers had held their commercial licenses, nor was information available about individual drivers' mileage since commercial licensing. This makes it difficult to interpret the age finding. The obvious interpretation is that age is correlated with experience as well as maturity, and thus the reduced event rate is a function of these. However, this cannot be stated with certainly without supporting data.

Results like this can only be obtained using longitudinal studies. However, designing and undertaking a naturalistic longitudinal experimental study is a challenging enterprise with multiple real-life factors making the approach cost-prohibitive or impractical. However, based on the current results, the use of secondary data analyses for a natural experiment may be a good middle-ground, allowing the analysis of objective measures of driver behavior over time. This has methodological advantages compared to more traditional designs (e.g. simulator training

studies validating training effectiveness on driver safety (Robin et al. 2005)), but also has some shortcomings. A discussion on different methods (driving simulator, naturalistic driving studies or field operational test studies) and contrasting methodological approaches with regard to utility, data collection, and validity can be found in Dotzauer et al. (2014).

The advantages include a) large sample size, b) little manipulation of conditions, c) the high professionalism of the commercial drivers (all are part of a regular commercial driver training program), d) high scenario versatility (all ranges of weather conditions, roadway environments, etc.), e) real commercial operation situations, and f) the fact that safety managers and coaches are naturally exposed to their real managerial environments, day-to-day business, and real organizational rules while coaching their drivers. Thus the fact that this driver behavior modification feedback program was deployed in a commercial, real-world setting, ensures significant external validity (as recommended by Knipling, Hickman, & Bergoffen, 2003). The shortcomings of the natural experiment approach are that a) there was little control over driving scenarios and conditions, b) there was little control over the nature and deployment of coaching. Moreover, c) coaching was not controlled over all sessions, and, d) the number of coaching sessions within the treatment period and the interval between them was not controlled either, as it would be the case using more traditional study methods. Also, given the secondary data analysis nature of this study, some variables could not be obtained retrospectively such as the stableness of a driver's employment history, years of driving experience of each driver, safety culture at the fleet, attitudes of drivers and coaches, or the drivers' records of moving violations (e.g., speeding tickets). Additionally, it was not possible to interview drivers on acceptance of the system and/ or the overall feedback program

The current analysis of available data leaves room for further analysis such as the investigation of potential retention effects or transfer effects (e.g. What happens once a driver gets used to a system and then has to drive without it?). In future steps it should also be investigated which of the event types (e.g. speeding, cell phone use, seat belt use, violation of traffic rules) accounted for the majority of events and why. Additionally, in any future study using similar methodology, the effect of combined vs. single feedback could be investigated. E.g., what impact does one feedback type (LED lights) have vs. no LED light feedback but only retrospective feedback.

In conclusion, this study shows that on-board safety monitoring devices and behavior-based safety management systems have tremendous potential to impact commercial driver safety performance and particularly to identify potential unsafe driving performance. In line with previous literature (e.g., Wege & Victor, 2013), it is shown that on-board safety management devices can serve as a basis for short- and long-term safety performance feedback. On-the-job-evaluations can thus be a powerful countermeasure against safety violations of commercial truck drivers. Safe driving performance can be promoted once meaningful safety-focused communication is in place and the feedback is objective, individual, constructive, and targeted.

ACKNOWLEDGMENTS

REFERENCES

Dotzauer, M., Berthon-Donk, V., Beggiato, M., Haupt, J., & Piccinini, G. (2014). Methods to assess behavioral adaptation over time as a result of ADAs use. In A. Stevens, C. Brusque, & J. Krems (Eds). *Driver adaptation to information and assistance systems*.

Dunning, T. (2008). Improving causal inference: Strengths and limitations of natural experiments. *Political Research Quarterly*, *61*(2), 282-293.

Federal Motor Carrier Safety Administration. (2007). The Large Truck Crash Causation Study (Publication no. FMCSA-RRA-07-017). *FMCSA's Office of Analysis, Research, and Technology. Retrieved from www.fmcsa.dot.gov/facts-research/research-technology/analysis/fmcsa-rra-07-017.htm*

Federal Motor Carrier Safety Administration. (2014). Pocket guide to large truck and bus statistics - October 2014 update 2. *FMCSA's Office of Analysis, Research, and Technology*. *Retrieved from http://www.fmcsa.dot.gov/safety/research-and-analysis/2014-pocket-guide-large-truck-and-bus-statistics*

Freedman, D., Pisani, R., & Purves, R. (1997). Statistics. 3rd ed. New York: Norton.

Horrey, W. J., Lesch, M. F., Dainoff, M. J., Robertson, M. M., & Noy, Y. I. (2012). On-board safety monitoring systems for driving: review, knowledge gaps, and framework. *Journal of safety research*, *43*(1), 49-58.

Knipling, R.R., Hickman, J.S., & Bergoffen, G. (2003). Effective commercial truck and bus safety management techniques. Commercial Truck and Bus Safety Synthesis Program, Synthesis 1. Washington, DC: Transportation Research Board.

National Highway Traffic Safety Administration. (2016). Traffic safety facts, large trucks 2014 data. *NHTSA's National Center for Statistics and Analysis. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812279*

Robin, J.L., Knipling, R.R., Derrickson, M.L., Antonik, C., Tidwell, S.A., & McFann, J. (2005). Truck simulator validation ("SimVal") training effectiveness study. Proceedings of the 2005 truck & bus safety & security symposium (pp. 475-483). Alexandria, VA: national Safety Council.

Simons-Morton, B. G., Bingham, C. R., Ouimet, M. C., Pradhan, A. K., Chen, R., Barretto, A., & Shope, J. T. (2013). The effect on teenage risky driving of feedback from a safety monitoring system: A randomized controlled trial. *J. of Adolescent Health*, *53*(1), 21-26.

Toledo, T., & Lotan, T. (2006). In-vehicle data recorder for evaluation of driving behavior and safety. *Transportation Research Record: Journal of the Transportation Research Board*, (1953), 112-119.

Wege, C., & Victor, T. (2013). Distraction and Inattention Prevention by Combining Behaviour-Based Safety with Advanced Driver Assistance Systems. In A. Stevens, C. Brusque, & J. Krems (Eds). *Driver adaptation to information and assistance systems*. IET published book. ISBN: 978-1-84919-639-0; E-ISBN: 978-1-84919-640-6.

Wrapson, W., Harré, N., & Murrell, P. (2006). Reductions in driver speed using posted feedback of speeding information: Social comparison or implied surveillance? *Accident Analysis & Prevention*, *38*(6), 1119-1126.