U.S. Department of Transportation

Federal Railroad Administration

Office of Research, Development and Technology
Washington, DC 20590

Evaluation of Habituation to Alerts in Locomotive Cabs

DOT/FRA/ORD-24/36 Final Report | September 2024

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
9/12/24

2. REPORT TYPE
Technical Report

3. DATES COVERED *(From - To)*
September 2019–May 2022

4. TITLE AND SUBTITLE
Evaluation of Habituation to Alerts in Locomotive Cabs

5a. CONTRACT NUMBER
693JJ6-19-C-000030

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kelly Ozdemir: 0000-0002-3238-4163, Kianna Pirooz: 0000-0001-6414-8336, Maura Campbell: 0000-0001-8272-7251, Timothy Allen: 0000-0002-8362-0023, Matthew Isaacs: 0000-0003-3889-8592, Lafcadio Flint: 0000-0001-9817-9819

5d. PROJECT NUMBER
FR19RPD34000000003

5e. TASK NUMBER
4

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
KEA Technologies Inc.
400 Donald Lynch Blvd., Suite 102
Marlborough, MA 01752

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DOT/FRA/ORD-24/36

12. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA website.

13. SUPPLEMENTARY NOTES
COR: Michael Jones

14. ABSTRACT
From September 2019 to May 2022, the Federal Railroad Administration (FRA) funded KEA Technologies Inc. to conduct a pilot study using a rail simulator to determine whether engineers can become habituated to auditory alerts while operating a locomotive. Habituation, in this case, is a behavior where the engineer becomes less responsive to an alarm or alert because of its highly repeated or too frequent sounding. Because of habituation, the design and presentation of alarms and alerts to the engineer and how he/she responds becomes an important safety issue. The team completed a short series of pilot studies with a limited number of engineer participants using FRA's Cab Technology Integration Lab. However, in the studies, the team did not observe habituation behavior. While conducting the studies in a simulated environment offered more controlled settings, further observations aboard locomotives in actual operational settings is recommended.

15. SUBJECT TERMS
Habituation, inattentional deafness, alarm fatigue, locomotive engineer, Positive Train Control, PTC, alerter, human factors

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		37	19b. TELEPHONE NUMBER *(Include area code)*

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

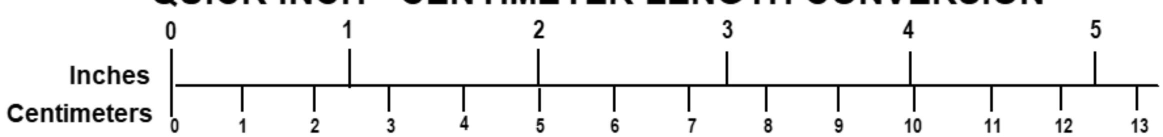
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

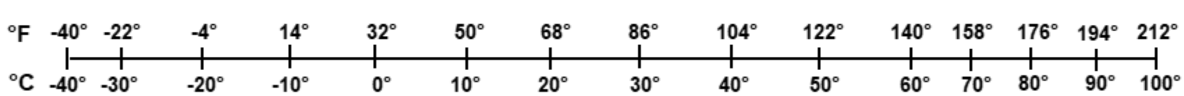
TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION



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

Updated 6/17/98

Acknowledgements

The research team thanks the staff at Volpe National Transportation Systems Center, specifically Matt Isaacs, for guidance, knowledge, and support on this project. The team also thanks all the locomotive engineers who provided their expertise and time during every phase of this study, but most importantly, during the pilot testing process.

Contents

Illustrations

Tables

Executive Summary

From September 2019 to May 2022, the Federal Railroad Administration (FRA) funded KEA Technologies Inc. (KEA) to conduct a pilot study using a rail simulator to determine whether engineers can become habituated to auditory alerts while operating a locomotive. Habituation to warning signals is not a new concept; however, the depth and volume of research is limited (Wogalter & Silver, 1995; Edworthy & Adams, 1996).

Habituation is defined as a reduced attentional response to repeated exposure to a stimulus. Warning signals may initially attract attention, but over time become perceived as less important. Habituation to warnings is problematic for several reasons: 1) the signal loses its alerting capability; 2) people may prematurely dismiss the warning before encoding or interpreting all the necessary information; and 3) subsequent warning signals may elicit a lowered or lesser response (Wogalter & Vigilante, 2006). Researchers conducted a literature review on habituation in other industries such as aviation, automotive, and healthcare, to understand and provide context whether this behavior was prevalent and document lessons learned. The literature review helped inform the design of the pilot study. The team then developed an experimental protocol to potentially observe habituation behavior using FRA's Cab Technology Integration Laboratory (CTIL) simulator and identified variables of interest for the study.

Nine engineers participated in the pilot study in CTIL for up to 6 hours each. Participants completed two surveys, a demographic survey and exit questionnaire.

Researchers analyzed more than 43 hours of video of participants in CTIL and did not observe evidence of habituation behavior. Findings suggested, however, that while habituation was not occurring, fatigue may have influenced results for some engineers due to long sessions in CTIL. Due to significant challenges in observing habituation in a simulated environment, future research should evaluate habituation in the field (i.e., on an actual locomotive).

1. Introduction

From September 2019 to May 2022, the Federal Railroad Administration (FRA) funded KEA Technologies Inc. (KEA) to investigate the potential for habituation to auditory alerts in the locomotive cab.

1.1 Background

In 2009, FRA released a report titled “Technology Implications of a Cognitive Task Analysis for Locomotive Engineers” that listed the cognitive demands and performance issues faced by locomotive engineers. The list included:

- Maintain broad situational awareness and develop an accurate current situation model.
- Generate expectations to think ahead, guide attention, and prepare for anticipated actions as well as plan for contingencies.
- Actively engage in sustained visual and auditory monitoring, including monitoring radio communication.
- Manage multiple demands on attention.
- Prioritize and manage multiple goals.

Given that a locomotive crew monitors the locomotive through visual and auditory means, there is a need to explore and understand the effectiveness of visual and auditory alerts and how locomotive crew respond. Evidence suggests that individuals are likely to ignore alarms when a high false alarm rate exists (Getty et al., 1995). This is more likely to occur when an individual has a heavy workload (Getty, et. al., 1995). Findings of this study may affect the designs of future alerts and warnings in locomotives.

1.2 Objectives

The objective of this study was to conduct pilot studies using a locomotive simulator to observe whether engineers can become habituated to alerts. For example, the auditory alerts and warnings inherent in a Positive Train Control (PTC) system present habituation challenges. According to FRA, “Positive Train Control. . . systems are integrated command, control, communications, and information systems for controlling train movements with safety, security, precision, and efficiency. . . PTC systems vary widely in complexity and sophistication based on the level of automation and functionality being implemented, the system architecture including wayside systems (e.g., non-signaled, block signal, cab signal, etc.), and the degree of train control.” The purpose of these systems is to prevent train-to-train collisions, overspeed derailments, incursions into established work zone limits, and the movement of a train through a main line switch in the wrong position. Inherent in the design of PTC systems are status alerts and warnings to operational personnel when there is a malfunction or a system anomaly. Researchers examined whether engineers exhibited habituation behavior to these alerts.

1.3 Overall Approach

The team drew on their backgrounds in neuroscience, psychology, and data science to investigate whether habituation to auditory alerts could be observed in locomotive engineers.

1.4 Scope

To complete this pilot study, KEA researchers conducted the following tasks:

- **Task 1: Define Study Scope and Develop Experimental Design** - The research team designed an experimental study to examine the potential for locomotive engineers to habituate to alerts in the locomotive cab. This involved a literature search of habituation habits observed in other transportation fields. Based on the literature search, the team vetted an experimental design to test the phenomenon of habituation.
- **Task 2: Testing and Data Collection** - Researchers arranged for the review and approval of testing human subjects through an Institutional Review Board (IRB). Once approved, the team recruited and tested a pilot group of locomotive engineers in CTIL. All study variables were recorded throughout the individual testing sessions.
- **Task 3: Data Analysis** – The team analyzed the data and produced a final report.

1.5 Organization of the Report

Section 2 details the literature review. Section 3 outlines the experimental design of the research. Section 4 reports data analysis of the findings. Section 5 contains a summary of the research conducted and recommendations for future work. Appendix A contains the screening questions and the exit questionnaire.

2. Literature Review

Auditory alerts present a common and daily encounter that captures someone's attention. Unlike visual alerts, a person does not need to be within visual range of the alert to respond (Wickens, Lee, Liu, & Gordon-Becker, 2004), thus making them a potentially effective warning system. For safety alerts, the ability to reach the user is important (e.g., fire alarms, car chimes, back up sensors, etc.). With the use of auditory alarms comes the potential for technical problems that can render the alert non-productive. For example, the sound of the alert may not be ideal so that detection is difficult or may be a distraction. Alerts may be ambiguous leading to confusion about what action should be taken. Additionally, the alerting system may malfunction, which could trigger a false alert or a failure to alert when necessary (Ulfvengren, 2000; Wickens & Hollands, 2000). All these problems have the potential to lead to decreased confidence in the alert.

Another potential problem with alerts is habituation, defined as a reduced attentional response to repeated exposure to a stimulus. Mackworth (1969) states that when exposed to a constant stimulus over long periods of time, an individual's sensitivity or response to the stimulus is reduced. Even though habituation to warning signals is not a new concept, research on this topic is limited (Edworthy & Adams, 1996; Wogalter & Silver, 1995).

It is important to distinguish habituation from muscle fatigue, which is when a muscle's ability to generate force declines because of exhaustion. Habituation is not a result of exhaustion but desensitization to the need to respond. Warning signals may initially attract attention, but over time can become perceived as less important. Habituation to warnings is problematic for several reasons: 1) the signal loses its alerting capability; 2) people may prematurely dismiss the warning before encoding all the necessary information; and 3) subsequent warning signals may elicit a lowered response (Wogalter & Vigilante, 2006).

2.1 Auditory Habituation of Locomotive Engineers

In reviewing the literature, the team found that habituation to alerts and warnings behavior has been observed across multiple transportation sectors (i.e., aviation, automotive), as well as the healthcare industry.

In the rail industry, many locomotive engineers have raised concern over excess audio warnings and the need to respond to them being a workload issue. These concerns may result in several outcomes, such as annoyance, disregard for the system, and distraction. For example, research on signals passed at danger (SPAD) in Italian railways has shown that in-cab audio alerts can fail to serve their desired function (Pasquini, Rizzo, & Save, 2004). Pasquini et al. (2004) studied a case where a locomotive engineer missed a stop signal despite an auditory warning that required a clear acknowledgement of the alert through pushing a button. Pasquini et al. found that the alert system in place encouraged engineers to instinctively press the button acknowledging the alert without processing its meaning. Researchers suggested that this was occurring because the alerts that the system typically delivered were uninformative and inaccurate. Therefore, many experienced locomotive engineers were observed holding their fingers on the acknowledgement button, ready to disengage the alert as soon as they heard it to remove the distraction. Many of them pressed the button without looking down at the in-cab display screen indicating the meaning behind the signal.

In the United States, alerters in the locomotive cab are designed to keep the engineer alert and attentive by periodically requiring acknowledgement of a frequently sounding alert; if the alerter is not acknowledged, the locomotive will come to a stop. Acknowledging the alerter can become an annoyance and a distraction. Given the technologies engineers must attend to and the demand for attentional resources during operation of a train, the engineer may be susceptible to habituation, distraction, annoyance, and subsequent disregard of these alerts. The literature surrounding how locomotive engineers deal with their auditory environment when operating a locomotive is limited.

3. Pilot Testing Program Details

This section discusses details of the current study's pilot test in CTIL.

3.1 Background

While automation has many performance benefits, the shift to monitoring a computer-based system introduces potential disadvantages (Sheridan et al., 1999), including the increased risk of habituation to alerts.

As more technology is developed to improve transportation safety (e.g., PTC, driver-assist features, and automated driving), there is the potential for error. Humans must be aware of and react to the warnings provided by new technologies. Researchers conducted a pilot study to examine the risk of habituation in locomotive engineers.

3.2 CTIL Track

Nine experienced engineers spent 4 to 6 hours in CTIL operating a train across the same 300-mile freight track. The simulation used novel tones that would act to dishabituate the engineer and included areas with high and low workload.

3.3 Participant Recruitment

Recruitment began with a referral from a railroad union member used as a recruiter. After the first few studies, recruitment snowballed when locomotive engineers referred their colleagues or friends who were locomotive engineers.

The nine participants were chosen according to several criteria:

- Participant can understand and comply with the protocol
- Participant is available for 6 to 8 hours on the study day
- Participant is in good physical and mental health
- Participant must be a current locomotive engineer (freight or passenger)

The participants also were chosen according to several exclusion criteria:

- Individual is not familiar with freight operations as locomotive engineer
- Individual is not a current locomotive engineer

Eligible participants could be current passenger engineers but were required to have been a licensed freight engineer for at minimum 1–2 years and be comfortable with running freight operations.

3.4 Data Collected

Pilot study participants answered a demographic questionnaire and an exit questionnaire related to the study. These questions can be found in Appendix A.

Three cameras collected footage in CTIL from behind, directly in front of, and to the side of the participant. This captured any head or body movements in response to stimuli. The track itself was also recorded. Additionally, a microphone designed for reducing background noise collected audio in the cab, automatically filtering out the noise of the simulator to better hear the participant. CTIL automatically produces timestamps of participants' button or control modifications. Participants were asked to press the "attendant call button" when a dishabituator sounded in the cab, a button not typically used in simulator operation.

3.5 Dishabituation Tones

The team selected two novel tones that were distributed along the track at varying intervals, in areas of high, medium, and low workload. The novel tones were distinct from other auditory alerts already present within the cab environment.

Each participant operated across as much CTIL territory as possible within the given timeframe. Tones were played in random intervals along the track. Each participant experienced between 8 and 10 dishabituator tones during the study.

3.6 Testing Adjustments

A few minor adjustments were made after the first two participants completed the study. These participants did not make it as far along the inbound route as was anticipated, so some elements of the track were adjusted for future participants, including increasing speeds in some areas and changing the restricted speed to 30 mph. Unlike normal operations, researchers told engineers that they could receive Northern Operating Rules Advisory Committee (NORAC) Form D's and fill out Temporary Speed Restriction Bulletins (TSRBs) while moving at restricted speed instead of stopping as required; this also increased the workload for the engineer. Also, CTIL does not permit some of the permanent signage on the track to be updated, so engineers were asked to ignore these signs and follow the PTC, which may have caused confusion for some participants. Finally, two to four tones were added to the segment of track that the engineers were most likely to complete in the allotted study time to account for the tones missed on the unused portion of track (i.e., most of the outbound track).

Payment for participation in the study also increased slightly, from $750 to $950. This change accounted for the additional risk of participation during the COVID-19 pandemic and helped with recruitment.

The first participant suggested that researchers provide participants with a brief orientation run in the simulator to familiarize themselves with the PTC system and the tones that would occur. Additionally, the participant requested a set of background materials they could keep with them throughout the study. This allowed the participants coming from different carriers or who did not currently operate freight rail (i.e., currently a passenger rail engineer) to become familiar with operating in the simulator under different sets of rules and equipment. The background materials included details about the train's composition, the route, rules for the route that may differ from NORAC (i.e., the northeast standard with which most participants were familiar), and reference charts of the permanent speed restrictions for the inbound and outbound routes.

3.7 Added Workload Events and Tone Placement

Table 1 and Table 2 detail the workload events and tone placement for each leg of the pilot test.

Table 1: Moving Towards Chicago: In-Bound Main 2

Track Section	Workload	NORAC Form D	Added Events (dispatch calls)	Add Tone
185–180	High	S&P- 181.46 [186.75]		Add tone 1 at 184.0
175–165	Medium		Cart on track at 165.16 [174.9]	Add tone 2 at 169.44
135–130	High		TSR 40 134.37-132.289 [142.00]	no
125–120	High	Work zone 123.988-121.96		Add tone 2 at 120.661
105–100	High		Add approach medium (yellow over green) 104.189**change as the train approaches	Add Tone 1 at 102.871 (Ridge Rd)
85–80	Low		Add chatter 185.81	no
70–65	Low			Add tone 1 at 67.254
40–39 (end of Aurora track)	Low		39.429—pull off and take new form D**check in with engineer	
35–30	Medium			Add tone 2 at 33.273 (SL-1-bridge-ABSL)
25–20	Medium		Approach and stop 20.411 [23.952]**do not include in PTC	Add tone 2 at 20.782
15–10	Medium	Work zone 14.796-13.08		Add tone 1 at 10.479

Table 2: Moving Away from Chicago: Outbound Main 1

Track Section	Workload	NORAC Form D	Added Events (dispatch calls)	Add Tone
25–30	Medium		Add approach medium (yellow over green) 29.808	Add tone 1 at 26.81
30–35	Medium			Add tone 2 at 31.001 (Bridge)
38–40 (end of Chicago track)	Low	38.38 pull off and take a new form D - - pull off into siding, for work zone at 170 and TSR at 25		No
40–50	Low			Add tone 1 at 47.189
70–75	Low			Add tone 2 at 72.269
105–120	High	TSR 40 116.491-117.913 [107.39]		Add tone 1 at 116.89
120–125	Low		Add chatter 122.988	
130–135	High		TSR 45 131.139-133.15 [127.27]** do not include in PTC	Add tone 1 at 131.926
155–160	Low			Add tone 2 at 158.206
170–175	Medium	Work zone 171.584-173.07		Add tone 1 at 174.978

4. Data Analysis

This section documents analysis of the data collected in the study.

4.1 Background

Researchers removed the first participant (HALC-01) from the dataset. The first participant's run was treated as a trial to inform and verify study and track design. The team analyzed data from eight participants, HALC-02 – HALC-09 using R software (for statistical computing). Data analysis focused on determining if the participants experienced habituation to the auditory alerts in the locomotive cab during the study. Additionally, the team analyzed more than 40 hours of video to understand participants' behavior and operational responsibilities while tones sounded during the study.

4.2 Data Collected

Data from CTIL, video recordings, participant surveys, and track workload were all analyzed to determine whether habituation was present.

4.2.1 CTIL Data

Data collected from CTIL includes:

- Time elapsed from start
- Distance in chainage (can be converted to milepost)
- Attendant call button press
- Grade
- Speed
- Acceleration
- Bell
- Horn
- Alerter
- Alerter penalty

4.2.2 Video Recording

Video data was collected from four different sources for each participant.

- Screen 1 showed the PTC display throughout the study
- Screen 2 showed the Main Display
- Screen 3 showed the Track Simulation
- Screen 4 showed video taken from behind the participant in CTIL

4.2.3 Surveys

Demographics: All participants completed a demographics questionnaire.

Exit Questionnaire: All participants completed an exit questionnaire that asked them to evaluate how similar their experience in the simulator was to their typical workday.

4.3 Results

This section documents the results from the data analysis.

4.3.1 Survey Analysis

The first questionnaire collected demographic information from participants such as age, sex, race, years of experience, years of exposure to PTC, and typical duration of current shift. The exit questionnaire asked participants to describe how similar they found the track simulation to a real workday and to note differences, including limitations of the simulator. Participants were also asked to compare the duration and signal workload of the study to a regular workday, and to report whether they experienced any fatigue during the study.

All participants were male. The average age was 48.7 years, with the youngest participant being 41 years old and the oldest being 63 years old. Seven participants identified as White or Caucasian, while one identified as Hispanic and another identified as Asian. The average years of experience as a locomotive engineer was 17.5 years, with the fewest being 9 years and the most being 28 years. The average years of PTC exposure among participants was 7.9 years, with the fewest being 0 years and the most being 17 years. Four participants exclusively operated freight locomotives, whereas the remaining five had experience running freight and passenger operations. The typical duration of participants' shifts in their real workdays was approximately 10.3 hours, although many gave ranges of possible shift times. Most participants operated routes in the Greater Boston area.

Based on responses from the exit questionnaire, most participants found the track simulation used in the study very similar to the real workday of a freight engineer. Several participants noted dissimilarities including unfamiliarity with the territory, the differences in rules, and the lab environment. Examples included using General Code of Operating Rules (GCOR) instead of NORAC rules and allowing participants to receive Form Ds at restricted speed instead of stopping. Additionally, restricted speed and speeds experienced generally in the simulator were higher than the speeds with which the participants were accustomed. The participants also noted that the simulator is not as audibly loud and does not have vibration or slack/buff force effects, particularly on hills and curves. One participants noted that it had been ten years since he operated a freight locomotive after switching to passenger operations, a significant amount of time. A few participants noted that they don’t experience beeping from the end-of-train device during normal day-to-day operations and that their PTC system was different from the one used in CTIL, which made it more challenging.

Six participants commented that they thought the study duration was shorter than the regular workday of a freight engineer, but the remaining three participants said that the duration was similar. One participant commented that while it was shorter, the mileage was close to what he typically experienced in a workday due to the increased speeds traveled in the simulator. Another participant commented that the simulation was operating more hours continuously than his regular workday because he usually switches trains and may have some downtime between switching; he still found the total duration was much shorter than his typical day, which is between 10.2–12 hours. Four of the nine participants found the signal workload in the simulation to be similar to their regular freight operations. One participant thought it was easier, while the remaining four believed that the simulator had a higher signal workload than what they typically experience and thus was busier than usual.

Five participants commented that they felt no fatigue during the study. Two of those five commented that, if anything, they felt more attentive during the study because of the unfamiliarity of the territory and simulator environment, so they were more focused and stimulated than usual. Two other participants said it was similar to running a train, and the fatigue they felt was "not bad." One participant commented that not being qualified on the territory made it "pretty tough" and tiring, making him more fatigued than a normal workday. The remaining participant commented that during periods of low stimulus, he experienced some fatigue, but it subsided during high stimulus periods.

4.3.2 Simulation Data Analysis

Table 3 shows the total number of novel tones and computer tones introduced to the participant. A successful response is indicated if the participant pressed the attendant call button within 15 seconds of the tone sounding. The rightmost column displays the number of false responses or times the participant pressed the attendant call button in response to a non-dishabituating tone inside the cab. The number of false responses varied substantially between participants, which may be due to differences in the audio alerts they are accustomed to versus the audio alerts used in the simulator.

Early in the data analysis process, researchers determined that several participants were responding to a tone intended to remind the experimenter running CTIL to act as the dispatcher. This tone, which appeared between 10–13 times depending on the participant, came from the experimenter's laptop and was audible over the radio used to correspond with the participant. Due to the difference of this tone from typical sound in the locomotive cab, these reminders were evaluated as if they are additional novel (i.e., dishabituating) tones. This resulted in some cases having double the intended amount of dishabituating tones.

Table 3: Response Rate to Dishabituating Tones and False Tones by Participant

Participant Number	Number of Computer Tones	Number of Responses	Response Rate (%)	Number of Novel Tones	Number of Responses	Response Rate (%)	Number of False Responses
HALC02	11	2	18.18	6	6	100	2
HALC03	11	1	9.09	8	8	100	1
HALC04	10	9	90	7	6	85.71	9
HALC05	11	6	54.55	9	8	88.89	10
HALC06	13	7	53.85	10	8	80	10*
HALC07	11	9	81.82	8	5	62.50	16
HALC08	11	6	54.55	10	9	90	2*
HALC09	11	9	81.82	10	9	90	16*

**Some of the extra false responses were due to researcher error, such as hitting a button on the computer used for the simulation and a Windows Microsoft error tone sounding. This led to two extra false responses in HALC-06, 1 extra false response in HALC-08, 1 extra false response in HALC-09.*

The total number of both computer tones and novel tones varied slightly for each participant. Participants 05, 06, 08, and 09 were the only ones to complete a small section of outbound track in the given period, so they had the potential to receive up to two more novel tones and two more computer tones compared to those who only completed the inbound route.

Any tones that the researchers could not hear while conducting the video analysis were not counted in the analysis, as the participant would not have been able to respond to a tone that was not audible or was not administered. For example, some of the computer tone reminders were disabled by the team, usually due to timing constraints. For participant 02, only six novel tones had been programmed for the study. However, two more novel tones were added for all subsequent participants. For participant 04, both a computer tone and novel tone were not audible to researchers. Similarly, participants 08 and 09 each had a computer tone that was not audible to researchers, and during participant 05's study, a tone occurred right as the simulation was paused for lunch, so the participant did not have a chance to respond.

Table 4 shows the aggregated successful response rate (%) for all tones among all participants as well as for each participant. The table also shows the successful response rate for each dishabituating tone introduced (i.e., two novel tones and the computer tone). The number of events to which each participant had a chance to respond is also presented, which varied across participants.

Table 4: Aggregate Response Rate to Dishabituating Tones

Response Type	Participant	Event Type	# of Events	# of Successful Responses	Successful Response Rate (%)
All Participants-All Tone Types	All Participants	All Tone Types	157	108	68.8
HALC-02-All Tone Types	HALC-02	All Tone Types	17	8	47.1
HALC-03-All Tone Types	HALC-03	All Tone Types	19	9	47.4
HALC-04-All Tone Types	HALC-04	All Tone Types	17	15	88.2
HALC-05-All Tone Types	HALC-05	All Tone Types	20	14	70.0
HALC-06-All Tone Types	HALC-06	All Tone Types	23	15	65.2
HALC-07-All Tone Types	HALC-07	All Tone Types	19	14	73.7
HALC-08-All Tone Types	HALC-08	All Tone Types	21	15	71.4
HALC-09-All Tone Types	HALC-09	All Tone Types	21	18	85.7
All Participants-Computer	All Participants	Computer	89	49	55.1
All Participants-Sound 2	All Participants	Sound 2	29	25	86.2
All Participants-Sound 2A	All Participants	Sound 2A	39	34	87.2

4.3.3 Workload

After the study, the research team consulted with an experienced engineer and study participant to retroactively label areas of the track as high, high-medium, medium, low-medium, or low workloads. This was then converted into a numerical score where 1 is the lowest and 5 is the highest workload. The engineer was given the track chart of the simulated track and all the events added by the research team, such as work zones and restricted speed areas. The engineer considered speed, grade, number of crossings, and number of events happening concurrently to rate and label sections of track with their corresponding workload. Table 5 and Table 6 show the track section with the corresponding workload score and the notes from the engineer around the justification of the workload score.

Table 5: Workload Scoring of Selected Inbound Track with Researcher Additions

Track Section	Workload	Engineer Notes
193.1 – 190	HIGH	Engineer is getting train up to speed and is performing routine tasks (i.e., releasing brakes, advancing throttle, ringing bell, etc.). A train with this consist and Horse Power per Ton (HPT) would take about 2–3 miles to achieve operating speed.
190 – 187	MEDIUM	Engineer has a good "feel" of the train and is settling into the run.
187 - 181	MEDIUM to HIGH	Engineer's run is elevated from Med to High due to Stop and Protect awareness at Mile Post (MP) 181.65. Engineer was notified at MP 186.74. The Engineer has 5 miles to bring the train to a full stop. Note: track chart shows a #11 turnout at this location, not a X-ing. First time "new" tone 1 is heard at MP 184.
181 – 175	MEDIUM	Routine running
175 – 165	MEDIUM	Engineer is notified of "cart on track" at MP 165.16, dispatcher alerted. Engineer at MP 174.9. Tone 2 at MP 169.44.
165 – 137	LOW	Undulating territory begins at MP 143
137 – 132	HIGH	Temporary Speed Restriction (TSR) of 40 mph between MP 134.46 and 132.36, engineer must reduce speed from 70 mph within undulating territory consistent with proper train handling, loads on rear of train may cause some "run in" or buff force depending on the course of braking action taken (power braking, fully dynamic or a combination). Engineer must allow at least 2 miles for train to meet target speed at TSR location. Engineer can use undulations in track to aid in reducing speed.
132 – 126	MEDIUM	Engineer is operating over undulating territory.
126 – 120	MEDIUM to HIGH	Work zone from MP 124.08 to MP 122.03. Engineer will have a heightened sense of awareness as they approach and pass through this area still within undulating territory. Tone 2 is also heard at MP 120.5.
120 – 105	LOW	Undulating territory, 70 mph.
105 – 85	LOW TO MEDIUM	Approach Medium signal at MP 104.199 (Engineer must begin to reduce speed from 70 mph to 40 mph). Tone 1 heard at MP 100.9 (note: "Ridge Rd." does not appear on track chart). Territory undulates and there are many X-ings.
85 – 82	MEDIUM	Multiple speed changes within this area and many X-ings.
82 – 40	LOW	Undulating territory. 50 MPH at "STEWARD" otherwise 70 mph. Tone 1 at MP 67.254 (pedestrian grade X-ing).
40 – 38	MEDIUM	End of one Subdivision, enter second Subdivision (MP 38.40), two speed changes within two miles. Grade separation from X-ings.
38 – 25	MEDIUM	High-speed running, few X-ings (mainly grade separated. Engineer copies a Form D at MP 35.84 (TSR MP 21.12 – 20.12). Fairly flat territory. Tone 2 at MP 33.273 EOLA Interlocking (signal bridge).
25 – 20	HIGH	Engineer anticipation of multiple actions needed to be taken within a short distance including controlling and stopping train from 70 mph. TSR MP 21.12 to MP 20.12 (Downers Grove to Fairview Ave), tone 2 at MP 20.782 (Fairview Ave signal bridge). Uphill grade will aid Engineer in slowing train.
20 – 15	MEDIUM TO HIGH	Engineer begins to start train after stopping. Anticipates remainder of run. Downhill grade with multiple X-ings in multiple track territory with stations.
15 – 9	MEDIUM TO HIGH	Downhill grade with multiple X-ings in multiple track territory with stations. Form D MP 14.51 to MP 13.79, tone MP 10.479.
9 – 1.2	MEDIUM	Grade begins to flatten out, X-ings become grade separated. Multiple track territory with stations.
1.2	END OF INBOUND	END

Table 6: Workload Scoring of Selected Outbound Track with Researcher Additions

Track Section	Workload	Notes
14.3 – 25	MEDIUM	70 mph. 19.782 is top of grade, Engineer will transition from pulling to controlling speed going downhill, for next 5 miles, using preferred braking method. Multiple X-ings in multiple track territory with stations.
25 – 30	MEDIUM TO HIGH	70 mph. Undulating territory. Tone 1 at MP 26.810 (Naper Blvd). Approach Medium signal at MP 29.808. Few X-ings in multiple track territory with stations.
30 – 35	LOW to MEDIUM	70 mph. Tone 2 at 31.001 (ROUTE 59 Interlocking). Undulating territory with no X-ings.

Researchers used this information to examine whether response rate was significantly affected by workload (Figure 1). The X-axis of this graph is a workload score rating where 1 = Low, 2 = Low-Medium, 3 = Medium, 4 = Medium-High, and 5 = High. A linear regression model showed no statistically significant linear relationship between workload and response rate.

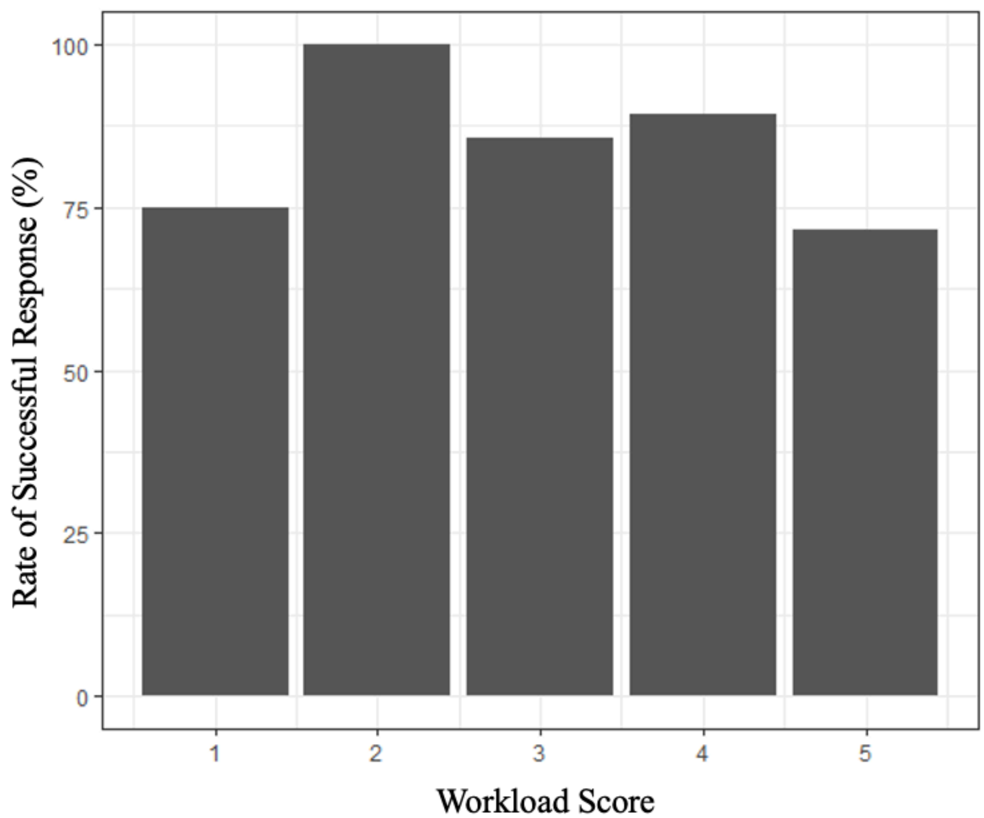


Figure 1: Analysis of the Response Success Rate as a Function of Workload Score

4.3.3.1 Analysis of Response Time

Table 7 shows the mean response time in seconds to the dishabituating tones to which participants successfully responded (within 15 seconds). The table shows the response times aggregated across all participants and tone types and then broken down by each participant individually and by sound type. The mean response time to the computer tones was longer compared to the intended novel tones selected as dishabituators. However, there was not a meaningful difference in mean response time between “Sound 2,” “Sound 2a,” or the computer tones. This data is also shown in Figure 2.

Table 7: Mean Response Time of Response to Tones by Participant and Sound Type

Response Type	Participant	Event Type	n	Mean Response Time (Seconds)
All Participants-All Tone Types	All Participants	All Tone Types	108	3.655
HALC-02-All Tone Types	HALC-02	All Tone Types	8	3.188
HALC-03-All Tone Types	HALC-03	All Tone Types	9	2.389
HALC-04-All Tone Types	HALC-04	All Tone Types	15	2.200
HALC-05-All Tone Types	HALC-05	All Tone Types	14	3.943
HALC-06-All Tone Types	HALC-06	All Tone Types	15	6.067
HALC-07-All Tone Types	HALC-07	All Tone Types	14	3.250
HALC-08-All Tone Types	HALC-08	All Tone Types	15	4.860
HALC-09-All Tone Types	HALC-09	All Tone Types	18	2.783
All Participants-Computer	All Participants	Computer	49	4.155
All Participants-Sound 2	All Participants	Sound 2	25	3.296
All Participants-Sound 2A	All Participants	Sound 2A	34	3.197
HALC-02-Computer	HALC-02	Computer	2	6.300
HALC-02-Sound 2	HALC-02	Sound 2	3	2.167
HALC-02-Sound 2A	HALC-02	Sound 2A	3	2.133
HALC-03-Computer	HALC-03	Computer	1	2.800
HALC-03-Sound 2	HALC-03	Sound 2	4	2.400
HALC-03-Sound 2A	HALC-03	Sound 2A	4	2.275
HALC-04-Computer	HALC-04	Computer	9	2.367
HALC-04-Sound 2	HALC-04	Sound 2	2	1.850
HALC-04-Sound 2A	HALC-04	Sound 2A	4	2.000
HALC-05-Computer	HALC-05	Computer	6	4.300
HALC-05-Sound 2	HALC-05	Sound 2	2	5.100
HALC-05-Sound 2A	HALC-05	Sound 2A	6	3.200
HALC-06-Computer	HALC-06	Computer	7	6.586
HALC-06-Sound 2	HALC-06	Sound 2	3	5.033
HALC-06-Sound 2A	HALC-06	Sound 2A	5	5.960
HALC-07-Computer	HALC-07	Computer	9	3.322
HALC-07-Sound 2	HALC-07	Sound 2	3	2.900
HALC-07-Sound 2A	HALC-07	Sound 2A	2	3.450
HALC-08-Computer	HALC-08	Computer	6	6.517
HALC-08-Sound 2	HALC-08	Sound 2	4	4.325
HALC-08-Sound 2A	HALC-08	Sound 2A	5	3.300
HALC-09-Computer	HALC-09	Computer	9	2.889
HALC-09-Sound 2	HALC-09	Sound 2	4	2.825
HALC-09-Sound 2A	HALC-09	Sound 2A	5	2.560

Figure 2 is a box-and-whisker plot showing an analysis of the response time in seconds by each tone that served as a dishabituator. Only successful responses, or responses within 15 seconds of the tone, are shown in this analysis. For all box-and-whisker plots in this task report, each colored dot corresponds to a unique participant’s responses. The box-and-whisker plots note which colored dots are statistical outliers with black dots showing the same y-value (i.e., at the same latitude). In Figure 2, no significant difference exists in response time between the different dishabituating tones

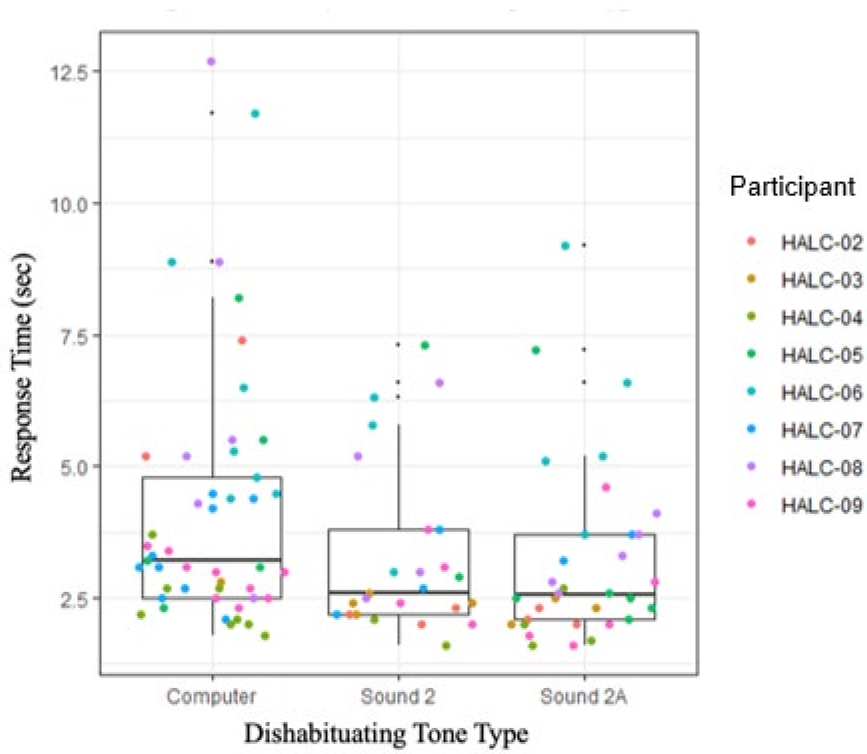


Figure 2: Analysis of the Response Time by Dishabituating Tone

Figure 3 is a box-and-whisker plot showing the response time in seconds to dishabituating tones under various workload conditions. A score of 1 is the lowest workload while 5 is the highest. No significant difference was found in the response time under different workloads.

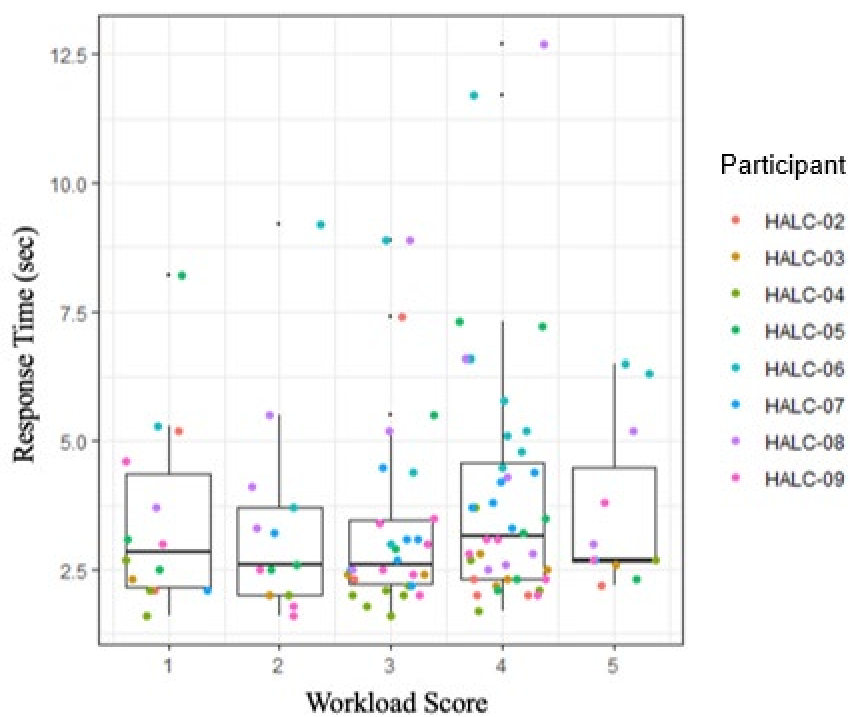


Figure 3: Analysis of the Response Time as a Function of Workload Score

4.3.3.2 Response Recovery Analysis

As previously noted, habituation is defined as reduced attentional response to repeated exposure to a stimulus. In the context of warning signals, over time the alert becomes coded as less salient and thus associated with less importance. Rankin et al. (2009) identified ten characteristics that specifically define habituation and distinguish it from fatigue or sensory adaptation. One of the essential criteria of habituation is that the response decrement observed must be progressive and needs to follow exponential or linear trends (Rankin et al., 2009). Additionally, habituation is stimulus specific and dishabituation must be possible. In auditory examples of habituation, a dishabituator would serve as a novel or deviant tone that is inserted into a set of repeated tones and likely varies in frequency and/or amplitude of the repeated tone. When a dishabituator is presented, response recovery is observed, meaning that the habituation process is interrupted and the response to the previously habituated tone should return to what the response was right before the presentation of the dishabituator (Rankin et al., 2009). In contrast, if the reduced response observed was attributed to fatigue or sensory adaptation, that trend of reduced response would continue even if a dishabituator was presented. Response recovery to the previously habituated tone would not occur in cases of fatigue or sensory adaptation (Rankin et al., 2009).

To determine if response recovery was observed, an analysis was performed on the response times to the alerter, which is a safety feature in the locomotive cab. The alerter warning appears when an engineer has not touched the controls in a specified set of time. The alerter is a known feature to any locomotive engineer. Typically, the alerter first appears on the main control screen as a visual countdown without sound for a certain amount of time (for this study, 5 seconds). An auditory alarm then starts to sound if the alerter is not acknowledged within that time. If it is ignored for a period after the alarm sounds, the locomotive initiates penalty brake application. No alerter penalty braking occurred in the study data set, meaning all alerters were acknowledged. The analysis of response recovery excluded responses for events where the participant did something other than push the alerter button to resolve the alarm. This includes the engineer using any of the controls on the control stand.

A participant's response time to the alerter was calculated as the time from the start of the alerter to the time the participants pressed the alerter button. Researchers identified alerter responses that were the "First" and responses that were the "Last" before a novel tone. All other responses were identified as "Other." Accordingly, if a response was both the first and last response between any two novel tones it was reclassified as "Other."

Only inbound data was analyzed because the length of the outbound rides (when completed) were short. Researchers only received a few points that could be analyzed and presumably there would have been no time to become habituated. The hypothesis is that in between tones the participant becomes habituated to the alerter and will respond more slowly over time. Dishabituation and response recovery would be seen if the last alerter response before the dishabituating tone is slower than the response to the first alerter.

Table 8 shows mean response time (in seconds) for all response types (first, last, and other) across all participants as well as for each. It also shows mean response time for the "First" and "Last" alerter responses aggregated across participants and then separated by each. None of these differences are statistically significant.

Table 8: Analysis of the Response Time to the Alerter Before and After Dishabituating Tone

Participant	Response Type	n	Mean Response Time (Seconds)
All Participants	All Response Types	288	3.66
HALC-02	All Response Types	24	4.79
HALC-03	All Response Types	16	3.56
HALC-04	All Response Types	35	4.11
HALC-05	All Response Types	46	3.10
HALC-06	All Response Types	37	3.50
HALC-07	All Response Types	36	3.70
HALC-08	All Response Types	49	3.09
HALC-09	All Response Types	45	4.04
All Participants	First	66	3.75
All Participants	Last	68	3.52
All Participants	Other	154	3.68
HALC-02	First	5	4.80
HALC-02	Last	6	4.63
HALC-02	Other	13	4.85
HALC-03	First	4	3.00
HALC-03	Last	4	3.98
HALC-03	Other	8	3.62
HALC-04	First	9	3.16
HALC-04	Last	9	4.17
HALC-04	Other	17	4.58
HALC-05	First	11	2.82
HALC-05	Last	11	3.09
HALC-05	Other	24	3.23
HALC-06	First	9	3.77
HALC-06	Last	9	3.36
HALC-06	Other	19	3.44
HALC-07	First	7	3.70
HALC-07	Last	8	3.55
HALC-07	Other	21	3.76
HALC-08	First	11	3.81
HALC-08	Last	11	2.53
HALC-08	Other	27	3.02
HALC-09	First	10	5.05
HALC-09	Last	10	3.79
HALC-09	Other	25	3.74

Figure 4 is a box-and-whisker plot showing response time (in seconds) aggregated across participants for the First alerter response and the Last alerter response before a novel tone. The colored dots represent each unique participant. There was no statistically significant difference between the response time to the First alerter after a dishabituating tone and the Last alerter before a dishabituating tone.

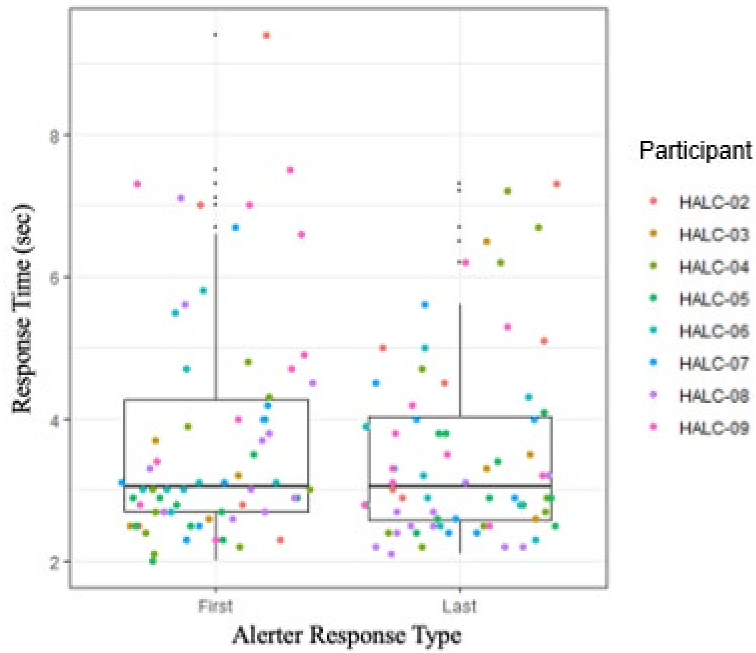


Figure 4: Analysis of the Response Time (Seconds) to the First and Last Alerter Before and After a Dishabituating Tone

Figure 5 is a box-and-whisker plot showing response time by participant ID for the First alerter response and the Last alerter response before a novel tone. No significant difference was found between the two groups. Response as a function of participant ID was statistically significant which means some participants responded faster than others. However, this is not important to the analysis and doesn't impact the conclusions.

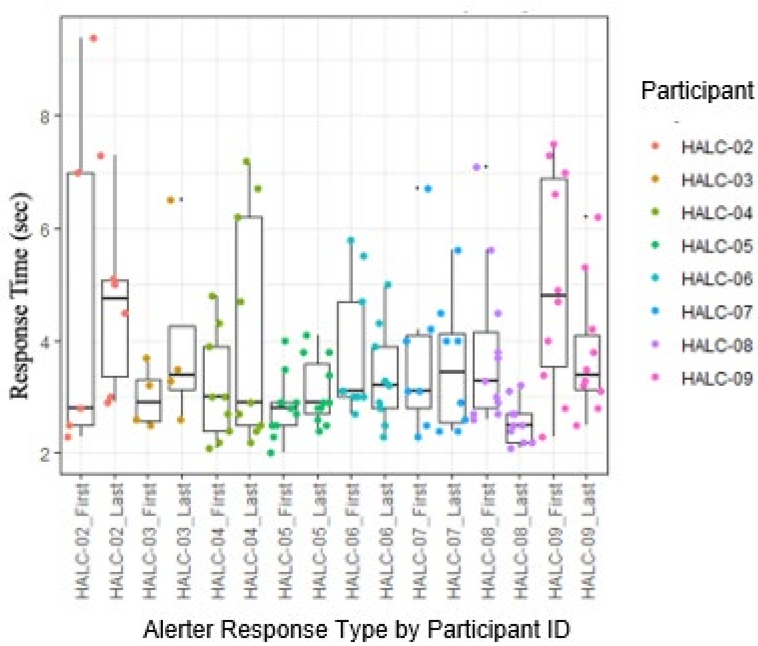


Figure 5: Analysis of the Response Time to the First and Last Alerter in Reference to a Novel Tone by Participant

4.3.3.3 Video Analysis

The time and distance at which the attendant call button was pressed was recorded in the CTIL data. Researchers reviewed video footage from the minute leading up to each press of the button to verify the tones or sounds to which the engineer was responding and evaluate the engineer's workload in that minute, as well as the engineer's perceived mood and confidence in pressing the button. The attendant call button press chainage (Table 7) was compared with the known chainages of the novel tones and reminder tones to calculate a response time to each. Approximately 43 hours of video footage was collected during this pilot study. The results of the video analysis are summarized in Table 9.

Participant	Missed Tones	Participant is not doing anything (e.g., hands in lap)	Participant is actively doing something else (e.g., sounding the horn and bell, taking a Form D, etc.)	Other auditory tones are sounding in the cabin	Participant seems to notice tone but chooses not to acknowledge	Researchers do not hear tone or see participant respond in video
HALC-04	Computer Tone #1	X				
	Novel Tone #6					X
	Novel Tone #7		X			
	Computer Tone #11					X
HALC-05	Computer Tone #1	X				
	Novel Tone #2		X			
	Computer Tone #4	X				
	Computer Tone #6		X			
	Computer Tone #7		X	X		
	Computer Tone #9		X	X		
	Novel Tone #7		X	X		
HALC-06	Computer Tone #3		X	X		
	Computer Tone #6		X			
	Novel Tone #5		X		X	
	Computer Tone #8		X	X		
	Computer Tone #9	X			X	
	Novel Tone #6		X			
	Computer Tone #12	X		X	X	
	Computer Tone #13	X		X		
HALC-07	Novel Tone #5		X	X		
	Computer Tone #8		X	X		
	Computer Tone #9			X	X	
	Novel Tone #7		X	X		
	Novel Tone #8		X	X		
HALC-08	Computer Tone #3			X	X	
	Computer Tone #5	X				
	Computer Tone #6		X			
	Computer Tone #7		X			
	Computer Tone #8		X			
	Computer Tone #9					X
	Novel Tone #9		X		X	
HALC-09	Computer Tone #1		X		X	
	Computer Tone #3		X	X		
	Computer Tone #8					X
	Novel Tone #9	X		X		

**The tones in red in the table are not counted toward the total number of computer and novel tones because researchers could not hear them using the video footage and assumed that the participant also missed them.*

Behavioral analysis was conducted, including categorizing whether a participant was busy actively doing something else (e.g., sounding the bell and horn, actively braking, etc.). The team also noted whether other auditory alerts and tones (e.g., the end-of-train device, alerter, or radio chatter) were sounding in the cabin at the time. False responses were also reviewed to understand where confusion or differences between participants was occurring.

Researchers examined whether the participant seemed to notice or hear the tone but elected not to respond. The team noted head movement indicating the participant looked at the control stand (i.e., toward the attendant call button), generally looked around the cabin toward the sounds, or moved their hand toward the attendant call button but stopped short of pressing it. At other times, the participant was facing forward and without other distractions but they did not seem to notice the auditory tone. There were four instances when a computer tone simply did not sound (researchers may have manually disabled the tone in these areas). These instances are noted in red in Table 9.

4.4 Study Limitations

As previously stated, one of the limitations of the study was the addition of many extra dishabituators in the form of experimenter reminders. Up to 13 additional computer tones were used, with meant participants may not have had time to experience habituation in between tones, so any effect on reaction time after the dishabituator may not be visible. This was not something the team could mitigate, because the first few participants did not respond as often to the reminders as subsequent participants.

The team experienced other study limitations that may have affected the data. Conducting the study in a simulator with unfamiliar personnel watching may have factored into the participants' performance. A simulator does not provide the same vibration, tilting, or momentum shift that an engineer would experience in a real locomotive. Additionally, while all the participants were locomotive engineers familiar with freight operations, they were all running different routes in their normal workdays that had varying degrees of similarity and difference to the simulated track. Typically, an engineer will be extensively trained on a route before running it alone. Engineers in this study were asked to run a simulated track unfamiliar to them and without extensive training. Also, the operating rules for the simulated track were different from typical operations (such as using GCOR rules instead of NORAC). To control for this limitation, researchers reviewed these unfamiliar rules with the engineer prior to entering CTIL and gave them the option to bring a printout of the rules into CTIL.

Researchers believe that most of the engineers were extra attentive during the study given the unfamiliarity of the territory, rules, and environment. Therefore, observing habituation was challenging.

4.5 Discussion

The comparison of the last alerter response time before a dishabituating tone versus the first alerter response after the dishabituating tone reveals that response recovery was not observed and thus habituation could not be observed. Researchers hypothesized that if response recovery was occurring, the first alerter response after a dishabituating tone should be faster than the alerter response just before the dishabituating tone. However, Table 8 shows the opposite trend, with the first alerter responses after a dishabituating tone being longer (3.75 seconds) compared to the last alerter response before a dishabituating tone (3.52 seconds) across all participants. This could also suggest that while habituation is not occurring, fatigue may be setting in. , given that response time did not return to what it was as one would expect before the dishabituating tone. Continued diminishing response time is an indicator of fatigue (Rankin et al., 2009). However, none of these differences were statistically significant and thus point to the need for a larger sample size.

While habituation was not observed in this study, the analysis suggests that analyzing workload and its effect may be an important factor in habituation. Figure 1 shows that as participants' workload increased, their response success rate to dishabituating tones across generally decreased. While these results were not statistically significant (as shown in Figure 3), the study's sample size of eight engineers was small. Future studies with an increased number of participants may yield different results.

Figure 5 shows that participants had a statistically significant difference in response time from each other, which also may point to needing a larger sample size for future studies. Additionally, the recruitment requirement that engineers must have at least one year of freight experience limited the candidate pool. Many engineers that participated in the study were previously freight engineers but had since switched to operating passenger locomotives and thus were less familiar with freight operations. Researchers examined differences between participants who were operating freight locomotives at the time of the pilot study, and did not find consistent trends regarding response times. Additionally, years of experience did not appear to play a role in response times nor did operating under the same rail carrier. For future studies, a larger sample size or more narrow recruitment criteria may reveal or bring clarity to these possible trends.

4.6 Suggested Future Research

Participants' feedback on the exit survey suggested that operating on unfamiliar track or in the presence of the researchers likely limited researchers' ability to study habituation in CTIL's simulated environment. Many engineers noted that they felt more attenuated and stimulated compared to their normal workday because of the unfamiliar track and environment. Given the challenges of observing habituation in a simulated environment, the team recommends future research be conducted in the field (i.e., on an actual locomotive) using video cameras and novel tones. Data collected should evaluate response time to various alerts during normal operation of a locomotive from experienced engineers before and after a dishabituating tone is introduced. Additionally, it may be useful to restrict recruitment to engineers who work for the same carrier or over similar territories and possibly with similar years of experience to limit the number of variables when assessing habituation.

5. Conclusion

Researchers conducted a literature review to compile techniques used in other industries to study habituation. The literature review also helped inform the experimental plan. Researchers used CTIL to conduct a pilot study with nine locomotive engineers; data from eight engineers was used in the study's final analysis.

Researchers collected 40 hours of video in CTIL and conducted data analyses on the potential effects of habituation to safety systems in the locomotive cabin (e.g., the alerter). Habituation to alerts was not observed. The team recommends further research that uses a larger sample size, has fewer dishabituating tones, and is conducted in a more realistic setting.

Researchers also observed diminishing response times to some auditory alerts, which is a sign of fatigue. While the differences in response time were not statistically significant, it is a potentially interesting finding and may suggest that fatigue is a greater risk to engineers than habituation. Additionally, as the workload during the study increased, the response success rate to dishabituating tones across all participants generally decreased. While these results were also not statistically significant (likely due to the small sample size), future studies with an increased number of participants may yield different results.

Since observing habituation is challenging in a simulated environment, researchers recommend that future research to evaluate habituation be conducted in the field. Based on the results from the study and exit questionnaires conducted with the locomotive engineers, the research team hypothesized that the more novel the situation, the less likely habituation will be observed. Therefore, future studies should be done by monitoring (through video) an engineer completing their regular route and their response times to certain alert systems throughout the ride to evaluate whether engineers become habituated in a more realistic setting.

6. References

Arrabito, G. R., Mondor, T. A., & Kent, K. J. (2004). Judging the urgency of non-verbal auditory alarms: A case study. *Ergonomics*, *47*(8), 821–840. https://doi.org/10.1080/0014013042000193282.

Bitan, Y., Meyer, J., Shinar, D., & Zmora, E. (2004). Nurses' reactions to alarms in a neonatal intensive care unit. *Cognition, Technology & Work*, *6*(4), 239–246. https://doi.org/10.1007/s10111-004-0162-2.

Bliss, J. P., Fallon, C. K., & Nica, N. (2007). The role of alarm signal duration as a cue for alarm validity. *Applied Ergonomics*, *38*(2), 191–199. https://doi.org/10.1016/j.apergo.2006.03.004.

Block, F. E., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing*, *15*(2), 75–83. https://doi.org/10.1023/A:1009992830942.

Bustamante, E. A., Bliss, J. P., & Anderson, B. L. (2007). Effects of varying the threshold of alarm systems and workload on human performance. *Ergonomics*, *50*(7), 1127–1147. https://doi.org/10.1080/00140130701237345.

Cvach, M. (2012, July). Monitor alarm fatigue: An integrative review. *Biomedical Instrumentation and Technology*, *46*(4). https://doi.org/10.2345/0899-8205-46.4.268.

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*, *124*(3), 367–372. https://doi.org/10.1016/j.cognition.2012.05.012.

Dehais, F, Roy, R. N., & Scannella, S. (2019). Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural Brain Research*, *360*, 51–59. https://doi.org/10.1016/j.bbr.2018.11.045.

Dehais, Frédéric, Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness. *Human Factors*, *56*(4), 631–644. https://doi.org/10.1177/0018720813510735.

Edworthy, J., & Adams, A. S. (1996). *Warning design : a research prospective*. Taylor & Francis.

Fabey, M. (2009). Analysis shows pilots often ignore Boeing 737 cockpit alarm. Retrieved from Travel Weekly.

Fagerlönn, J. (2011). Urgent alarms in trucks: Effects on annoyance and subsequent driving performance. *IET Intelligent Transport Systems*, *5*(4), 252–258. https://doi.org/10.1049/iet-its.2010.0165.

Roth, E., & Multer, J. (2007). Technology Implications of a Cognitive Task Analysis for Locomotive Engineers. Technical Report No. DOT/FRA/ORD-09/03, Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Federal Aviation Administration. (2006). *Airworthiness Directives; Boeing Model 737 Airplanes*. Washington, DC: U.S. Department of Transportation, Federal Register.

Fenn, K. M., Shintel, H., Atkins, A. S., Skipper, J. I., Bond, V. C., & Nusbaum, H. C. (2011). When less is heard than meets the ear: change deafness in a telephone conversation. *Quarterly Journal of Experimental Psychology (2006)*, *64*(7), 1442–1456. https://doi.org/10.1080/17470218.2011.570353.

Görges, M., Markewitz, B. A., & Westenskow, D. R. (2009). Improving alarm performance in the medical intensive care unit using delays and clinical context. *Anesthesia and Analgesia*, *108*(5), 1546–1552. https://doi.org/10.1213/ane.0b013e31819bdfbb.

Haas, E. C., & Casali, J. G. (1993). Perceived urgency and detection time of multi-tone and frequency-modulated warning signals. In *Proceedings of the Human Factors and Ergonomics Society* (Vol. 1, pp. 544–548). Published by Human Factors and Ergonomics Society, Inc. https://doi.org/10.1177/154193129303700906.

Harper, C. D., Hendrickson, C. T., & Samaras, C. (2016). Cost and benefit estimates of partially-automated vehicle collision avoidance technologies. *Accident Analysis and Prevention*, *95*(2016), 104–115. https://doi.org/10.1016/j.aap.2016.06.017

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Jones, D. M., & Vachon, F. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 539–553. https://doi.org/10.1037/a0029064.

Kobayashi, L., Gosbee, J. W., & Merck, D. L. (2017). Development and Application of a Clinical Microsystem Simulation Methodology for Human Factors-Based Research of Alarm Fatigue. *Health Environments Research and Design Journal*, *10*(4), 91–104. https://doi.org/10.1177/1937586716673829.

Koreimann, S., Gula, B., & Vitouch, O. (2014). Inattentional deafness in music. *Psychological Research*, *78*(3), 304–312. https://doi.org/10.1007/s00426-014-0552-x.

Korniewicz, D. M., Clark, T., & David, Y. (2008). A national online survey on the effectiveness of clinical alarms. *American Journal of Critical Care*, *17*(1), 36–41.

Lawless, S. T. (1994). Crying wolf: False alarms in a pediatric intensive care unit. *Critical Care Medicine*, *22*(6), 981–985. https://doi.org/10.1097/00003246-199406000-00017.

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). *Human Factors Issues Associated with Limited Ability Autonomous Driving Systems: Drivers' Allocation of Visual Attention to the Forward Roadway*. Bolton Landing, NY: Public Policy Center. https://doi.org/10.17077/drivingassessment.1472

Mackworth, J. F. (1969). *Vigilance and Habituation: A Neuropsychological Approach*. Baltimore, MD: Penguin.

Mumaw, R. J. (2017). Analysis of alerting system failures in commercial aviation accidents. In *Proceedings of the Human Factors and Ergonomics Society*, *61*(1), 110–114. Human Factors an Ergonomics Society Inc.

Parasuraman, R., & Molloy, R. (2009). The International Journal of Aviation Psychology Performance Consequences of Automation-Induced Complacency. *The International Journal of Aviation Psychology*, *3*(1), 1–23. https://doi.org/10.1207/s15327108ijap0301.

Pasquini, A., Rizzo, A., & Save, L. (2004). A methodology for the analysis of SPAD. *Safety Science*, *42*(5), 437–455). https://doi.org/10.1016/j.ssci.2003.09.010.

Patterson, R. D. (1982). Guidelines for Auditory Warning Systems on Civil Aircraft. https://doi.org/10.1001/archderm.143.8.1080.

Rankin, C. H., Abrams, T., Barry, R. J., et al. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012.

Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *7*(2), 59–76. https://doi.org/10.1016/j.trf.2004.02.001.

Sheridan, T. B., Gamst, F. C., & Harvey, R. A. (1999). *White Paper: 'Reliance' and 'Distraction' Effects in PTC Automation*, Report No. STD-SCD-DEC-02-00, Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Ulfvengren, P. (2000). Natural warning sounds in safety critical human-machine systems - A cognitive engineering approach. In *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Association, "Ergonomics for the New Millennium"* (pp. 742–745). https://doi.org/10.1177/154193120004402264.

Warner, H. W., & Åberg, L. (2008). The long-term effects of an ISA speed-warning device on drivers' speeding behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*(2), 96–107. https://doi.org/10.1016/j.trf.2007.08.002.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd Edition). Pearson Prentice Hall.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2004). *An introduction to human factors engineering*. Pearson Prentice Hall.

Wiener, E. L., & Curry, R. E. (1980). Automation in the Cockpit: Some Generalizations. *Proceedings of the Human Factors Society Annual Meeting*, *24*(1), 38–38. https://doi.org/10.1177/107118138002400111.

Wogalter, M. S., & Silver, N. C. (1995). Warning signal words: Connoted strength and understandability by children, elders, and non-native English speakers. *Ergonomics*, *38*(11), 2188–2206. https://doi.org/10.1080/00140139508925262.

Wogalter, M. S., & Vigilante, W. J. (2006). Attention switch and maintenance. In *Handbook of Warnings* (M.S. Wogal, p. 841). Lawrence Erlbaum Associates.

Wreathall, J., Woods, D. D., Bing, A. J., & Christoffersen, K. (2007). *Relative Risk of Workload Transitions in Positive Train Control*, Technical Report No. DOT/FRA/ORD-07/12, Washington, DC: U.S. Department of Transportation, Federal Railroad Administration.

Young, K. L., Regan, M. A., Triggs, T. J., Stephan, K., Mitsopoulos-Rubens, E., & Tomasevic, N. (2008). Field operational test of a seatbelt reminder system: Effects on driver behaviour and acceptance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*(6), 434–444. https://doi.org/10.1016/j.trf.2008.04.004.

Appendix A.
Screening Questions and Exit Questionnaire

Demographic Questionnaire

To be conducted morning of study:

1. What is your name?
2. What is your age?
3. What is your race or ethnicity?
4. What is your sex?
5. How many years of experience do you have as locomotive engineer?
6. Do you drive passenger or freight trains?
    a. How many years of each if both?
7. How many years of PTC exposure do you have?
8. How many years of TO exposure do you have?
9. What is your usual route?
10. What is the typical duration of your shifts?
11. When was the last time you operated a train?

Exit Questionnaire

1. How similar did you find this to a real workday?
    a. If not similar: Why? What are the limitations of the simulator? Please note any unrealistic aspects of the simulation
2. How does the duration compare?
3. How does signal workload compare?
4. How fatigued did you feel at various points in the study?

Abbreviations and Acronyms

ACRONYM	DEFINITION
CTIL	Cab Technology Integration Laboratory
FRA	Federal Railroad Administration
GCOR	General Code of Operating Rules
KEA	KEA Technologies Inc.
MP	Mile Post
NORAC	Northern Operating Rules Advisory Committee
PTC	Positive Train Control
TSR	Temporary Speed Restriction
TSRB	Temporary Speed Restriction Bulletin