# CRC Report No. CM-138-18-1

# CRC DRIVEABILITY WORKSHOP

**April 2020**

**COORDINATING RESEARCH COUNCIL, INC.**
**5755 NORTH POINT PARKWAY • SUITE 265 • ALPHARETTA, GA 30022**

# CRC DRIVEABILITY WORKSHOP

**FINAL REPORT**

**CRC Project No. CM-138-18-1**

**SwRI® Project No. 03.24785**

**Prepared for:**

**Dr. Christopher J. Tennant**

**Coordinating Research Council**
**5755 North Point Parkway, Suite 265**
**Alpharetta, GA 30022**
**ctennant@crcao.org**

**Prepared by:**

**Stanislav Gankov, Research Engineer**
**Sankar Rengarajan, Group Leader**
**James Fritz, Sr. Technician**

**Southwest Research Institute**
**6220 Culebra Road**
**San Antonio, TX 78238**

**April 17, 2020**

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## EXECUTIVE SUMMARY


CRC has used trained raters to assess vehicle driveability performance and has conducted rater workshop programs in the past to train and calibrate raters. Such a workshop has not taken place since 2002 (see CRC Report No. 631 2002 CRC Driveability Workshop), and there is a limited number of available trained raters. Southwest Research Institute (SwRI) has developed a "trick car" vehicle under CRC Project CM-138-17 that could be used to trigger driveability events on-demand. This trick car was used to conduct a driveability workshop as a part of this program.

The workshop was conducted at the Continental test track in Uvalde, TX over the course of one week. Novice and inexperienced personnel were trained and calibrated on the CRC Driveability Procedure E-28-94. A trained rater who has rated on previous CRC programs conducted the training. The training included a classroom discussion, followed by demonstrations of driveability malfunctions. The trainees were then tested on their ability to detect the malfunctions.

Four driveability events at three different severity levels, as shown in Table 1, were demonstrated by the vehicle. Test data was collected for each trainee and a performance analysis was completed. Results are summarized in this report. In most cases, the accuracies achieved by the trainees during the test were 5-15% lower than those for expert raters who helped calibrate the trick car. Challenges were noted with the stumble malfunction. To help further understand the new raters' performance, and to identify potential issues, additional training and analysis took place beyond the required scope of work and provided possible explanations for performance variations.

The trick car has proved to be an important training tool that can be incorporated into future driveability programs. The car can help the group of raters become accustomed with a baseline demonstration of driveability events at different levels of severity. This could be done as practice for driveability maneuvers, ahead of raters gaining familiarity with the test vehicles. The participants also highlighted the importance of the classroom portion of the training, as documented in Appendix 3. The participants may have benefitted by having a round table discussion on their experiences conducting maneuvers and experiencing different events and severities on the track. This sharing of knowledge between participants is typical during a CRC driveability program but was not feasible due to our compressed schedule at the track. Based on their feedback, we conclude that a combination of classroom, in vehicle (trick car) and round table discussion would benefit rater performance and the overall accuracy of the ratings for a CRC driveability program.

# 1. INTRODUCTION

CRC has developed a system and nomenclature to discriminate between different driveability events. The E-28-94 procedure was used for this workshop is shown in Figure 1, and the malfunctions and severities that can be assigned to a procedure are shown in Table 1. The method has been used for decades, but no new raters have been trained to replace existing CRC raters. The trick car, developed by SwRI for CRC project CM-138-17, is shown in Figure 2. This vehicle allows for executing malfunctions on-demand through a tablet-style computer interface. A second tablet is used to record a rater's grading of each maneuver. A detailed description of the vehicle and driveability controller functionality can be found in the final report for CM-138-17, which can be accessed at [http://crcsite.wpengine.com/wp-content/uploads/2019/05/CM-138-17-1-FINAL-REPORT_Oct.-2018.pdf](http://crcsite.wpengine.com/wp-content/uploads/2019/05/CM-138-17-1-FINAL-REPORT_Oct.-2018.pdf).

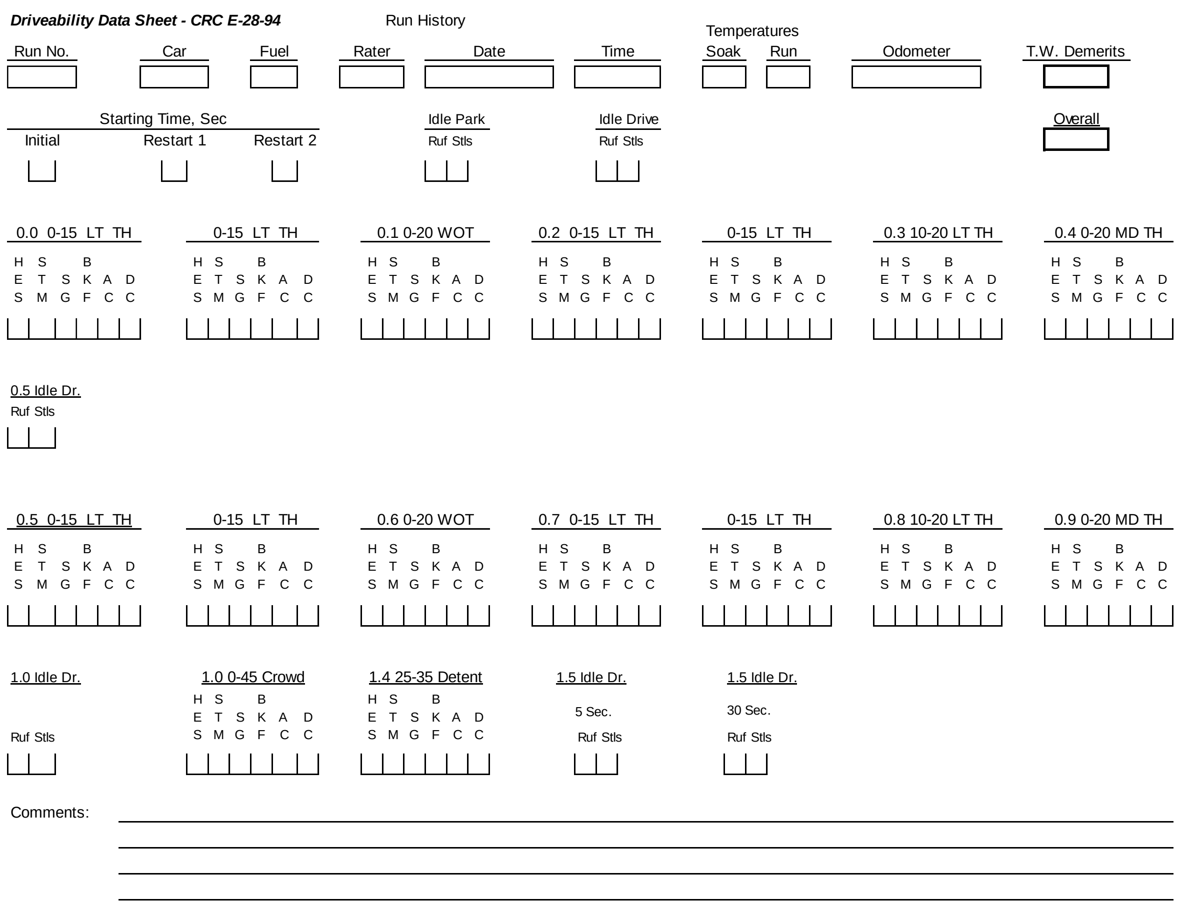


**Figure 1: CRC E-28-94 Driveability Data Sheet**

**Table 1: Driveability Events and Severity Levels**

| EVENT | SEVERITY | | |
|---|---|---|---|
| Hesitation | Trace | Moderate | Heavy |
| Stumble | Trace | Moderate | Heavy |
| Surge | Trace | Moderate | Heavy |
| Idle Quality | Trace | Moderate | Heavy |

The workshop was organized to familiarize untrained personnel in becoming driveability raters using the trick car as a training tool. An experienced rater went over the procedure, malfunctions, accelerator pedal positions, and other guidelines with each trainee. A technician was seated in the back of the vehicle and controlled the trick car malfunctions. The experienced rater demonstrated driving the procedure to each trainee while describing each malfunction as it was executed. The trainee was then asked to drive the vehicle, repeatably follow the procedure, and identify any malfunctions they experience. The trainees were initially corrected by the experienced rater and the technician. To complete the training, they were tested on their ability to identify a randomized set of malfunctions.

A stall event does not categorize into the severity levels as mentioned in Table 1, since the event results in a stopped engine. These were not demonstrated during the workshop, though they are obvious to detect and were discussed with the trainees.

**Figure 2: Ford Fusion 2014 Vehicle used for the Program**

Figure 3 shows the tablet interface of the technician who controls the trick car malfunctions. This interface allows them to modify the behavior of the vehicle by selecting the malfunction severity and associated maneuver throttle. The technician selects a malfunction, which is activated when the driver pedal input exceeds a specified threshold. The technician is also able to monitor the accelerator pedal position and provide feedback on a driver's position performance (tip-in vs. roll-in, magnitude, stability).

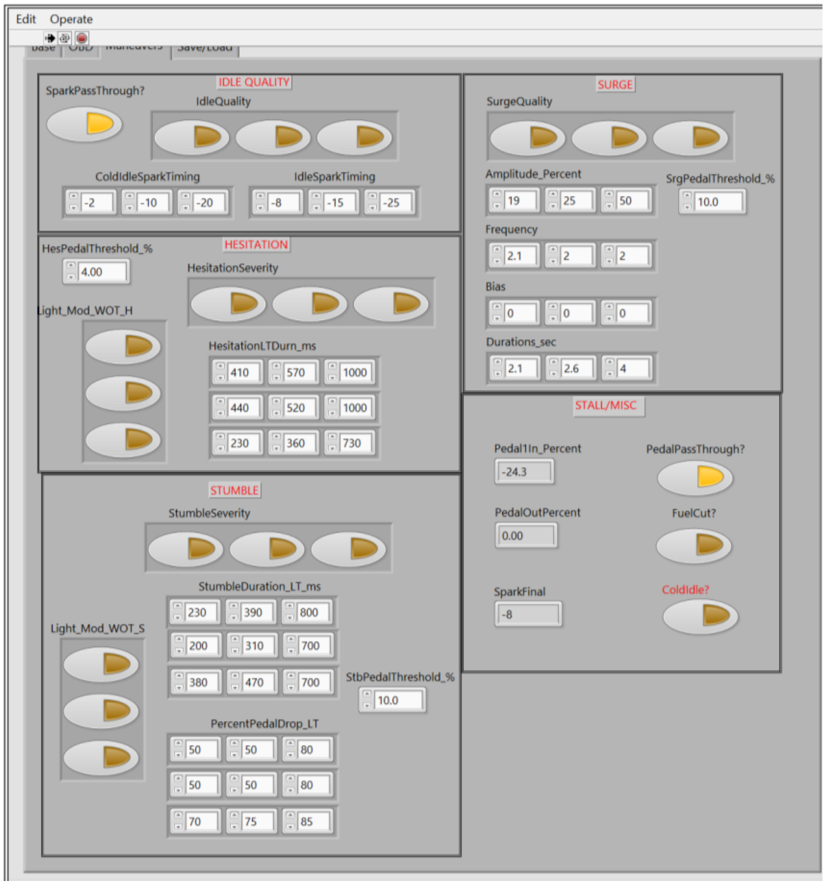


**Figure 3: Controller Tablet Interface**

Figure 4 shows two views of the rater tablet interface. During the training, the experienced rater would record the trainees' responses on this tablet while seated in the front passenger seat. The tablet presents events per the CRC E-28-94 procedure (Figure 1). An explanation of these events can be found in Appendix 1 (Definitions and Explanations, and Driveability Quick Reference). During idle events, only idle severities are presented as available options. During acceleration events, raters have the option of selecting a hesitation, stumble, and surge malfunctions, as well as their associated severities. Data from this tablet is synchronized with the vehicle controller tablet (Figure 3). This is useful for data processing, where triggered controller events are compared to the logged rater responses.

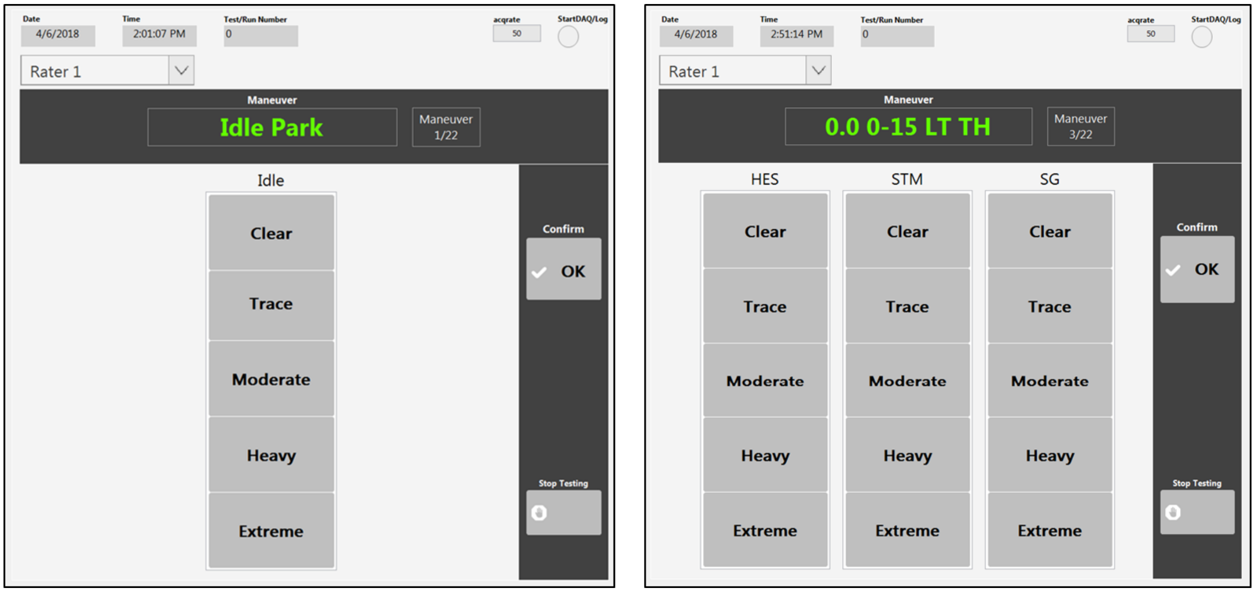


**Figure 4: Rater Tablet Interface (Left: Idle Rating; Right: Hesitation, Stumble, Surge Rating)**

## 2. OBJECTIVES

1. Lead a workshop to train novice personnel to be driveability raters using the CRC trick car
2. Track and provide the progress of participants to CRC project leadership

### 2.1. Vehicle Preparation

In preparation, basic maintenance (oil and filter change, wipers, visual inspection) and replacement of both rear tires were completed. The vehicle was then driven to the SwRI test track to verify driveability malfunction controller operation. Both operator and rater tablets were connected to the vehicle, and the vehicle was test driven, with the driver varying pedal positions and testing all the malfunctions. The vehicle was not driven by a trained rater, but malfunctions of varying severities were observed by both the driver and technician operating the malfunctions.

### 2.2. Track Selection

The track selected for this program is the Continental Uvalde Proving Grounds, located at 6969 FM 117 (Batesville Rd), Uvalde, Texas 78801. Specifically, the Dry Handling #1 course was used, and is shown in Figure 5. The track is one mile long and provides an approximately 1700 foot straight-away (highlighted in yellow) with available turnaround points show in in red. The whole straight-away was to be used for the workshop, but physical inspection of the track revealed inadequate road conditions at the extreme end of this path. The entirety of the “continuous” track surface was thus utilized for the workshop. Cones were placed at every one-tenth of a mile to indicate maneuver starting positions. These were organized to allow for crowd maneuvers to take place on the straight-away (see Appendix 2 for maneuver descriptions). Exclusive access was requested to allow for travel in either direction.

**Figure 5: Dry Handling #1 Track**

### 2.3. Training Material

Training guides and references were sourced or created for this workshop and used to help train the raters. They are listed below and are included in Appendix 2 of this document.

- Training Outline

A high-level approach to the training.

- Training Agenda

A breakdown of the workshop by the hour. This was the process followed to train new raters.

- CRC Driveability Procedure

Detailed document describing the driveability procedure step-by-step, including engine operation, maneuvers, descriptions of throttle positions, malfunctions, severities, and demerits.

- Driveability Quick Reference

A compact version of the Driveability Procedure document, this outlines throttle positions, malfunctions, and severity levels.

- Q&A from rater

A list of questions was compiled by workshop leadership, and these were answered by the experienced rater. These responses provide a valuable conversational reference to help in understanding the procedures and what to look for during rating.

- Driveability Worksheet

CRC E-28-94 Data Sheet

## 2.4. Trainee Availability and Schedule

During vehicle preparation, it was noted that a full day of training might be overwhelming for trainees (see section 5.2). Still, spending as much time in the vehicle as possible is necessary to properly calibrate a rater. A training schedule was compiled to allow for a reasonable length of training (enough for a basic understanding of the procedure, maneuvers, and malfunctions) and to allow for breaks between sessions. A final version of the schedule can be seen in Table 2.

**Table 2: Training Schedule**

| Hour | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 |
|---|---|---|---|---|---|
| 8 am | Trainee #1 | Trainee #1 | Trainee #3 | Trainee #3 | Trainee #4 |
| 9 am | | | | | |
| 10 am | | | | | |
| 11 am | | | | | |
| 12 pm | Lunch | | | | |
| 1 pm | Trainee #2 | Trainee #2 | Trainee #4 | Trainee #5 | Trainee #5 |
| 2 pm | | | | | |
| 3 pm | | | | | |
| 4 pm | | | | | |

A workshop agenda (included in Appendix 2) was formulated with inputs from SwRI engineers, CRC project leadership, and the expert rater. It was decided that a trained rater would lead workshop activities and use the trick car as a tool for demonstrating the procedure and malfunctions. The final agenda called for a short "classroom" session that included discussing general information about the procedure. The expert rater would then demonstrate throttle positions and run the cycle while describing the rating process. The trainee would then practice driving the cycle. Malfunctions would be introduced at a consistent severity level (to learn the vehicle's behavior during each malfunction), and then different severities would be introduced for each malfunction. This would be followed by driving the cycle with randomized malfunctions and severity levels. During this process, feedback was provided by both the controller technician (as to what malfunctions were triggered), as well as the expert rater (how to follow the procedure, things to look and feel for, and overall performance).

Bruce Henderson participated as the trained rater for the program. Bruce spent nearly 35 years in roles for the BP-Amoco Fuels Technology group. He has a background in conducting vehicle performance tests including octane and driveability evaluations and developing and improving rating methodologies. He contributed to many programs testing fuel composition effects on vehicle driveability and octane requirements.

The following individuals (in no particular order) participated as driveability trainees. Their company affiliation is indicated in parenthesis:

- Marie Valentine (Toyota Motor North America, Inc.)
- Lucio Dominguez III (Intertek)
- Joe Lohmann (Ret. Independent consultant)
- Chris Eisenhauer (Southwest Research Institute)
- Sergio Gonzalez (Southwest Research Institute)

**Table 3: Results of Fuel Testing**

| *ASTM Method* | Test | Units | Value |
|---|---|---|---|
| *D5191*[1] | RVP[2] | psi | 7.61 |
| | DVPE | psi | 7.48 |
| *D2699Mdp* | RON | ON | 93.8 |
| *D2700Mdp* | MON | ON | 84.5 |
| *D381* | UnWshdGm | mg/100mL | 6.5 |
| | WashdGum | mg/100mL | <0.5mg/100mL |
| *D4052* | API@60F | | 57.49 |
| | SPGr@60F | | 0.7487 |
| | Dens@15C | g/ml | 0.7484 |
| *D5599* | EtOHVol | Vol% | 9.8312 |
| | EtOHWt | Wt% | 10.4248 |
| | TtlWt | Wt% | 3.62 |
| *D86* | IBP | deg F | 106.7 |
| | Evap_5 | °F | 129.1 |
| | Evap_10 | °F | 135.1 |
| | Evap_15 | °F | 139.6 |
| | Evap_20 | °F | 143.8 |
| | Evap_30 | °F | 151.2 |
| | Evap_40 | °F | 162.2 |
| | Evap_50 | °F | 215.4 |
| | Evap_60 | °F | 238.6 |
| | Evap_70 | °F | 259.2 |
| | Evap_80 | °F | 289.5 |
| | Evap_90 | °F | 328.6 |
| | Evap_95 | °F | 357.1 |
| | FBP | °F | 417.3 |
| | Recoverd | mL | 98.7 |
| | Residue | mL | 0.9 |
| | Loss | mL | 0.4 |

[1] D5191 Scope: This test method covers the use of automated vapor pressure instruments to determine the total vapor pressure exerted in vacuum by air-containing, volatile, liquid petroleum products and liquid fuels, including automotive spark-ignition fuels with or without oxygenates and with ethanol blends up to 85 % (volume fraction) (The precision using 1 L containers was determined in a 2003 interlaboratory study (ILS); the precision using 250 mL containers was determined in a 2016 ILS.). This test method is suitable for testing samples with boiling points above 0 °C (32 °F) that exert a vapor pressure between 7 kPa and 130 kPa (1.0 psi and 18.6 psi) at 37.8 °C (100 °F) at a vapor-to-liquid ratio of 4:1. Measurements are made on liquid sample sizes in the range from 1 mL to 10 mL. No account is made for dissolved water in the sample.

[2]The RVP result reported is calculated with the EPA equation.

## 2.5. Fuel Analysis and Transportation

A single test fuel was to be used for testing to ensure malfunctions are not caused by the fuel. Fuel for the trick car was obtained from SwRI's dispensing facility. Sixty gallons of gasoline were obtained, stored in five-gallon metal containers, and kept on-site in cold storage, while a sample was sent SwRI's Petroleum Products Research Department (PPRD) for testing. Results of the fuel analysis are shown in Table 3.

The fuel was then transported to the Uvalde Proving Grounds test track facility for use in the vehicle during the workshop by a certified technician. Between five and eight gallons of gas were used each day during the workshop, and the vehicle was topped-off at the end of each day.

# 3. WORKSHOP

## 3.1. Overview

The workshop followed the processes outlined in section 2. Each trainee who signed up was present for the training. The training process was mostly identical for each trainee, with minor modifications to allow for individual differences (e.g. spending more time on a specific maneuver). The expert rater, who was not previously exposed to the trick car or rater tablets, was able to quickly adjust to the vehicle without issue.

The expert rater was able to provide thorough, detailed information on the driveability procedure, necessary techniques, and other important information. The trainees were able to understand the procedure and spent most of their time in the trick car following maneuvers and getting used to malfunctions, proper tip-in movements, and appropriate pedal positions.

Detailed workshop results from tests conducted at the end of the trainees' training sessions are found in Section 4. At a high level, test performance was mixed. This can be attributed to the vehicle as well as the participants (see Sections 3.2 and 4.4).

At an event level, in most cases, hesitation accuracy for the workshop trainees was comparable to that of the trained raters – above 90%. Hesitations are easily reproducible and are felt before the vehicle starts moving (indeed because the vehicle doesn't move). This fact at least removes the variability of track conditions. Stumble malfunctions were problematic and were often not felt (or rated as something else) by the raters; stumbles are discussed in more detail below. The surge malfunction accuracy was above 70% and is 10-15% lower than that of the trained expert raters. Idle event severity accuracy was around 50%. This is equal to, or only somewhat lower than the experts, depending on which expert is used for comparison.

**Figure 6: Entrance to Dry Handling 1 Course at the Proving Grounds**

**Figure 7: Technician (James Fritz) Controlling Malfunctions Seated in Trick Car**

### 3.2. Issues and Concerns

Though weather conditions were favorable overall (partly cloudy skies without precipitation), differences in temperature conditions must be noted as they likely attributed to a variation in trick car vehicle performance. Morning temperatures began in the low 70s (°F) with high relative humidity of above 80%. At its hottest in the late afternoon, temperatures reached into mid 90s with lower humidity.

The biggest concern during training was the effectiveness of the stumble malfunctions. Though hardware and calibrations were unchanged from CM-138-17, triggered stumbles were often not felt, particularly for trace and moderate severities. Partly as a result of these concerns, further testing was done with two of the trainees and is discussed in Section 4.4. Results from this testing link weather conditions as the likely cause behind this difference in performance, though other considerations were brought up. For example, transmission shift points affect the vehicle response and what the driver ultimately feels. This is dependent, for one, on throttle position. Slightly different throttle positions that both qualify as a “light throttle” could result in the transmission shifting at different vehicle speeds. This shift may coincide with a triggered stumble malfunction, producing unexpected behavior. Indeed, when the accelerator pedal tip-in position corresponded to the upper end of what would be considered a “light throttle”, the transmission shifted into a higher gear during the stumble.

Another point is the age of the vehicle and its fluids (hydraulic, power steering, etc…), which could change vehicle behavior. This is compounded with other variables, such as temperature, which also changes fluid viscosities. Adaptive learn on the vehicle has not been explored but may also have an effect. Further, the vehicle is only used for infrequent testing and training; it is otherwise kept outdoors and idle. Hence, while the malfunction controller and its calibrations perform in a repeatable manner, these factors are likely to contribute to differences in trick car vehicle performance.

Many of these challenges would be present in most vehicles and in driveability testing of any other vehicles. Despite this, it should be noted that the issue of most concern affects just one of the malfunctions. Additional testing (included in 4.4) showed better vehicle performance. The conclusion drawn in 4.4 suggests the trick car will provide the best performance when calibrated and tested at the same environmental conditions. Controller enhancements to compensate for some of these factors could be developed in the future to provide even better repeatability.

# 4. WORKSHOP FINAL TEST DATA ANALYSIS

The analysis will compare triggered events with rater responses. The data analysis and associated conclusions will be explained in detail for Rater 1. This pattern for analysis can be used to examine the other data sets. Each trainee was tested on the same pattern of malfunctions and severities, though more runs were completed by some trainees if time allowed.

## 4.1. Key Quantitative Review Concepts, Process, and Evaluation

The test events were distributed as shown in Figure 8. A random event generator algorithm, developed in CM-138-17, was used to achieve the more even event distribution.

- A hesitation, stumble, or surge is listed in the figure if one of those malfunctions was executed, regardless of the severity.
- Idle events are executed by default in six of the total 22 maneuvers, as listed in the E-28-94 procedure. Therefore, the percentage of idle events could deviate somewhat from that of the other malfunctions.
- A “clear” is executed if no malfunction is enabled by the operator.
- Data errors occur when there is an undefined or unhandled exception; these occur in relatively few instances and are therefore discarded from the analysis.

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 17 | 15% |
| Stumbles | 20 | 18% |
| Surges | 26 | 24% |
| Idles | 24 | 22% |
| Clears | 22 | 20% |
| Data Errors | 1 | 1% |
| | | |
| Total Maneuvers | 110 | 100% |

**Figure 8: Rater 1 Event Distribution**

Figure 9 shows the overall results for Rater 1 for each event, where the percentages are a weighted accuracy. The weights were calculated by dividing the number of times a specific event was triggered by the total number of events.

$$Weighted\ Mean\ Accuracy\ Hesitation = \frac{Number\ of\ hesitations}{Total\ Number\ of\ events} * accuracy_{hesitation}$$

"Correct Event" means the event was identified correctly, regardless of severity. For example, when a hesitation event was triggered, the rater reported a hesitation event 94% of the time, as seen in Figure 9. If the rater marked an executed event as "clear", it is considered that the rater did not correctly identify the event. Therefore, a lower percentage, such as that for the stumbles, could indicate either a confusion with other events (a stumble rated as a hesitation or surge), or lower rates of detection (rater did not feel a malfunction).

The idle events do not have a "correct event" accuracy since raters cannot rate idle maneuvers as another event. Therefore, Rater 1 correctly identified the severity of the idles 54% of the time. It is worth noting, however, that the vehicle's stock idle can be rough. Even if the controller did not enable a rougher idle (marked as "clear" on the data log), the rater would then be rating the vehicle's stock idle, which could be more severe than a "clear". Further, the stock vehicle sometimes produces an idle that could be rated as a "moderate"; requesting a "trace" idle from the controller does not produce a smoother idle. If the rater deems the idle to be indeed "moderate", this would lead to an "incorrect severity" rating as recorded on the data log.

The "Clear" percentage reflects the number of times a rater called "clear" when the operator did not trigger a malfunction. In most cases, idle events are the cause of an incorrect "clear" rating for reasons explained above. Other times during pedal maneuvers, the stock vehicle may itself produce a jitter that could be picked up as a driveability malfunction.

"Correct Severity" means that both the triggered event and its severity were identified correctly. For Rater 1, they detected a triggered hesitation 94% of the time, regardless of how they rated its severity. However, they correctly identified both the hesitation event and its triggered severity 29% of the time.

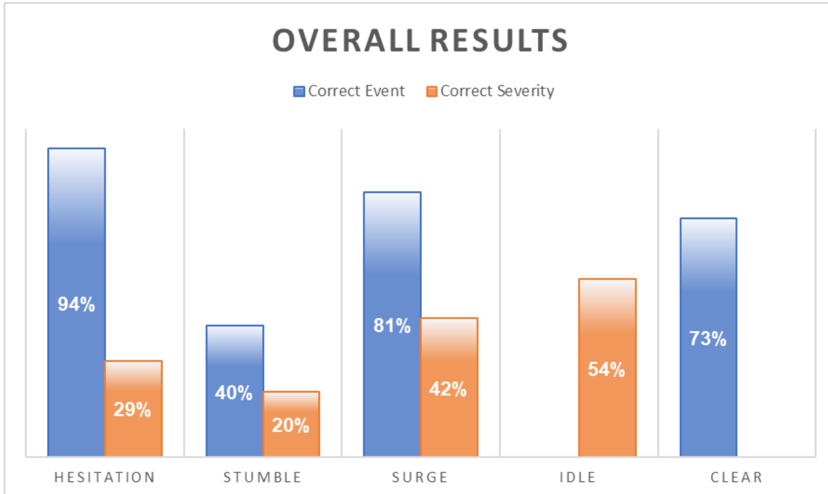


**Figure 9: Rater 1 Overall Results**

While potentially useful at a high level, there are many limitations to drawing conclusions based on a rater's overall weighted average. Therefore, the data should also be analyzed at an event level.

*Hesitations*

Calibrations used for testing hesitations are shown in Figure 10. Like stumbles, the difference between trace and moderate severity events was about 150 milliseconds, while the heavy hesitations were more spread apart.

| HESITATION (milliseconds) | | | | |
|---|---|---|---|---|
| | | Severity | | |
| | | *Trace* | *Moderate* | *Heavy* |
| **Throttle** | *Light* | 410 | 570 | 1000 |
| | *Moderate* | 440 | 520 | 1000 |
| | *Wide Open* | 230 | 360 | 730 |

**Figure 10: Final Calibration used for Hesitation Event Testing**

Figure 11 shows a detailed breakdown of Rater 1's responses to the hesitation events. The "executed event" columns represent the event that was triggered by the controller, while the "rater response" rows represent the logged responses of the raters. For example, the second column titled "Hesit Mod" indicates every time a moderate hesitation malfunction was enabled. 14% of the time, the rater labeled it a trace. They correctly identified the moderate hesitation 29% of the time, and rated it heavy 43% of the time, while 14% of the time they rated it as a different event altogether, regardless of severity.

It is worth pointing out that the number of events executed is relatively small. Only five trace, seven moderate, and five heavy hesitations were executed during the entire test. These can skew first impressions about the results. Although 20% of heavy hesitations were rated as a trace hesitation, this represents just one rated malfunction (20% of 5 runs). Thus, this data is useful for seeing patterns but not necessarily drawing concrete conclusions.

| | | Executed Event | | |
|---|---|---|---|---|
| | | Hesit. Trace | Hesit. Mod | Hesit. Hvy |
| Rater Response | Clear | 0% | 0% | 0% |
| | Trace | **0%** | 14% | 20% |
| | Moderate | 100% | **29%** | 20% |
| | Heavy | 0% | 43% | **60%** |
| | Incorrect Event Rated | 0% | 14% | 0% |
| | Total Number of Tests | 5 | 7 | 5 |
| | Average Rating (Value) | **3.0** | **3.3** | **3.4** |
| | Deviation | 0.00 | 0.75 | 0.80 |
| | Correct Event (Excl. Clears) | | | **94%** |
| | Correct Event and Severity | | | **29%** |

**Figure 11: Rater 1 Hesitation Event Results**

Some inferences can still be made from Figure 11. First, the rater was able to pick up on the hesitations overall with a high accuracy. Rarely (in fact, only once) did the rater incorrectly call a hesitation as either a stumble or a surge, and they were able to detect every single malfunction (never calling a triggered hesitation “clear”). However, their precision is off and implies they are more sensitive than the trick car calibration: trace hesitations were rated more severely (as moderates), and moderates tended to be rated closer to “heavy”. Heavy hesitations were more consistently rated the highest. Extreme severities were never called during the testing and are therefore not included in the graphics.

To help with understanding these tendencies, an average rating (value) has been assigned to each severity, as shown in Table 4. Trace hesitations, ideally averaging at 2.0, were rated such that their average value is 3.0, implying the rater tended to rate these more heavily. The 3.3 rating for moderate hesitations (ideally 3.0) again says the rater tended to rate these more heavily as well. The 3.4 rating for heavy hesitations (ideally 4.0) suggests the rater didn’t rate these heavy enough but did perceive them to be more severe than the others.

The accuracy, or the consistency, of responses is captured in the “deviation” value, described in Equation 1. A deviation of zero implies all values match the mean. This can be seen in the trace hesitations of Figure 11, where all responses, although not accurate, were very consistent. In contrast, a high deviation value would suggest a large spread in responses. If half of responses were “clear”, and half were “heavy”, the deviation value would be greatest at 1.5.

**Equation 1: Response Deviation Calculation**

$$\sqrt{\sum^{Each\ Severity} \frac{Weighted\ Severity\ Occurence}{(Severity\ Rating - Average\ Rating)^2}}$$

**Table 4: Severity Key Established to Capture Sensitivity Differences**

| | |
|---|---|
| 1.0 | Clear |
| 2.0 | Trace |
| 3.0 | Moderate |
| 4.0 | Heavy |
| 5.0 | Extreme |

*Stumbles*

The final calibration for the stumble event is shown in Figure 12. The upper part of the chart shows the duration of power drop, while the bottom part shows the actual drop in pedal percentage. The raters have a tight band in which to distinguish a trace versus a moderate severity. For example, trace severity for a light throttle maneuver has a duration of 230 milliseconds. For the same light throttle maneuver, a moderate severity has a duration of 390 milliseconds, a difference of 160 milliseconds. A moderate to heavy severity has a more noticeable difference of about 400 milliseconds.

| STUMBLE DURATION (MILLISECONDS) | | | | |
|---|---|---|---|---|
| | | Severity | | |
| | | *Trace* | *Moderate* | *Heavy* |
| **Throttle** | *Light* | 230 | 390 | 800 |
| | *Moderate* | 200 | 310 | 690 |
| | *Wide Open* | 380 | 470 | 700 |

| STUMBLE PEDAL DROP (PERCENT) | | | | |
|---|---|---|---|---|
| | | Severity | | |
| | | *Trace* | *Moderate* | *Heavy* |
| **Throttle** | *Light* | 50 | 50 | 80 |
| | *Moderate* | 50 | 50 | 80 |
| | *Wide Open* | 70 | 75 | 85 |

**Figure 12: Calibrations Tested for a Stumble Event**

Figure 13 shows the detailed breakdown of Rater 1's responses to stumble events. The detection of stumbles for all raters during the workshop in general is quite low. Trace stumbles were almost never detected. Moderate stumbles were detected more often, but also had the tendency of being called a different event altogether. Heavy stumbles were more perceptible, but the deviation value shows that the severity is not consistent. Possible issues with stumbles on the trick car during the workshop are discussed in sections 3.2, 4.4, and 5.1.

| | | Executed Event | | |
|---|---|---|---|---|
| | | Stumble Trace | Stumble Mod | Stumble Hvy |
| Rater Response | Clear | 80% | 25% | 29% |
| | Trace | **0%** | 13% | 14% |
| | Moderate | 20% | **25%** | 14% |
| | Heavy | 0% | 0% | **29%** |
| | Incorrect Event Rated | 0% | 38% | 14% |
| | Total Number of Tests | **5** | **8** | **7** |

| | | | |
|---|---|---|---|
| Average Rating (Value) | **1.4** | **2.0** | **2.5** |
| Deviation | 0.80 | 0.89 | 1.26 |

| | |
|---|---|
| Correct Event (Excl. Clears) | **40%** |
| Correct Event and Severity | **20%** |

**Figure 13: Rater 1 Stumble Event Results**

*Surges*

Final calibrations used for the surges are shown in Figure 14. The amplitude dictates the attenuation from the base pedal value. A frequency of around two hertz was used across all severities. The total duration dictated how long the surge event lasted.

| SURGE | | | |
|---|---|---|---|
| | Severity | | |
| | *Trace* | *Moderate* | *Heavy* |
| *Amplitude (percent)* | 19 | 25 | 50 |
| *Frequency (Hertz)* | 2.1 | 2 | 2 |
| *Total Duration (seconds)* | 2.1 | 2.6 | 4 |

**Figure 14: Final Calibration used for Surge Event Testing**

Rater 1's detailed surge breakdown is shown in Figure 15. The average rating for trace stumbles, although a perfect 2.0, is misleading because of the inconsistency of responses. Nearly half of executed trace surges were not detected at all, while those that were detected were not consistently rated. Moderate surges were all detected, but more consistently rated higher than the executed malfunction; some more rater calibration is likely needed. Heavy surges were well detected with high accuracy and precision. It is interesting to note that none of the surge events were rated as another event (as a hesitation or stumble).

| | | Executed Event | | |
|---|---|---|---|---|
| | | Surge Trace | Surge Mod | Surge Hvy |
| Rater Response | Clear | 45% | 0% | 0% |
| | Trace | **18%** | 13% | 0% |
| | Moderate | 27% | **38%** | 14% |
| | Heavy | 9% | 50% | **86%** |
| | Incorrect Event Rated | 0% | 0% | 0% |
| | Total Number of Tests | **11** | **8** | **7** |

| | | | |
|---|---|---|---|
| Average Rating (Value) | **2.0** | **3.4** | **3.9** |
| Deviation | 1.04 | 0.70 | 0.35 |

| | |
|---|---|
| Correct Event (Excl. Clears) | **81%** |
| Correct Event and Severity | **42%** |

**Figure 15: Rater 1 Surge Event Results**

*Idle Quality*

The final idle calibration is shown in Figure 16.

| IDLE | | | |
|---|---|---|---|
| | Severity | | |
| | *Trace* | *Moderate* | *Heavy* |
| *Spark Timing (degrees before firing TDC)* | -8 | -15 | -25 |
| *Random noise limits* | ± 5 | ± 10 | ± 15 |

**Figure 16: Final Calibration used for Idle Quality Testing**

The idle events for Rater 1 are shown in Figure 17. The heavy idles are obvious enough that they are consistently rated correctly. Moderate idles likely need a bit more rater calibration, as it would be preferable to see fewer of these rated as "trace". Trace idles are a bit more problematic due to reasons discussed at the beginning of this analysis, but overall performance is reasonable.

| | Executed Event | | |
|---|---|---|---|
| | Idle Trace | Idle Mod | Idle Hvy |
| Rater Response: Clear | 44% | 0% | 0% |
| Trace | **22%** | 33% | 0% |
| Moderate | 33% | **56%** | 0% |
| Heavy | 0% | 11% | **100%** |
| Incorrect Event Rated | 0% | 0% | 0% |
| Total Number of Tests | 9 | 9 | 6 |

| | | | |
|---|---|---|---|
| Average Rating (Value) | 1.9 | 2.8 | 4.0 |
| Deviation | 0.87 | 0.63 | 0.00 |

| | |
|---|---|
| Correct Event (Excl. Clears) | 83% |
| Correct Event and Severity | 54% |

**Figure 17: Rater 1 Idle Event Results**

*Clears*

The combined accuracy when no events were triggered was about 73%. This means that when the operator did not trigger a malfunction, the rater correctly called "clear" 73% of the time. Trace, moderate, or heavy responses of any malfunction (hesitation, stumble, surge, or idle) rated during a "clear" event are listed in Figure 18. As previously explained, most of the incorrect clear responses can be attributed to the idle events: enabling a clear idle on the controller does not produce a smoother idle on the vehicle. For this rater, excluding incorrect idle responses results in a "clear" accuracy of 94%. Other raters would also see a 20-30% improvement in the clear results if the idle events were not included.

| | Clear |
|---|---|
| Rater Response: Clear | **73%** |
| Trace | 14% |
| Moderate | 14% |
| Heavy | 0% |
| Incorrect Event Rated | |
| Total Number of Tests | 22 |

**Figure 18: Rater 1 No Triggered Event Results**

## 4.2. Participant Test Results

Detailed test results can be found in Appendix 1. The following generalizes the outcomes:

- Most raters were able to correctly identify hesitations over 95% of the time, and correctly identified the severity in roughly 30-40% of cases.

- Raters were able to correctly identify stumbles 20-40% of the time, and correctly identified the severity in roughly 20% of cases.
- Raters were able to correctly identify surges over 70% of the time, and correctly identified the severity in roughly 35% of cases.
- Raters were able to correctly identify the idle severities in over 50% of cases, though this does include "incorrectly" rated clears.

These high-level numbers carry significant nuances and should not be solely relied upon to draw conclusions. It is recommended that the detailed results are reviewed for a more accurate representation of the success of the workshop.

### 4.3. Qualitative Results and Feedback

A survey was conducted at the conclusion of the workshop. Feedback was provided by the trainees and are summarized below. Complete questions and responses to the surveys can be found in Appendix 3.

Some common matters addressed in the responses highlight the importance of:

- The classroom portion of the training to the trainees in understanding the procedure and recognizing the meanings of terminology
- Interactions with the trained rater, to include
- Demonstration by the instructor of procedures and maneuvers
- Help from the instructor in identifying events and their severities

Other responses focused on:

- Overall confidence in conceptual understanding of driveability testing, maneuvers, and severities
- Good experience with the expert rater
- Positive feedback on the track, but some confusion with cone placement
- Comfort with safety but concerns about not wearing seatbelts. This practice was carried over from the expert raters in CM-138-17.
- Overall the trainees were somewhat confident in their ability in rating vehicles, with everyone asking for more seat-time in the vehicle

### 4.4. Additional Consistency Testing

Beyond the training workshop, SwRI undertook a separate training and testing study with two of the raters – Rater 1 and Rater 4. Fixed severity testing was performed to further confirm rater consistency and understand issues with the stumbles during the workshop. This mirrors an exercise performed with expert raters in CM-138-17, in which only moderate severity events (and clears) were triggered. Raters were not aware of this test methodology.

The testing took place on the SwRI test track. Though track conditions are not favorable for driveability rating, cone placement was modified slightly from the standard convention of 0.1 mile-increment placement to avoid rough patches and minimize feedback from the road. Some cones were thus placed further, and some closer, than the recommended 0.1-mile increment. Nonetheless, trials and testing confirmed that drivers had plenty of driving and stopping distance.

The first rater (Rater 4) was tested in the morning, when the ambient air temperature was in the 70s. They were given two practice runs during which the procedure was reviewed and were given immediate feedback on which malfunctions were triggered (which included trace, moderate, and heavy severities). They were then tested on several test runs with all malfunctions at the moderate severity.

It was noted by both the driver and the malfunction technician that vehicle performance appeared to be different. The rater's overall results can be seen in Figure 19, while their detailed numbers are found in Figure 20. The moderate stumbles were perceived more often by both the driver and technician.

- Hesitations were correctly identified 98% of the time (compared to 100% during the workshop)
- Stumbles were correctly identified 57% of the time (compared to just 20% overall and 25% for moderate stumbles)
- Surges were correctly identified 93% of the time (compared to 72% overall and 89% for moderate surges)

Overall, results seem to indicate improvement over the workshop test results, but this could be attributed to either increase in performance of the driver or the trick car. However, accuracy was still quite low. Their deviation (except for the stumbles) was close in accuracy to their performance at the workshop.

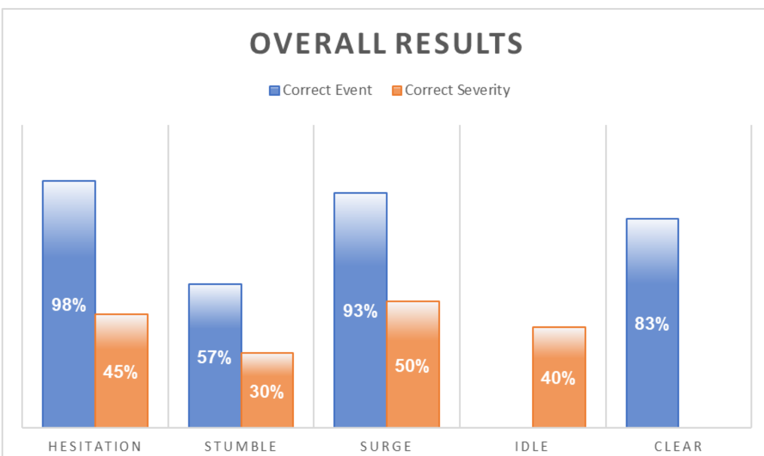


**Figure 19: Rater 4 Extended Testing Overall Results**

| | | Executed Event | | | | |
|---|---|---|---|---|---|---|
| | | Hesit. Mod | Stumble Mod | Surge Mod | Idle Mod | Clear |
| Rater Response | Clear | 2% | 39% | 2% | 22% | 83% |
| | Trace | 41% | 18% | 23% | 38% | 15% |
| | Moderate | 45% | 30% | 50% | 40% | 2% |
| | Heavy | 12% | 9% | 20% | 0% | 0% |
| | Incorrect Event Rated | 0% | 5% | 5% | 0% | |
| | Total Number of Tests | 49 | 44 | 44 | 45 | 59 |

| | | | | |
|---|---|---|---|---|
| Average Rating (Value) | 2.7 | 2.1 | 2.9 | 2.2 |
| Deviation | 0.71 | 1.04 | 0.74 | 0.77 |

**Figure 20: Rater 4 Extended Testing Detailed Results**

Performance seemed to change yet again when the next rater (Rater 1) drove the vehicle later in the afternoon during their practice runs, when ambient temperatures had climbed into the upper 80s. At this time, the moderate stumbles were not perceived at all, and many of the heavy stumbles were not as severe as they should have been (see Figure 21).

| | | Stumble Mod | Stumble Hvy |
|---|---|---|---|
| Rater Response | Clear | 100% | 33% |
| | Trace | 0% | 50% |
| | Moderate | 0% | 17% |
| | Heavy | 0% | 0% |
| | Incorrect Event Rated | 0% | 0% |
| | Total Number of Tests | 6 | 6 |

**Figure 21: Rater 1 Extended Testing Practice Run**

At this point the Rater was asked to continue testing on a different, cooler day, when ambient temperatures were in the upper 60s to low 70s. Overall results for this testing are shown in Figure 22. Again, performance improved, particularly for the stumble malfunctions.

- Hesitations were correctly identified 95% of the time (compared to 94% during the workshop overall and 86% for moderate hesitations)
- Stumbles were correctly identified 75% of the time (compared to 40% overall and 38% for moderate stumbles)
- Surges were correctly identified 84% of the time (compared to 83% overall and 100% for moderate surges)

More surges during this testing were incorrectly identified by the rater as stumbles and could indicate a need for more training. Like Rater 4, severity accuracy was still low, and could indicate a need for more training.

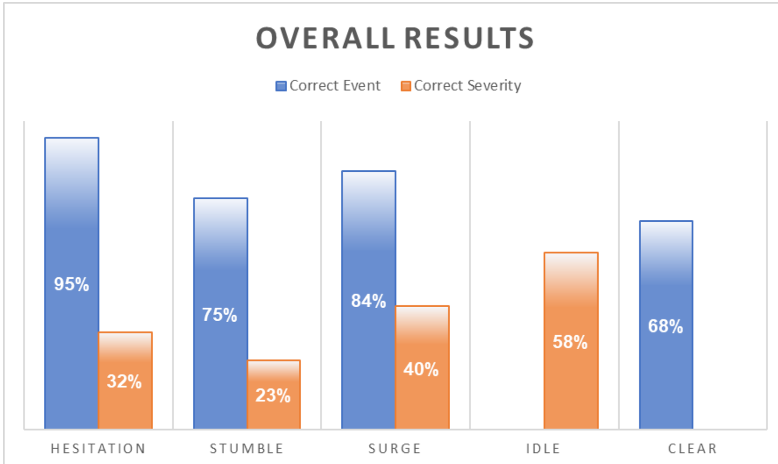


**Figure 22: Rater 1 Extended Testing Overall Results (excl. Initial Run)**

| Rater Response | Executed Event: Hesit. Mod | Stumble Mod | Surge Mod | Idle Mod | Clear |
|---|---|---|---|---|---|
| Clear | 0% | 18% | 0% | 8% | **68%** |
| Trace | 0% | 40% | 12% | 25% | 19% |
| Moderate | **32%** | **23%** | **40%** | **58%** | 10% |
| Heavy | 63% | 13% | 32% | 10% | 3% |
| Incorrect Event Rated | 5% | 8% | 16% | 0% | |
| Total Number of Tests | **19** | **40** | **25** | **40** | **31** |

| | | | | |
|---|---|---|---|---|
| Average Rating (Value) | **3.7** | **2.3** | **3.2** | **2.7** |
| Deviation | 0.47 | 0.93 | 0.68 | 0.75 |

**Figure 23: Rater 1 Extended Testing Detailed Results**

The extended testing appears to point to the difference in ambient temperatures as the reason behind poor surge performance during the workshop, since performance improved when tests were conducted at lower ambient temperatures (to note, the trick car was calibrated when temperatures were in the 60s). The raters still showed accuracy deviation at the severity level, but overall accuracies are not significantly worse when compared to expert raters' fixed severity testing (see CM-138-17).

# 5. TAKEAWAYS AND RECOMMENDATIONS

## 5.1. Vehicle

The trick car provided a valuable tool for the training of raters. Its ability to enable malfunctions on-demand was important for showing different malfunctions and severities to trainees. Further, the connected controller and rater tablets provide a good method for data collection and synchronization. As data analysis methods for these tests evolve and improve, this allows for relatively quick numerical feedback on the raters.

Vehicle performance was not flawless, as discussed in Sections 3.2 and 4.1. Ambient temperature variations appeared play a role in vehicle behavior, especially during the stumble malfunction. There are many variables within the vehicle itself that make it difficult to achieve a perfectly reproducible execution, and test results are just a snapshot of a certain rater in very specific conditions. There are many degrees of freedom, and conditions will cause a difference in vehicle performance.

The expert rater provided feedback that the vehicle and rater tablet were simple to get used to. The experienced technician works to control the vehicle's malfunctions, but the rating procedure can otherwise be followed as if the trick car was a normal vehicle.

## 5.2. Raters

Learning the cycle is important but driving the cycle correctly (i.e. with proper pedal positioning and movement) is paramount for consistent driveability rating. Raters must be experts in precisely controlling their foot position. Of further note to the physical capabilities of raters, differences in physique and driving positions need to be considered. The procedure itself involves many abrupt accelerations and decelerations and rating many vehicles would involve significant seat time. A rater should be capable of enduring these conditions, which can lead to exhaustion and motion sickness.

During this workshop, raters were professional and performed to a suitable level. They were able to understand the procedure, asked questions about unclear actions and situations, and worked well with the experienced rater to understand the procedures.

## 5.3. Recommendations for Future Training and Workshops

Certain considerations must be recognized to further optimize training in future workshops. In order to reduce variability between workshops, it would be optimal to both calibrate and use the trick car in similar climatic conditions. Besides trying to match weather conditions outdoors, the trick car could be used on a dynamometer in a climate-controlled facility if a dyno is deemed acceptable for rating.

One possibly useful addition to the trick car is a throttle position gauge. Such a device could be useful feedback to the rater on their absolute pedal position, the manner in which they press into the pedal, and the stability of their pedal during acceleration. Similar to the vacuum gauges used in other driveability workshops and studies, this could be both a useful tool and a distraction: drivers may become fixated on the gauge, alter their throttle based on the reading, and pay less attention to vehicle performance.

If a new vehicle were considered, a different model with other components could be studied. For example, the transmission in this trick car was noted as a possible problem spot. Gear changes are somewhat rough and mirror a triggered malfunction in the form of a stumble. A smoother transmission could eliminate this variable. Also, a turbocharger produces non-linear acceleration. A naturally aspirated engine may reduce this effect. These changes could allow the driver to focus solely on the created malfunction, and less on vehicle behavior that is specific to the trick car.

This first workshop with this trick car showed that it can be a very useful training tool. Beyond just training raters, the trick car could be used as a reference vehicle in future driveability studies. For example, after rating some number of other vehicles, a rater could drive the trick car and experience a set of malfunctions to help them "recalibrate" to a common standard.

## 6. CONCLUSIONS

The driveability rating workshop conducted as part of this program was successful. The experience helped five trainees understand the CRC E-28-94 driveability procedure, proper driving technique for rating vehicles, and the behavior of a vehicle undergoing specific malfunctions. Analysis of the trainees' tests, as well as additional testing done beyond the workshop, showed results that were in some cases comparable to those of experienced raters. Causes for lower accuracies for some situations were investigated, and key variables that likely resulted in trick car performance differences were identified.

Several future developments would be valuable for the advancement of driveability rating efforts:

- A request common to all trainees was for more training time in the trick car. The schedule of this workshop allowed for a basic understanding of driveability rating, but more practice is needed to adjust and calibrate raters. A new workshop may be scheduled to provide more training to existing or new participants.

- Efforts may be made to explore the effectiveness of performing driveability training and studies on a dynamometer, particularly in a climate-controlled facility. This would remove some intractable variables that have been shown to cause performance variation in the trick car.

- The E-28-94 procedure is outdated. It was developed in 1994 for vehicles of the time. Cars have evolved substantially, and a redesign may be necessary. While the procedure sequences maneuvers in a way that's easy to implement and doesn't generate too much engine heat too quickly, it does not consider newer technologies like auto-crank starting, AWD/FWD, selectable transmission modes, tractions control, and CVTs. Further, the method for establishing throttle positions needs review.

- Finally, the trick car was calibrated from the combined feedback of two expert raters, who themselves have different sensitivities. It may be desirable to develop a new trick car calibration that is reflective of a different audience, such as the general consumer. For example, more raters (to include the newly trained group) could be surveyed to determine if there is a consensus on rating severities. In addition, polling the general public could be beneficial, since such a group represents the end user.

## 7. CLOSURE

SwRI would like to thank the CRC and its members for funding this effort and is excited to participate in training new driveability raters. If you have any further questions, please contact Stanislav Gankov at sgankov@swri.org or at (210) 522-6206.

# APPENDIX 1
# PARTICIPANT TEST RESULTS

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 17 | 15% |
| Stumbles | 20 | 18% |
| Surges | 26 | 24% |
| Idles | 24 | 22% |
| Clears | 22 | 20% |
| Data Errors | 1 | 1% |
| | | |
| **Total Maneuvers** | **110** | **100%** |

**Figure 24: Rater 1 Event Distribution**

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 30 | 23% |
| Stumbles | 18 | 14% |
| Surges | 26 | 20% |
| Idles | 26 | 20% |
| Clears | 30 | 23% |
| Data Errors | 2 | 2% |
| | | |
| **Total Maneuvers** | **132** | **100%** |

**Figure 25: Rater 2 Event Distribution**

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 25 | 16% |
| Stumbles | 32 | 21% |
| Surges | 33 | 21% |
| Idles | 31 | 20% |
| Clears | 31 | 20% |
| Data Errors | 2 | 1% |
| | | |
| **Total Maneuvers** | **154** | **100%** |

**Figure 26: Rater 3 Event Distribution**

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 24 | 18% |
| Stumbles | 25 | 19% |
| Surges | 29 | 22% |
| Idles | 29 | 22% |
| Clears | 25 | 19% |
| Data Errors | | 0% |
| | | |
| **Total Maneuvers** | **132** | **100%** |

**Figure 27: Rater 4 Event Distribution**

| Totals (All Severities) | | |
|---|---|---|
| Hesitations | 38 | 22% |
| Stumbles | 29 | 16% |
| Surges | 34 | 19% |
| Idles | 29 | 16% |
| Clears | 43 | 24% |
| Data Errors | 3 | 2% |
| | | |
| **Total Maneuvers** | **176** | **100%** |

**Figure 28: Rater 5 Event Distribution**

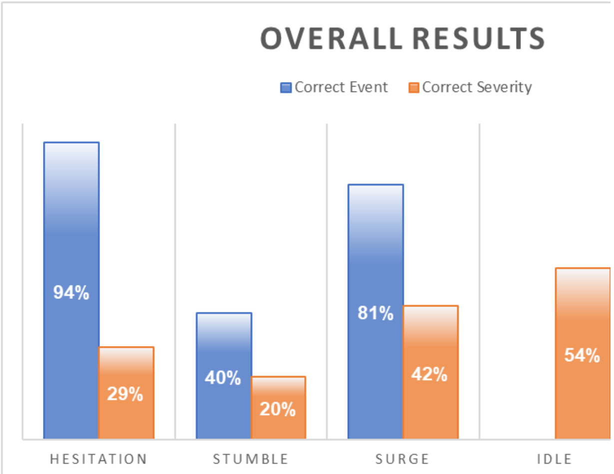


**Figure 29: Rater 1 Overall Results**

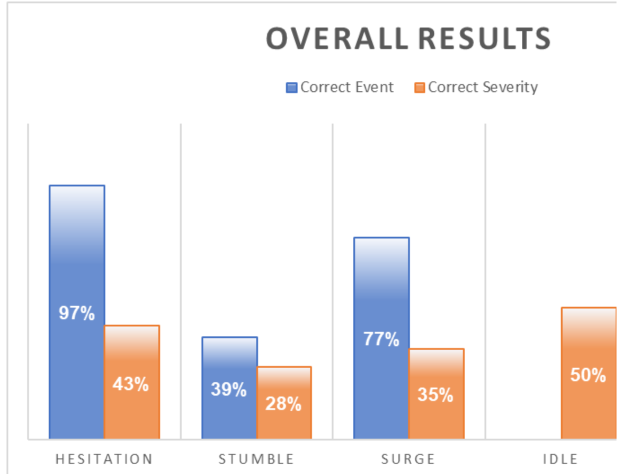


**Figure 30: Rater 2 Overall Results**

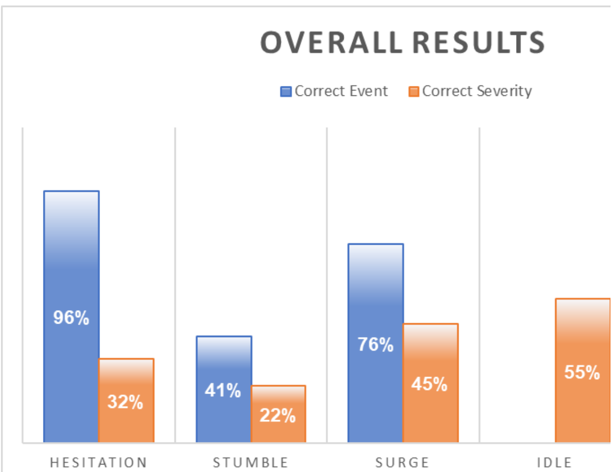


**Figure 31: Rater 3 Overall Results**

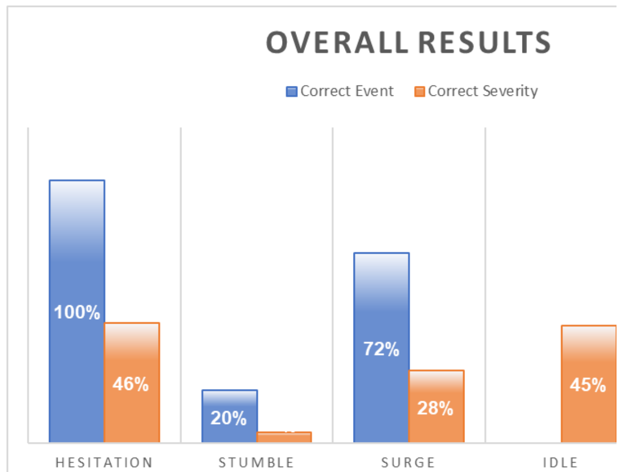


**Figure 32: Rater 4 Overall Results**

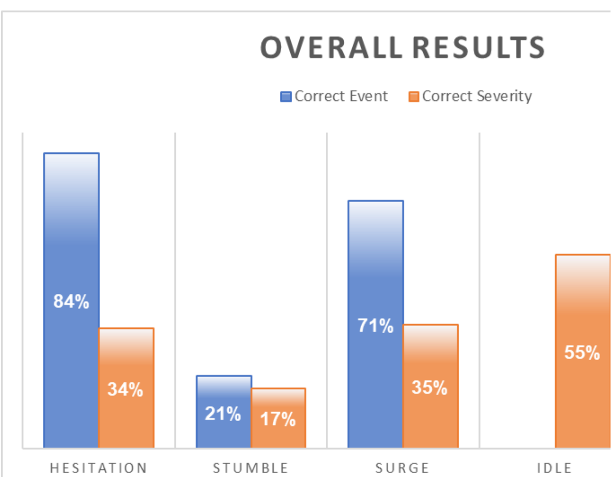


**Figure 33: Rater 5 Overall Results**

| Rater Response | Executed Event | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Hesit. Trace | Hesit. Mod | Hesit. Hvy | Stumble Trace | Stumble Mod | Stumble Hvy | Surge Trace | Surge Mod | Surge Hvy | Idle Trace | Idle Mod | Idle Hvy | Clear |
| Clear | 0% | 0% | 0% | 80% | 25% | 29% | 45% | 0% | 0% | 44% | 0% | 0% | 73% |
| Trace | 0% | 14% | 20% | 0% | 13% | 14% | 18% | 13% | 0% | 22% | 33% | 0% | 14% |
| Moderate | 100% | 29% | 20% | 20% | 25% | 14% | 27% | 38% | 14% | 33% | 56% | 0% | 14% |
| Heavy | 0% | 43% | 60% | 0% | 0% | 29% | 9% | 50% | 86% | 0% | 11% | 100% | 0% |
| Incorrect Event Rated | 0% | 14% | 0% | 0% | 38% | 14% | 0% | 0% | 0% | 0% | 0% | 0% | |
| Total Number of Tests | 5 | 7 | 5 | 5 | 8 | 7 | 11 | 8 | 7 | 9 | 9 | 6 | 22 |
| Average Rating (Value) | 3.0 | 3.3 | 3.4 | 1.4 | 2.0 | 2.5 | 2.0 | 3.4 | 3.9 | 1.9 | 2.8 | 4.0 | |
| Deviation | 0.00 | 0.75 | 0.80 | 0.80 | 0.89 | 1.26 | 1.04 | 0.70 | 0.35 | 0.87 | 0.63 | 0.00 | |
| Correct Event (Excl. Clears) | | | 94% | | | 40% | | | 81% | | | 83% | 73% |
| Correct Event and Severity | | | 29% | | | 20% | | | 42% | | | 54% | |

**Figure 34: Rater 1 Test Detailed Event Results**

| Rater Response | Executed Event | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Hesit. Trace | Hesit. Mod | Hesit. Hvy | Stumble Trace | Stumble Mod | Stumble Hvy | Surge Trace | Surge Mod | Surge Hvy | Idle Trace | Idle Mod | Idle Hvy | Clear |
| Clear | 0% | 0% | 0% | 83% | 20% | 43% | 13% | 10% | 0% | 71% | 11% | 0% | 73% |
| Trace | 33% | 0% | 0% | 17% | 20% | 0% | 13% | 10% | 13% | 29% | 56% | 10% | 20% |
| Moderate | 33% | 33% | 22% | 0% | 20% | 0% | 25% | 20% | 13% | 0% | 22% | 0% | 3% |
| Heavy | 33% | 67% | 67% | 0% | 20% | 43% | 0% | 60% | 75% | 0% | 11% | 90% | 3% |
| Incorrect Event Rated | 0% | 0% | 11% | 0% | 20% | 14% | 50% | 0% | 0% | 0% | 0% | 0% | |
| Total Number of Tests | 12 | 9 | 9 | 6 | 5 | 7 | 8 | 10 | 8 | 7 | 9 | 10 | 30 |
| Average Rating (Value) | 3.0 | 3.7 | 3.8 | 1.2 | 2.5 | 2.5 | 2.3 | 3.3 | 3.6 | 1.3 | 2.3 | 3.8 | |
| Deviation | 0.82 | 0.47 | 0.43 | 0.37 | 1.12 | 1.50 | 0.83 | 1.00 | 0.70 | 0.45 | 0.82 | 0.60 | |
| Correct Event (Excl. Clears) | | | 97% | | | 39% | | | 77% | | | 77% | 73% |
| Correct Event and Severity | | | 43% | | | 28% | | | 35% | | | 50% | |

**Figure 35: Rater 2 Test Detailed Event Results**

| Rater Response | Hesit. Trace | Hesit. Mod | Hesit. Hvy | Stumble Trace | Stumble Mod | Stumble Hvy | Surge Trace | Surge Mod | Surge Hvy | Idle Trace | Idle Mod | Idle Hvy | Clear |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Executed Event | | | | | | | | | | | | |
| Clear | 0% | 0% | 14% | 83% | 67% | 27% | 50% | 0% | 0% | 0% | 0% | 0% | **65%** |
| Trace | **25%** | 30% | 0% | **17%** | 27% | 9% | **33%** | 27% | 10% | **100%** | 67% | 29% | 35% |
| Moderate | 25% | **10%** | 14% | 0% | **0%** | 9% | 0% | **27%** | 10% | 0% | **20%** | 0% | 0% |
| Heavy | 50% | 60% | **71%** | 0% | 0% | **55%** | 0% | 45% | **80%** | 0% | 13% | **71%** | 0% |
| Incorrect Event Rated | 0% | 0% | 0% | 0% | 7% | 0% | 17% | 0% | 0% | 0% | 0% | 0% | |
| Total Number of Tests | 8 | 10 | 7 | 6 | 15 | 11 | 12 | 11 | 10 | 9 | 15 | 7 | 31 |
| Average Rating (Value) | 3.3 | 3.3 | 3.4 | 1.2 | 1.3 | 2.9 | 1.4 | 3.2 | 3.7 | 2.0 | 2.5 | 3.4 | |
| Deviation | 0.83 | 0.90 | 1.05 | 0.37 | 0.45 | 1.31 | 0.49 | 0.83 | 0.64 | 0.00 | 0.72 | 0.90 | |
| Correct Event (Excl. Clears) | | | 96% | | | 41% | | | 76% | | | 100% | 65% |
| Correct Event and Severity | | | 32% | | | 22% | | | 45% | | | 55% | |

**Figure 36: Rater 3 Test Detailed Event Results**

| Rater Response | Hesit. Trace | Hesit. Mod | Hesit. Hvy | Stumble Trace | Stumble Mod | Stumble Hvy | Surge Trace | Surge Mod | Surge Hvy | Idle Trace | Idle Mod | Idle Hvy | Clear |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Executed Event | | | | | | | | | | | | |
| Clear | 0% | 0% | 0% | 67% | 58% | 43% | 45% | 0% | 0% | 40% | 0% | 14% | **60%** |
| Trace | **38%** | 22% | 0% | **0%** | 25% | 0% | **18%** | 22% | 0% | **30%** | 42% | 0% | 28% |
| Moderate | 63% | **33%** | 29% | 0% | **0%** | 14% | 18% | **22%** | 56% | 30% | **42%** | 14% | 12% |
| Heavy | 0% | 44% | **71%** | 0% | 0% | **14%** | 0% | 44% | **44%** | 0% | 17% | **71%** | 0% |
| Incorrect Event Rated | 0% | 0% | 0% | 33% | 17% | 29% | 18% | 11% | 0% | 0% | 0% | 0% | |
| Total Number of Tests | 8 | 9 | 7 | 6 | 12 | 7 | 11 | 9 | 9 | 10 | 12 | 7 | 25 |
| Average Rating (Value) | 2.6 | 3.2 | 3.7 | 1.0 | 1.3 | 2.0 | 1.7 | 3.3 | 3.4 | 1.9 | 2.8 | 3.4 | |
| Deviation | 0.48 | 0.79 | 0.45 | 0.00 | 0.46 | 1.26 | 0.82 | 0.83 | 0.50 | 0.83 | 0.72 | 1.05 | |
| Correct Event (Excl. Clears) | | | 100% | | | 20% | | | 72% | | | 83% | 60% |
| Correct Event and Severity | | | 46% | | | 4% | | | 28% | | | 45% | |

**Figure 37: Rater 4 Test Detailed Event Results**

| | | Executed Event | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Hesit. Trace | Hesit. Mod | Hesit. Hvy | Stumble Trace | Stumble Mod | Stumble Hvy | Surge Trace | Surge Mod | Surge Hvy | Idle Trace | Idle Mod | Idle Hvy | Clear |
| Rater Response | Clear | 25% | 0% | 8% | 80% | 57% | 30% | 54% | 0% | 0% | 0% | 0% | 0% | **40%** |
| | Trace | **19%** | 20% | 0% | **20%** | 0% | 10% | **8%** | 0% | 8% | **30%** | 0% | 0% | 9% |
| | Moderate | 31% | **0%** | 8% | 0% | **14%** | 0% | 8% | **13%** | 15% | 50% | **55%** | 13% | 42% |
| | Heavy | 19% | 80% | **83%** | 0% | 0% | **20%** | 8% | 88% | **77%** | 20% | 45% | **88%** | 9% |
| | Incorrect Event Rated | 6% | 0% | 0% | 0% | 29% | 40% | 23% | 0% | 0% | 0% | 0% | 0% | |
| | Total Number of Tests | 16 | 10 | 12 | 5 | 14 | 10 | 13 | 8 | 13 | 10 | 11 | 8 | 43 |
| | Average Rating (Value) | 2.5 | 3.6 | 3.7 | 1.2 | 1.4 | 2.2 | 1.6 | 3.9 | 3.7 | 2.9 | 3.5 | 3.9 | |
| | Deviation | 1.09 | 0.80 | 0.85 | 0.40 | 0.80 | 1.34 | 1.02 | 0.33 | 0.61 | 0.70 | 0.50 | 0.33 | |
| | Correct Event (Excl. Clears) | | | 84% | | | 21% | | | 71% | | | 100% | 40% |
| | Correct Event and Severity | | | 34% | | | 17% | | | 35% | | | 55% | |

**Figure 38: Rater 5 Test Detailed Event Results**

# APPENDIX 2

## WORKSHOP TRAINING MATERIAL
## TRAINING OUTLINE

## Training Outline

- The trainees will practice maneuvers and procedures.The trick car to demonstrate hesitation, stumble, and surge
- A single trainee will operate the vehicle at a time
- A SwRI operator will be seated in the vehicle to trigger driveability events
- An experienced rater will be seated in the vehicle to give instruction and direction
- Once the trainees have experienced all the malfunctions and severity levels, the SwRI operator will string together several malfunctions
- The trainees will then be "tested" on their ability to identify the malfunctions and discern the severity level



## Training Outline

1. Rater outline and discussion of approach to rating vehicles
2. Expose student to most obvious malfunctions to allow understanding of, and differentiation between malfunctions
3. Allow student to experience all malfunctions at different severities
4. Have student follow the CRC Driveability Data Sheet while exposing to different maneuvers at consistent severity levels
5. Randomized events, randomized severities, while following the CRC Driveability Data Sheet
6. Test student abilities to differentiate maneuvers and severities using the Driveability Data Sheet

# TRAINING AGENDA

Driveability Workshop Agenda

| Segment | Title | Topic | Notes |
|---|---|---|---|
| Hour One | Introduction/ How it's done | Introduce workshop approach. Discuss driving procedure, malfunctions, severity levels | Establish expectations and what the approach will be. Discuss general information about how the procedure is implemented along with Bruce's experiences with it. Review the entire E-28-94 Procedure. |
| Hour Two | Driving the Cycle | Determine throttle positions and drive the cycle | Seat time! Bruce demonstrates how to determine throttle positions and run the cycle. Then the trainee determines throttle positions and drives the cycle. |
| Hour Three | Driveability Malfunction Introduction | Use trick car to produce driveability malfunctions and severity levels | Trainee drives vehicle. Each type of malfunction is created at a moderate level on each type of maneuver. Then all three severity levels of each type of malfunction are demonstrated for each maneuver. |
| Hour Four | Driving the Cycle | Use trick car to produce randomized driveability malfunctions and severity levels | Trainee drives vehicle through entire cycle. Randomized driveability faults and severity levels are introduced. Trainee receives coaching where needed. Repeat as time allows. |
| Hour Five | Session Two Start | Recap first session, Q&A, discuss typical/historical programs | Discussion on whatever needs more work. Answer any questions. Bruce will give a brief overview on how procedure has been used and past programs. |
| Hour Six | Driveability Malfunction Practice | Use trick car to produce randomized driveability malfunctions and severity levels | Seat time! Trainee drives the cycle with randomized malfunctions and severity levels. |
| Hour Seven | Driveability Malfunction Practice | Use trick car to produce randomized driveability malfunctions and severity levels | Results versus the standard are discussed and areas for focus are determined and practiced. |
| Hour Eight | Final Assessment | Use trick car to produce randomized driveability malfunctions and severity levels | Trainee drives vehicle through entire cycle multiple times. Randomized driveability faults and severity levels are introduced. Results vs. actual are assessed. Discuss final scores and final questions/concerns. |

Training Outline r1.xlsxbeh9/19/2019

# CRC DRIVEABILITY PROCEDURE

## REVISED CRC COLD-START AND WARMUP DRIVEABILITY PROCEDURE E-28-94

A. Record all necessary test information at the top of the data sheet.

B. Turn key on for 2 seconds before cranking to pressurize fuel system. Make sure defrost is on and fan is in "low" position. Start engine per Owner's Manual Procedure. Record start time.

C. There may be a total of three starting attempts recorded. If the engine fails to start within 5 seconds on any of these attempts, stop cranking at 5 seconds and record "NS" (no start) in the appropriate starting time box on the data sheet. After the first and second unsuccessful attempts to start, turn the key to the "off" position before attempting to restart per the Owners Manual procedure. If the engine fails to start after 5 seconds during the third attempt, record an "NS" in the Restart2 box, then start the engine any way possible and proceed as quickly as possible to Step D without recording any further start times.

Once the engine starts on any of the first three attempts, idle in park for 5 seconds and record the idle quality. If the engine stalls during this 5-second idle, record a stall in the Idle Park "Stls" box, then restart per the above paragraph, subject to a combined maximum (in any order) of three no-starts and Idle Park stalls. After all the start-time boxes are filled, no further starts should be recorded.

D. Apply brakes (right foot), shift to "Drive" ("Overdrive" if available) for 5-second idle, and record idle quality. If engine stalls, restart immediately. <u>Do not record restart time.</u> Record number of stalls.

A maximum of three Idle Drive stalls may be recorded; however, only one stall contributes to demerits. If the engine stalls a fourth time, restart and proceed to the next maneuver as quickly as possible. It is important to complete the start-up procedure as quickly as possible to prevent undue warmup before the driving maneuvers and to maintain vehicle spacing on the test track.

E. After idling 5 seconds (Step D), make a brief 0-15 mph light-throttle acceleration. Light-throttle accelerations will be made at a constant throttle opening beginning at a predetermined manifold vacuum. This and all subsequent accelerations throughout the procedure should be "snap" maneuvers: the throttle should be depressed immediately to the position that achieves the pre-set manifold vacuum, rather than easing into the acceleration. Once the throttle is depressed, no adjustment should be made, even if the pre-set vacuum is not achieved. Use moderate braking to stop. Idle for approximately 3 seconds without rating it. Make a brief 0-15 mph light-throttle acceleration. Both accelerations together should be made within 0.1-mile. If both accelerations are completed before the

0.1-mile marker, cruise at 15 mph to the 0.1-mile marker. Use moderate braking to stop; idle for approximately 3 seconds without rating it.

F. Make a 0-20 mph wide-open-throttle (WOT) acceleration beginning at the 0.1-mile marker. Use moderate braking to achieve 10 mph and hold 10 mph until the 0.2-mile marker (approximately 5 seconds). Use moderate braking to stop; idle for approximately 3 seconds without rating it.

G. At the 0.2-mile marker, make a brief 0-15 mph light-throttle acceleration. Use moderate braking to stop. Idle for approximately 3 seconds without rating it. Make a brief 0-15 mph light-throttle acceleration. If accelerations are completed before the 0-3-mile marker, cruise at 10 mph to the 0.3-mile marker.

H. At the 0.3-mile marker, make a light-throttle acceleration from 10-20 mph. Use moderate braking to make a complete stop at the 0.4-mile marker in anticipation of the next maneuver. Idle for approximately 3 seconds at the 0.4-mile marker without rating the idle.

I. Make a 0-20 mph moderate acceleration beginning at the 0.4-mile marker.

J. At the 0.5-mile marker, brake moderately and pull to the right side of the roadway. Idle in "Drive" for 5 seconds and record idle quality. <u>Slowly</u> make a U-turn.

K. Repeat Steps E through J. At the 0.0-mile marker, brake moderately and <u>slowly</u> make a U-turn.

**NOTE: Items L-N may be useful only at colder temperatures.**

L. Make a crowd acceleration (constant predetermined vacuum) from 0-45 mph. Four-tenths of a mile is provided for this maneuver. Decelerate from 45 to 25 mph before the 0.4-mile marker.

M. At the 0.4-mile marker, make a 25-35 mph detent position acceleration.

N. At the 0.5-mile marker, brake moderately. Idle for 30 seconds in "Drive," recording idle quality after 5 seconds and after 30 seconds, and record any stalls that occur. This ends the driving schedule. Proceed to the staging area.

Definitions of light-throttle, detent, and WOT accelerations are attached. During the above maneuvers, observe and record the severity of any of the following malfunctions (see attached definitions):

1. Hesitation
2. Stumble
3. Surge
4. Stall
5. Backfire

It is possible that during a maneuver, more than one malfunction may occur. Record all deficiencies observed. Do not record the number of occurrences. If no malfunctions occur during a maneuver, draw a horizontal line through all boxes for that maneuver. Also, in recording subjective ratings (T, M, H or E), be sure the entry is legible. At times, M and H recordings cannot be distinguished from each other.

Record maneuvering stalls on the data sheet in the appropriate column: accelerating or decelerating. If the vehicle should stall before completing the maneuver, record the stall and restart the car as quickly as possible. Bring the vehicle up to the intended final speed of the maneuver. Any additional stalls observed will not add to the demerit total for the maneuver, and it is important to maintain the driving schedule as closely as possible.

## DEFINITIONS AND EXPLANATIONS

Test Run

Operation of a car throughout the prescribed sequence of operating conditions and/or maneuvers for a single test fuel.

Maneuver

A specified single vehicle operation or change of operating conditions (such as idle, acceleration, or cruise) that constitutes one segment of the driveability driving schedule.

Cruise

Operation at a prescribed constant vehicle speed with a fixed throttle position on a level road.

Wide Open Throttle (WOT) Acceleration

"Floorboard" acceleration through the gears from prescribed starting speed. Rate at which throttle is depressed is to be as fast as possible without producing tire squeal or appreciable slippage.

Part-Throttle (PT) Acceleration

An acceleration made at any defined throttle position, or consistent change in throttle position, less than WOT. Several PT accelerations are used. They are:

1. Light Throttle (Lt. Th) - All light-throttle accelerations are begun by opening the throttle to an initial manifold vacuum and maintaining *constant throttle position* throughout the remainder of the acceleration. The vacuum selected is the vacuum setting necessary to reach 25 mph in 9 seconds. The vacuum setting should be determined when the vehicle is cold. The vacuum setting is posted in each vehicle.

2. Moderate Throttle (Md. Th) - Moderate-throttle accelerations are begun by immediately depressing the throttle to the position that gives the pre-specified vacuum and maintaining a *constant throttle position* throughout the acceleration. The moderate-throttle vacuum setting is determined by taking the mean of the vacuum observed during WOT acceleration and the vacuum prescribed for light-throttle acceleration. This setting is to be posted in the vehicle.

3. Crowd - An acceleration made at a constant intake manifold vacuum. To maintain *constant vacuum*, the throttle-opening must be continually increased with increasing engine speed. Crowd accelerations are performed at the same vacuum prescribed for the light-throttle acceleration.

4. Detent - All detent accelerations are begun by opening the throttle to just above the downshift position as indicated by transmission shift characteristic curves. Manifold vacuum corresponding to this point at 25 mph is posted in each vehicle. *Constant throttle position* is maintained to 35 mph in this maneuver.

Malfunctions

1. Stall

   Any occasion during a test when the engine stops with the ignition on. Three types of stall, indicated by location on the data sheet, are:

   a. Stall; idle - Any stall experienced when the vehicle is not in motion, or when a maneuver is not being attempted.

   b. Stall; maneuvering - Any stall which occurs during a prescribed maneuver or attempt to maneuver.

   c. Stall; decelerating - Any stall which occurs while decelerating between maneuvers.

2. Idle Roughness

   An evaluation of the idle quality or degree of smoothness while the engine is idling. Idle quality may be rated using any means available to the lay customer. The rating should be determined by the worst idle quality experienced during the idle period.

3. Backfire

   An explosion in the induction or exhaust system.

4. Hesitation

   A temporary lack of vehicle response to opening of the throttle.

5. Stumble

   A short, sharp reduction in acceleration after the vehicle is in motion.

6. Surge

   Cyclic power fluctuations.

Malfunction Severity Ratings

The number of stalls encountered during any maneuver are to be listed in the appropriate data sheet column. Each of the other malfunctions must be rated by severity and the letter designation entered on the data sheet. The following definitions of severity are to be applied in making such ratings.

1. Trace (T) - A level of malfunction severity that is just discernible to a test driver but not to most laymen.

2. Moderate (M) - A level of malfunction severity that is probably noticeable to the average laymen.

3. Heavy (H) - A level of malfunction severity that is pronounced and obvious to both test driver and layman.

4. Extreme (E) - A level of malfunction severity more severe than "Heavy" at which the lay driver would not have continued the maneuver, but taken some other action.

Enter a T, M, H, or E in the appropriate data block to indicate both the occurrence of the malfunction and its severity. More than one type of malfunction may be recorded on each line. If no malfunctions occur, enter a dash (-) to indicated that the maneuver was performed and operation was satisfactory during the maneuver.

## DEMERIT CALCULATION SYSTEM

A numerical value for driveability during the CRC test is obtained by assigning demerits to operating malfunctions as shown. Depending upon the type of malfunction, demerits are assigned in various ways. Demerits for poor starting are obtained by subtracting one second from the measured starting time and multiplying by 5. The number of stalls which occur during idle as well as during driving maneuvers are counted separately and assigned demerits as shown. The multiplying x factors of 7, 28, 50, and 100 for idle and maneuvering stalls account for the fact that stalls are very undesirable, especially during car maneuvers. A maximum of three total Idle Park stalls and No-Starts are permitted.

Other malfunctions, such as hesitation, stumble, surge, idle roughness, and backfire, are rated subjectively by the driver on a scale of trace, moderate, heavy, or extreme. For these malfunctions, a certain number of demerits is assigned to each of the subjective ratings. Since all malfunctions are not of equal importance, however, the demerits are multiplied by the weighting factors shown to yield weighted demerits.

Finally, weighted demerits, demerits for stalls, and demerits for poor starting are summed to obtain total weighted demerits (TWD), which are used as an indication of driveability during the test. As driveability deteriorates, TWD increases.

A restriction is applied in the totaling of demerits to insure that a stall results in the highest possible number of demerits within a given maneuver. When more than one malfunction occurs during a maneuver, demerits are counted for only the malfunction which had the largest number of weighted demerits.

**When all the factors are multiplied together the following chart of demerit levels is generated.**

**Demerit levels for: Hesitation/Stumble/Surge/Backfire/Stall**

| Maneuver | Stall | Extreme Clear | Heavy | Medium | Trace | |
|---|---|---|---|---|---|---|
| Light Throttle | 50 | 16 | 8 | 4 | 2 | 0 |
| Medium Throttle | 100 | 32 | 16 | 8 | 4 | 0 |
| WOT | 100 | 32 | 16 | 8 | 4 | 0 |
| Detent | 50 | 16 | 8 | 4 | 2 | 0 |
| Crowd | 50 | 16 | 8 | 4 | 2 | 0 |

**For Idle Roughness**

| Extreme | Heavy | Medium | Trace | Clear |
|---|---|---|---|---|
| 8 | 4 | 2 | 1 | 0 |

**For Idle Stalls**

Idle-in- Park | Starting-in-Drive | Other Idle (after moderate throttle or at end of test)

7 each | 28 | 7

For Starting

**No Start | Slow Start|**

25 each | t-1*5

The Start time, t, is in seconds.

Only the results (start, start + stall, no-start) of the first three starting attempts in park count toward demerits.

Only the first stall in drive prior to maneuvering counts toward demerits

Only the first stall in each maneuver, or in each idle subsequent to the start of the maneuver is counted toward demerits.

Only the highest weighted demerit score from each maneuver is counted.

# DRIVEABILITY QUICK REFERENCE

**Coordinating Research Council (CRC) Vehicle Driveability Rating Workshop**
**Uvalde, Texas September, 2019**
**Quick Reference to be used with CRC Procedure E-28-94**

## Maneuver Throttle Positions

(note: the term throttle is used as an equivalent to accelerator pedal)

**Light Throttle Acceleration** - In a single movement, snap and hold the throttle position required to accelerate from 0 to 25mph in nine seconds with the engine soaked to test temperature.

**Wide Open Throttle (WOT Acceleration)** - Accelerator pedal is pushed rapidly to the floor as quickly as possible without producing tire slippage (maximum obtainable throttle).

**Moderate Throttle Acceleration** - Snap and hold the throttle at the midway point between light throttle and WOT.

**Crowd Acceleration** - Performed at a constant manifold vacuum pressure. These accelerations are achieved using the same vacuum prescribed for the vehicle's light throttle acceleration. To maintain constant vacuum, the throttle opening may need to be adjusted with increasing engine speed and automatic transmission shifts/lockups.

**Detent Acceleration** - Constant throttle position and associated initial manifold vacuum setting that is the maximum obtainable without downshift from the prescribed road speed with a warm to hot engine. Constant throttle position is maintained to the prescribed 35mph speed.

**Bunny Hops** - WOT accelerations and hard braking decelerations to/from a prescribed speed. Bunny Hops are used to heat up the engine during Hot Fuel Handling (Vapor Lock) testing. Typically, ten Bunny Hops to 35mph are performed before parking the vehicle in a soak shed.

**Coordinating Research Council (CRC) Vehicle Driveability Rating Workshop**
**Uvalde, Texas September, 2019**
**Quick Reference to be used with CRC Procedure E-28-94**

## Malfunctions

**Stall -** Engine ceases to run during park idle, drive idle, acceleration, or deceleration.

**Hesitation** - A temporary lack of power immediately upon throttle opening that is not characteristic of the vehicle's warmed-up performance.

**Stumble** - A temporary or continuous reduction in power after the throttle opening.

**Surge** - A cyclical decrease and increase in power and engine speed.

**Backfire** - An explosion in either the intake or exhaust system.

**No Start (NS)** - The engine fails to start in five seconds of cranking the starter.

**Coordinating Research Council (CRC) Vehicle Driveability Rating Workshop**
**Uvalde, Texas September, 2019**
**Quick Reference to be used with CRC Procedure E-28-94**

## Severity Levels

**Clean or Clear (horizontal line through the evaluation area of data sheet)** - Clean has separate definitions for idles and maneuver ratings.
**Clean Idle Roughness** - At worst, the trained rater can barely feel or hear the engine running.
**Clean Maneuver** - No malfunctions detected during maneuver.

**Trace (T) Idle Roughness or Maneuver Malfunction** - A level of malfunction just discernable to a trained rater and unlikely to be detectable by an average driver.

**Moderate (M) Idle Roughness or Maneuver Malfunction** - A level of malfunction noticeable to an average driver.

**Heavy (H) Idle Roughness or Maneuver Malfunction** - A level of malfunction pronounced and obvious to any driver.

**Extreme (E)** - Extreme has separate definitions for idles and maneuver ratings.
**Extreme/Idle** - Severe operational abnormalities where the engine is threatening to stall and/or the malfunction continues, reoccurs or is excessive for an extended period of the prescribed idle duration.
**Extreme/Maneuver** - Severe operational abnormalities at which the average driver would not complete the maneuver but would take some other action.

## Q&A FROM EXPERIENCED RATER

- Describe your overall approach to rating driveability events in the test vehicles.
    - Use senses to feel for engine roughness, typically through the steering wheel and seat. Listen for backfires. Pay attention to how the vehicle responds to accelerator pedal movements.
- Discuss preparing to execute the test procedure
    - What preparation steps are important for accurately identifying and recording driveability events?
        - No distractions to rater (no radio, no talking, clear windshield, no non-test vehicles using track, dry pavement etc.)
    - What do you do before you start driving the vehicle?
        - Record mileage, test temp, ready stopwatch, clear windshield, make sure track is clear of vehicles, etc..
        - There is another aspect to "preparation" prior to the actual rating. This includes familiarizing oneself with the vehicle, fuel handling, pre-conditioning, staging, etc. Although this isn't the focus of workshop it should get a short, 5-10 minute, discussion at the training.
- Discuss your experiences driving vehicles through the test procedure.
    - I've done many ratings done on a wide variety of vehicles, also training and oversight of other raters and participation in CRC programs. Driveability ratings are useful to determine performance changes due to fuel composition such as volatility and/or components such as oxygenates, and also engine cleanliness such as IVD and/or PFI deposits.
    - When driving the vehicle, are there any driveability events that are more difficult to identify than others?
        - Stalls are obvious and backfires don't require much judgement. Hesitations occur at the beginning of a maneuver. After the vehicle is launched, hesitations aren't a factor, but surges and stumbles are. Surges are cyclic and stumbles are a temporary reduction in power. I've seen people get hung up on distinguishing between the two but there's no reason they should; surges a cyclical, like several stumbles one right after another. In the past it didn't really matter if a rater called it a surge or a stumble because they both generated the same demerits but it's best to get it correct because one never knows how the data will be analyzed.
    - Have you ever had challenges discern between two events, or been faced with a scenario where more than one event occurred at the same time?

        - This happens frequently in some programs using bad fuels...just pay attention to the most severe malfunction which will be recorded.
    - What do you do when the vehicle stalls?
        - Key off and re-time idle stalls only (two plus initial). Depends where in the maneuver procedure it stalls...does it restart easily and run fine or continue to stall? This can be a tough question but fortunately hardly happens. The best advice is to keep it moving through the maneuvers the best you can which sometimes takes decision making because you need to keep the engine temperature in mind that is fair to the test run. Drive restarts are never timed/recorded or are idle stalls once the maneuvers have started.
    - What do you do when a driveability event is so severe that you need to shut off the vehicle (from risk of damage to the vehicle)?
        - Keep in mind you can destroy the catalyst if it loads up with fuel, lights up and melts...plus it is a safety issue, vehicle components outside of the catalyst can burn too like undercoating and carpeting inside the vehicle (it happened at the Brainerd program). Shut down engine if you think the malfunction may be harmful.
- Discuss recording and clarifying events in the Driveability Data Sheet.
    - What is your process for filling in the driveability rating sheet?
        - A novice rater should use an observer so the rater can concentrate on the vehicle performance and not run into anything. A seasoned rater can both rate and record malfunctions, but it is not recommended for best testing results on a test track; it can be done on a dyno without too much trouble. I've only seen data recorded via a data sheet/clipboard. I think it would be very challenging/dangerous for the rater to use a computer to record observations as they operate the vehicle on the road or track. For the rater to use a computer directly while on a dyno they'd have to be very pc-nimble and have a bulletproof program, but it seems there is a risk of losing a test. A clean running vehicle doesn't have many malfunctions whereas a poor fuel can cause multiple entries in the data sheet which require more attention.
    - If you have a question regarding an event you experienced in the test procedure, how did you seek clarification and / or get advice on quantifying the event?
        - Consult the CRC E-28-94 driveability procedure or ask an experienced rater. Don't worry about demerits while you are performing the procedure.
- Share your experience with the Trainees.
    - Describe your most challenging experience in a Driveability Workshop?
        - I have a funny story about a cold driveability in Brainerd. The most challenging tests are when there are 5-second No-starts and retiming stalls

during initial idles with key-offs.  Also, multiple stalls that happen on opening of the throttle during maneuvers and what you do to keep moving throughout the procedure without heating up the engine abnormally.

- What questions should a Trained Rater be asking Driveability Workshop Panel before the test?  -- during the testing?
  - The raters need to know if there are any changes to the driving cycle. The logistics are generally site specific.
- In your experience, when would you expect to see a new person display adequate competency as a trained rater?
  - Often, experienced raters and trainers can tell within a day if a person has the potential to become a good rater. Some people catch on quickly to rating, others never will. The more seat time and more varied driveability experiences a rater witnesses, the better they get and it becomes more natural moving through the procedure. It's a difficult balance; programs are costly yet raters need to gain experience somehow. I might be willing to begin trusting a rater if they showed competence and confidence in several dozen ratings but if I'm being truthful I'd have the most confidence in a rater that has done hundreds of ratings and the results line up with what's expected by either engine cleanliness, fuel composition, or trick settings.
- Any words of wisdom that you'd like to share with the Trainees?
  - Don't be offended if someone offers suggestions to help you learn.  Sometimes even the trainer learns new things.  After the crowd maneuver, don't wait until the last second to get down to 25 mph before performing the detent maneuver, get down to 25 mph soon after the crowd to let the transmission find it's happy gear.  Remember, the CRC procedure was written when vehicle automatic transmissions didn't have so many gears and combinations of lock-up gears, it may be impossible to perform a detent maneuver without gear shifts.

# DRIVEABILITY WORKSHEET

***Driveability Data Sheet - CRC E-28-94***

Run History

| Run No. | Car | Fuel | Rater | Date | Time | Temperatures Soak | Temperatures Run | Odometer | T.W. Demerits |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

Starting Time, Sec

| Initial | Restart 1 | Restart 2 | Idle Park Ruf Stls | Idle Drive Ruf Stls | Overall |
|---|---|---|---|---|---|
| | | | | | |

| 0.0 0-15 LT TH | 0-15 LT TH | 0.1 0-20 WOT | 0.2 0-15 LT TH | 0-15 LT TH | 0.3 10-20 LT TH | 0.4 0-20 MD TH |
|---|---|---|---|---|---|---|
| H S B | H S B | H S B | H S B | H S B | H S B | H S B |
| E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D |
| S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C |

0.5 Idle Dr.

Ruf Stls

| 0.5 0-15 LT TH | 0-15 LT TH | 0.6 0-20 WOT | 0.7 0-15 LT TH | 0-15 LT TH | 0.8 10-20 LT TH | 0.9 0-20 MD TH |
|---|---|---|---|---|---|---|
| H S B | H S B | H S B | H S B | H S B | H S B | H S B |
| E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D | E T S K A D |
| S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C | S M G F C C |

| 1.0 Idle Dr. | 1.0 0-45 Crowd | 1.4 25-35 Detent | 1.5 Idle Dr. | 1.5 Idle Dr. |
|---|---|---|---|---|
| | H S B | H S B | 5 Sec. | 30 Sec. |
| | E T S K A D | E T S K A D | | |
| Ruf Stls | S M G F C C | S M G F C C | Ruf Stls | Ruf Stls |

Comments:

# APPENDIX 3

# RATER SURVEY

| Topic/Question | Question Specifics | Rater Responses | | | | |
|---|---|---|---|---|---|---|
| | | Rater 1 | Rater 2 | Rater 3 | Rater 4 | Rater 5 |
| **1. Can you describe the purpose of CRC driveability programs** | **a. Why do we conduct them?** | We conduct CRC ratings for clients to determine driveability differences in a vehicle. This can be done to help clients test how products affect a vehicles performance. | Purpose was to test the different effects that a fuel can have on a vehicle. Through testing vehicles with consistent temperatures and maneuvers, a Rater can Identify any malfunctions these fuels might | Determine the performance of new fuels and additives for the average driver. Also determine vehicle performance for the average driver. | To assess fuel performance issues, not mechanical issues. | To determine the fuel effects on volatility classes, and vehicle performance to those fuels. |
| | **b. What is the role of the Rater?** | A rater gives a consistent unbiased summary of a vehicles performance during the CRC driveability runs. | To be consistent in performing maneuvers while using good judgement to identify malfunctions. | Determine vehicle performance for the average driver while using different fuels or additives. | To identify malfunctions that are fuel related and determine their severity level. | To run the test cycle, which contains set maneuvers. The rater needs to provide level of severity (clear, trace, moderate, …) of driveability events (hesitations, stumble, surge, …). |
| | **c. How does a CRC driveability program result in improvements to fuels?** | CRC driveability program results in improvements to fuels because it helps clients determine how a fuel is affecting a vehicle and they can make the appropriate changes if needed. | With the different maneuvers it simulates certain driving routines, so companies can know which fuels will perform where with best efficiency. | Creates a standardized approach to measure the performance of fuels and provide feedback. | The formal objective of CRC is to encourage and promote the arts and sciences by directing scientific cooperative research to develop the best possible combinations of fuels and lubricants. | A CRC program only provides data that could be used by others (e.g., ASTM) to make changes to fuel specifications. |
| **2. Did the classroom training help you understand the meaning of the terms below, and how to recognize them during vehicle operation?** | **a. Hesitation** | Agree | Strongly Agree | Strongly Agree | Strongly Agree | Strongly Agree |
| | **b. Stumble** | Agree | Strongly Agree | Strongly Agree | Strongly Agree | Strongly Agree |
| | **c. Surge** | Agree | Agree | Strongly Agree | Strongly Agree | Strongly Agree |
| | **d. Stall** | Agree | Strongly Agree | Strongly Agree | Strongly Agree | Agree |
| | **e. Idle quality** | Strongly Agree | Strongly Agree | Strongly Agree | Strongly Agree | Disagree |

| | | | | | |
|---|---|---|---|---|---|
| **3. After the classroom training, can you describe the differences in severity as Heavy, Moderate, or Trace for each of the driveability events?** | The class did not give me enough time to become comfortable describing the differences in all the driveability events. A trace surge felt light a stumble at times and overall I feel that more that a few hours of driving the trick car is needed to be consistent. | I felt I was doing well towards the end, but due to technical issues it was hard to sense the trace malfunctions. | Trace: Very minor, the average driver may not detect it. Moderate: Average driver may notice it, take note of it, but continue driving and schedule to get car in. Heavy: Average driver would notice it, would take car to nearest shop to have it looked at. | Yes | There was an understanding of the differences. - trace, moderate, and heavy were low, medium, and high intensities, respectively. However, the classroom portion was not enough to determine the difference during the test cycle. |
| **4. Did the instructor provide you with a description of the course, and how you will be expected to operate the vehicle during the different test cycles?** | The instructor did provide with a description of the course and did describe how I should operate the vehicle during the different test cycles. | We were shown a data sheet of whats to be expected, but it would of been nice to see the experienced rater perform more, that way we can get a better understanding of the process as a whole. | Yes. Bruce was very thorough on making sure I understood what was expected. | Yes | A verbal description was provided, but not the flow of the maneuvers. A diagram with the maneuvers, including indication where to start/end would have been useful. |
| **5. Did the instructor provide you with information on what they will be doing during the test cycles?** | The instructor did say what he would be doing during the test. | Yes, we went through each maneuver. | Yes. | Yes | This was not provided in the classroom portion. It was provided once in the vehicle an on the track when the instructor (experienced rater) went through the test cycle. |
| **6. Do you know what to do if there is a safety issue at the track or while you are driving the vehicle?** | Depends on severity of the safety issue. | Besides not having a seat belt, I felt comfortable enough to know the proper safety precautions would be implemented if needed. | Yes. | Yes | SwRI explained the 'kill' switch, and the need to evacuate the vehicle if there was an issue |

| | | | | | |
|---|---|---|---|---|---|
| **1. Did you feel comfortable with the vehicle seating position, vehicle cabin temperature, and configuration?** | Yes | Yes | Yes | Yes | Yes |
| **Was there anything that made you uncomfortable in the vehicle? (Note that the HVAC needs to be set at a specific temperature for the idle evaluation.)** | No. HVAC settings were not changed for comparison. | Just the no seat belt was weird, but once I adjusted my seat I was fine. | No. | No | Due to the high temperatures (100°F), adjustments were made for the comfort of all occupants. |
| **2. Did you achieve a safe driving position, mirror adjustment, familiarity of the controls, etc. before driving the vehicle?** | Yes | Yes | Yes | Yes | Yes |
| **Did anything prevent you from feeling safe in the vehicle?** | No. | No Seat Belt, Cones could have also been managed better. | No. | No | The only issue was that the seatbelts were not being used. Since on a track with no other vehicles, this was not a major issue. |
| **3. How did the vehicle function in normal drive mode? Did it feel normal or modified?** | Vehicle felt normal. | Felt Fine. | Felt normal. | Normal | The vehicle did perform as normal. Although some hardware made you aware that changes were made, none effected the driving mode. |
| **4. What was the condition of the track?** | Good. | Track was good and dry when I drove, Cone were unorganized but we managed. | Dry and clear. | Great | This track was acceptable for the test cycle and the training. However, it had cones for two programs (CRC & Continental), which was confusing at first. |
| **5. Were the courses well laid out?** | Yes | No | Yes | Yes | Yes |
| **6. Were their any distractions or conditions which made driving on the course difficult for you?** | No | No | No | No | extra cones |
| **7. During the in-vehicle training, did the instructor help you with identifying each driveability event?** | Yes | Other | Yes | Yes | Yes |

| | | | | | | |
|---|---|---|---|---|---|---|
| **8. During the in-vehicle training, did the instructor help you with identifying severity for each driveability event?** | | Yes | Yes | Yes | Yes | Yes |
| **a. Could you differentiate between the calibration of heavy, moderate, and trace severities?** | | Some better that others. More hands on training required. | Yes | Yes | Most of the time. | On some maneuvers, but not well on hesitation on WOT |
| **b. Did the different severities feel appropriate? i.e. heavy felt like a heavy severity.** | | A trace surge felt like a stumble. Over all I believe more hands on training is required to be accurage and consistent. | For the most part, except when trace didn't work properly. | Most of the time. James mentioned that there were some events that didn't trigger as heavy. | Most of the time | Yes |

| | | | | | | |
|---|---|---|---|---|---|---|
| **9. Did you experience any difficulty identifying moderate or trace severity for driveability events? If so, which ones?** | | Yes. Some where a trace surge, a trace stumble, a hesitation when comparing WOP to a light or moderate throttle. | I was more confused on Moderate to Heavy. But I it's just the lack of practice. | Some moderates felt heavy. | Yes, some trace, and moderate stumbles felt like surges, and some trace, and moderate surges felt like stumbles. Though, some of that could have been in my throttle position setting. | I had issue with identifying WOT hesitation, & detent (any). When asked about the crowd, it was some stated that the heat was having some impact. All in all, I think additional time in the vehicles would allow better identification. Once the test cycle is understood, repeats of maneuvers would be beneficial, especially if they could be done in a row. |
| **10. Overall, do you feel that you would have benefitted from more time for in vehicle training before the test?** | | Yes | Yes | Yes | Yes | Yes |
| **11. Did anything cause your training to be disrupted?** | | No | No | No | No | No |
| **1. Did you establish a safe and comfortable driving position and access to the vehicle controls before starting the test?** | | Yes | Yes | Yes | Yes | Yes |

| | | | | | | |
|---|---|---|---|---|---|---|
| **How confident are you that you correctly identified the driving events listed below?** | **a. Hesitation** | Very Confident | Not Confident | Very Confident | Very Confident | Somewhat Confident |
| | **b. Stumble** | Very Confident | Not Confident | Somewhat Confident | Somewhat Confident | Very Confident |
| | **c. Surge** | Somewhat Confident | Not Confident | Very Confident | Very Confident | Very Confident |
| | **d. Idle quality** | Very Confident | Not Confident | Very Confident | Very Confident | Very Confident |
| **3. How confident are you that you correctly identified the severity of the driving events listed below?** | **a. Hesitation** | Somewhat Confident | Not Confident | Very Confident | Very Confident | Not Confident |
| | **b. Stumble** | Somewhat Confident | Somewhat Confident | Somewhat Confident | Somewhat Confident | Somewhat Confident |
| | **c. Surge** | Somewhat Confident | Not Confident | Very Confident | Somewhat Confident | Very Confident |
| | **d. Idle quality** | Very Confident | Somewhat Confident | Very Confident | Very Confident | Very Confident |

| | | | | | |
|---|---|---|---|---|---|
| **4. Did anything during the test make you feel nervous or unsafe?** | The fact that I only had one hour of driving on day one and a few hours on day two. | No seat belts, car stalled at times | No. | No | During the evaluation (6 cycles in a row) I was a little nervous, since I was not confident of the event/severity in each of the maneuvers, especially during the detent. |
| **5. Did you have to abort the test? If so, what was the cause?** | No. | Technical issues here and there but they got it settled. Car would stall on re-ignition | No. | No | I do not remember if any of the 6 cycles were aborted in the final test. However, I think a couple of the maneuvers were not executed correctly. |
| **1. What was good about the training?** | I learned something. I am not as confident as I would like to be but the course did help. | Was a good overall experience, brought awareness to these types of tests. Felt both people I trained with knew how to handle any situation. | I appreciated the depth of knowledge from Bruce, James and Stas. There wasn't a question I had that went unanswered. | Bruce was great at providing his experiences and recommendations on his involvement in previous CRC driveability programs. In which made myself very confident in my abilities to perform these programs as a trained driveability rater. | The one-on-one time was good. Since I was a novice, some background on the test cycle would have helped especially for the crown and detent. The trick car performed as well as it could for the ambient temperature that week. The experience trainer, and the SwRI staff were all helpful. |
| **2. What could use improvement?** | More time driving and more that two half days training. | A better course, then just more explanation on what will actually be used for the workshop, and what information is more for the actual testing. Just felt I was given a lot of information to process, and I kept trying to guess when I would need it in the course. | Some more time in the car would be good, as long as it isn't a full day. I think that would be too long. there's a benefit to getting out of the car and coming back to it later. A one day follow-up check might be beneficial. | Maybe a little more seat time. | If a small screen could show the speed & throttle position to the driver, the training may move faster. Since the one needs to learn the difference with a light and medium throttle. The detent was call ‘archaic’ since most vehicles have higher transmission gears now. So this maneuver needs to be investigated if there is still value, or if it should be removed from the cycle. |
| **3. If you were in a CRC driveability Program, could you provide an accurate evaluation of driveability events?** | No | Yes | Yes | Yes | Yes |
| **4. Would you recommend this training program to others?** | No | Yes | Yes | Yes | Yes |

# APPENDIX 4

# EXPERT RATER FEEDBACK

Bruce Henderson
Bruce.Henderson@bp.com
630-688-1127

October 7, 2019

Stas Gankov, Research Engineer
Southwest Research Institute
210-522-6206 | sgankov@swri.org

Dear Stas:

Thank you for enlisting me to serve as the trainer for the CRC driveability workshop held in Uvalde, TX on September 23-27, 2019. I believe the workshop was successful in teaching five new driveability raters the procedure to a level they can build on in order to become proficient in its application. Please allow me to provide the following additional feedback:

Each of the five students received two, half-days, of individualized training comprised of a "classroom" explanation of the driving technique, malfunctions, and severity levels, then a demonstration of the technique and malfunctions, followed by driving the technique and sensing the malfunctions repeatedly with critique and coaching throughout. Working with each student individually allowed the training pace and focus areas to be tailored to the students' range of vehicle testing experience and their demonstrated level of competency. The students all started with different amounts of vehicle performance appraisal know-how, learned at different rates, and ended with somewhat different levels of mastery. At the conclusion of the students' training, each of them had gained enough skill in the application of the procedure that they could continue to develop as they practice the techniques and apply them to additional vehicles that have varied driveability performance.

The track used at the Continental Proving Grounds, Dry Handling #1, was slightly less than optimum because it had several curves that made the application of the acceleration maneuvers challenging. It was made to work adequately by laying out the mileposts so the straight-away was utilized for the 0-45mph Crowd acceleration and stabilization cruise prior to the detent acceleration. After some acclimation, the track provided a suitable place to conduct the training.

The "trick car" vehicle, while extremely useful in this training, did not perform flawlessly. There were times Trace and Moderate level malfunctions could not be produced repeatedly, particularly for Hesitations. Several theories were advanced on why this occurred including: sensitivity to rate of throttle application; an ambient temperature effect as compared to the initial calibration; and the vehicle had "learned" a revised calibration.

Finally, I'd be remiss if I didn't address the need for updates in the procedure itself in order to improve it for current vehicle technology. As the name indicates, the E-28-94 procedure has been in place for over 25 years. The procedure does a lot of things very well: the maneuvers are sequenced in a way which is easy to implement and doesn't generate too much engine heat too quickly. Unfortunately, the procedure doesn't take in consideration newer vehicle technologies including continuous auto-crank starting, AWD/FWD, selectable transmission modes, traction control, and continuously variable transmissions. The method for establishing throttle positions must also be reviewed. The approach used at this workshop was consistent with the Quick Reference document first produced as a result of the 2002 Workshop and updated prior to this workshop and it was perfectly adequate for the "trick car", a 2014 Ford Fusion.

I appreciate the organizational and logistical work you did to facilitate the success of the workshop. I also want to recognize the significant contribution James Fritz made to the training. His ability to read the throttle position from the control computer and then give feedback to the students was very useful and it was done in a constructive way that helped them progress. I also want to express my appreciation to CRC and the committee members for their interest and encouragement for this workshop.

Please feel free to contact me if I can help in any way.

Best Regards,

Bruce