Acknowledgements

This thesis was part of a broader joint project between General Electric and the Massachusetts Institute of Technology and was funded by the Federal Railroad Administration (FRA) Office of Research and Development. The experiment was conducted in a locomotive simulator at the Volpe National Transportation Systems Center in Cambridge, MA.

I would like to thank the members of the GE/MIT project team, as well as our collaborators at the Volpe Center, and acknowledge their individual contributions. James Brooks, the project Principal Investigator at GE, was in charge of all phases of the project and implemented both current and advanced versions of GE’s Trip Optimizer in the locomotive simulator. He also handled data processing and generally assisted with many parts of the project. Bradford Miller (GE) helped in developing the task hierarchy model and took the lead on computing the static workload metric. Neeraja Subrahmaniy (GE) also played a significant role in developing the task model. Hullas Sehgal (GE) ran simulations to determine the simulated train makeup and helped implement a new pacing function for the advanced Trip Optimizer prototype. Paul Houpt (GE) contributed his knowledge of Trip Optimizer and rail operations to the task model development. Robert Repola, a retired locomotive engineer, also contributed his expert knowledge on the role of the engineer and conductor to the development of the task model. From Volpe, Matthew Isaacs was the key individual who ensured the experiment ran smoothly and successfully. Amanda Mattson and Bianca Mejia also assisted in the experiment. I thank my thesis advisor Charles Oman, Alan Natapoff, and Andrew Liu, all at MIT, for their input and guidance. Chuck and Andy helped conceive the experiment. Alan guided the statistical analysis on the workload data. Andy held responsibility for detailed experiment design, developed a conductor display, and assisted with data analysis paths. I contributed to the concept mapping and task modelling, developed heuristics to predict engineer workload along the experimental route, developed the conductor and dispatcher scripts for the experiment, served as cab conductor or dispatcher during experiment runs, and performed the regression based evaluation of the static workload metric using secondary task data. I thank everyone for making the project so successful. Additional details on the entire project are available in the project final report to FRA.

Thank you also to our eleven professional locomotive engineer subjects who participated in the experiment and their Class I railroad companies that supported their time and travel to Cambridge, MA. I extend acknowledgement to the CORYS Inc. manufacturer that developed the simulator and helped implement and troubleshoot the experimental versions of Trip Optimizer.

Finally, I thank the DOT FRA for project financial support through contract DTFR53014-C-00009, and particularly FRA’s Mike Jones for his visionary development of the Cab Technology Integration Laboratory (CTIL) locomotive simulator at the Volpe Center.

As a disclaimer, the opinions contained in this thesis are my own and do not necessarily reflect those of the FRA or GE.
Abstract

As technological advancement continues and the usage of automation in the cab of a locomotive grows, the role of the locomotive engineer and conductor changes. There is a need then to better understand the impact of different task allocations, or configurations of tasks, among the human and automation agents.

The technical objective of this thesis was to investigate whether the workload of the locomotive engineer could be predicted from a task model of the role of the engineer and conductor and their interactions with automation in rail operation. A task model was developed by reviewing existing Cognitive Task Analyses, and also by using concept mapping to identify information gaps which were then filled via expert interviews. Engineer and conductor goals and priorities were formalized in an Abstraction Hierarchy. This included controlling train speed, train forces, and dispatcher communications, while maintaining awareness of train and signal states, track conditions, and nearby traffic. Hierarchical Task Analysis was used to identify the lower level tasks of the engineer and conductor when driving manually in traditional fashion, or when using two alternative levels of throttle, brake, horn, in cab signaling and pacing automation. Engineer mental workload was predicted based on the number of information inputs and actions required, and an analysis of when the engineer would be engaged in the specific tasks while driving a route. To validate the task model based workload prediction, a human-in-the-loop experiment was conducted in a DOT locomotive simulator. Eleven subjects drove the same route manually and under the two automation conditions. Response time to a visual secondary task was used as a proxy measure for mental workload. Hierarchical mixed regression analysis was performed to compare the secondary task response time with the workload predications from the task model, which incorporated the three automation effects. Response time correlated significantly with the modeled workload. The ratio of experimental response time to predicted workload depended significantly on subject, automation condition, and along track distance effects. The task model slightly (2%) under/overestimated mean effects of automation and average response time. However other un-modeled effects also contributed to the regression residual, as discussed.

A second objective of this thesis was to consider whether the qualitative information contained in the concept maps, CTA analyses, and formal task models on the role of the engineer and conductor and their interaction with automation could be useful in policy discussions. Several potential benefits of using this information were identified: it can educate people on the job of the engineer and conductor in rail operations, it can be a shared language for stakeholders to use in discussion of the future of rail automation, and it can promote knowledge translation between researchers and policymakers. However, several steps need to be taken to get the information from the task model into an accessible form for use in policy discussions. These include creating a visual and interactive tool to display the information and involving stakeholders in the development process.

Supported by DOT FRA contract DTFR53014-C-00009.
# Table of Contents

List of Figures ......................................................................................................................... 6
List of Tables .......................................................................................................................... 8

1. Introduction .................................................................................................................... 9
   1.1. Thesis Overview ...................................................................................................... 11

2. The Task Model .............................................................................................................. 14
   2.1. Background on Task Analysis Techniques ............................................................ 14
   2.2. Synthesizing Information from CTA Reports Through Concept Maps ................. 15
   2.3. Task Model on the Role of the Engineer and Conductor in Rail Operations ........... 20
   2.4. Developing Heuristics to Relate Static Workload Prediction to Tasks Along a Route 24
   2.5. Using the Task Model for Automation System Prototype Development ............... 27

3. Validation of Task Model Workload Prediction Through a Simulator Experiment ...... 29
   3.1. Previous Rail Human Factors Studies ................................................................... 29
   3.2. Methods .................................................................................................................. 31
       3.2.1. Locomotive Simulator ........................................................................................ 31
       3.2.2. Human Subjects ................................................................................................ 33
       3.2.3. Automation Conditions ...................................................................................... 33
       3.2.4. Secondary Task Workload Measurement .......................................................... 38
       3.2.5. Experiment Scenarios ........................................................................................ 39
       3.2.6. CTIL Simulator Cab Configuration .................................................................... 42
       3.2.7. Subject Training and Experiment Schedule ....................................................... 43
       3.2.8. Experiment Design ............................................................................................ 44
       3.2.9. Statistical Analysis Methods .............................................................................. 44
   3.3. Experiment Results ............................................................................................... 45
       3.3.1. Data Processing .................................................................................................. 45
       3.3.2. Establishing Sample Independence .................................................................... 46
       3.3.3. Hierarchical Mixed Regression Result ............................................................... 48
   3.4. Discussion ................................................................................................................. 54

4. Utilizing Human Factors Work in the Policy Sphere ..................................................... 61
   4.1. Key Events in Rail Automation Discussion .............................................................. 61
   4.2. Can the qualitative information from task modeling be useful in the policy sphere? 65
   4.3. What steps need to be taken to make the task model information accessible to policy makers? ........................................................................................................................................ 66
   4.4. Limitations and Conclusion .................................................................................... 69

5. Conclusion ...................................................................................................................... 71

References ............................................................................................................................ 73

Appendix A. Concept Maps for the Conductor and Dispatcher ........................................... 77
Appendix B. Route Timetable ............................................................................................... 79
Appendix C. Scenario Summaries .................................................................81
Appendix D. Conductor and Dispatcher Script for Manual Scenario ...............83
response time, and the static model workload prediction (shown on the right y-axis), for the TO condition of Subject 4, there is still significant autocorrelation at lag 4 in the TO condition.

Autocorrelation in the response time data falls and remains below the 95% confidence interval for Subject 10 at lag= 4 in the TO condition.

No signification autocorrelation is seen in the secondary task response time data for the TO+ condition of Subject 6.

Autocorrelation in the response time data falls and remains below the 95% confidence interval for Subject 10 at lag= 4 in the TO condition.

For the TO condition of Subject 4, there is still significant autocorrelation at lag 18 of the secondary task response time data.

Distribution of static workload prediction data.

Distribution of secondary task response time data.

Distribution of the ratio of the secondary task response time data to the static workload prediction data.

The residuals of the regression fit and a histogram of the residuals plotted against the predicted value of the log(response time/static model).

For the Manual condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment.

For the TO condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment. Also shown at the top are the time intervals in auto mode.

For the TO+ condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment.

Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the Manual condition.

Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the TO condition.
Figure 29. Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the TO+ condition.
List of Tables

Table 1. Example of mapping model tasks to the experiment scenarios. The start of the TO+ trip is shown with the mapping to two operational scenarios (M1 and M2)..................................... 25
Table 2. With the log(response time/static model) values in ascending order, the grey rows show the points on the extreme ends that were excluded from the regression model. ....................... 51
Table 3. Number of data points for each subject and condition after down-sampling and removing five additional outliers. ........................................................................................................ 52
Table 4. Effect coefficients for each subject............................................................................ 52
Table 5. Coefficient, standard error, and p-value for all the effects in the regression model..... 53
Table 6. Results of the tests for normally distributed residuals and equal variance across the data. ........................................................................................................................................ 54
1. Introduction

The rail industry has seen numerous changes over the course of its history. In the days of steam engines, operating a train required 5-7 men. These include the engineer, whose primary responsibility is to drive the train; the conductor, who handles communication, paperwork, and is responsible for the entire train ("consist"); the fireman, who stoked the fire and maintained steam pressure; and the brakemen, who climbed from car to car activating hand brakes and jumped off to throw manual switches on the tracks (Sperandeo and Keefe 2006). With the adoption of the diesel locomotive in the mid-1900’s, the train crew was reduced to only two people—the engineer and the conductor. The labor unions did not quietly watch the reduction in crew size brought on by advances in locomotive technology. They pushed for the continued employment of firemen even after the steam engine was phased out, as engineers tended to come from their ranks. However, the Chicago Tribune reported in 1985 that a presidential mediation board deemed the job of a fireman “no longer needed for the safe operation of diesel trains” (United Press International 1985).

Crew size is not the only change the rail industry has seen. Throughout the history of the railroad, the controls and technology that the engineer interacts with have changed drastically as well. What started as a job with virtually zero automation and used hand signals and lanterns for communication has morphed into a highly complex system involving various automation systems connecting the engineer and the locomotive to dispatch centers across the country. Sensors on various parts of the train provide information and diagnostics to the engineer on a digital display in the cab. Automation systems are in place, like General Electric’s Trip Optimizer (which will be discussed in detail later in this thesis), that control the throttle and dynamic brake automatically. Safety systems, like Positive Train Control, offer a redundancy that helps to ensure the signals and speeds are followed. The technology used to control trains is only going to continue to advance, and likely at faster rates, in the coming years.

Automated systems are prevalent in domains outside of rail as well. Aviation and space flight have seen advances and increases in automation to perform tasks previously done by human operators. Automation has the potential to ease the workload on the human operator, freeing up capacity to do other tasks. The precision of automated systems can increase safety in situations that could be prone to small errors (Wiener and Curry 1980).
However, automation can come at a cost. While automation systems are meant to decrease overwhelmingly high workload that can hinder performance, low workload situations, with low “arousal,” can contribute to decreased performance as well (Yerkes and Dodson 1908). Automation can also negatively impact an operator’s situation awareness, making it difficult to respond to emergency situations (Wickens 2008). When the human operator acts as a monitor or supervisor of automation systems, there is concern that the operator can become complacent, especially when the systems have consistent reliability. Complacency leads to an inability to identify when the automation systems are malfunctioning and take over control of the system if necessary (Parasuraman et al. 1993). This kind of out-of-the-loop performance decrement, when the operator is removed from direct control over the system, indicates that tasks are often allocated inappropriately between humans and automation systems. Again, automation has many benefits, but the human operator is a part of the system too. The strengths of both the human and the automation can be utilized if the tasks are allocated according to the capabilities of each (Kaber and Endsley 1997).

In the rail domain, the engineer and conductor are still very much a part of operating a train. With the addition of automation systems in the cab, and considering the potential disadvantages of human-automation interaction, natural questions emerge about the role of the engineer and conductor in interacting with these systems. Broadly, this thesis considers how the role of the engineer and conductor changes with the addition of automation systems and looks at how tasks can be assigned to each operator and an automation system to make the whole system run more efficiently (e.g. fuel savings or trip time) and effectively, as in the skills particular to humans and computers are utilized.

One way to study human-automation interaction is through an experiment using human subjects. However, these experiments are costly and time consuming. There is a need then to develop models of human-automation interaction that are able to predict outcomes and identify changes without having to run a full-fledged experiment.

One approach to modeling human-automation interaction is described in So Young Kim’s PhD work at the Georgia Institute of Technology. Kim focuses exclusively on function allocation which she defines as “the design decision which assigns function to all agents in a team, both human and automated.” Then “if functions are allocated properly, function allocation maximizes mission performance by best utilizing the capabilities of each agent and provides the
environment that fosters their individual performance by promoting effective interactions within the team” (Kim 2011, 1). The issue with function allocation then is determining if and when functions are allocated “properly.” Kim acknowledges that there can be incoherency in function allocation and vagueness in the guidelines about function allocation typical in the literature. To overcome these issues, Kim presents her own guidelines for creating function allocations and metrics to assess different function allocation models.

A function allocation model must first start by analyzing the team and the tasks each agent contributes to the work. Kim points to Cognitive Work Analysis methods as a way to “identify hidden, complex relationships between goals, functions, information requirements, work environment, and agents” (48). Specifically, these methods include Abstraction Hierarchy, Strategy Analysis, and Social Organization and Cooperation Analysis. Once a function allocation is established, Kim recommends using metrics to assess the effectiveness of the allocation. Two of the metrics highlighted for assessing different function allocations are workload and coherency. Kim suggests looking at the changes in operator workload, also called their taskload, by assigning values to tasks and seeing if the taskload estimated by the model exceeds the operator’s maximum capacity (69). The other metric, coherency, looks at how connected (coherent) or disparate (incoherent) the different tasks are for each agent. In a hierarchical model, this could be measured by looking at whether or not all the tasks for one agent are contained in one branch of the hierarchy. Additionally, identifying task overlap between agents could indicate incoherency.

1.1. Thesis Overview

It is this recent work in modeling human-automation interactions by So Young Kim – now at General Electric (GE) - that motivated the broader project, a collaboration between GE and the Massachusetts Institute of Technology (MIT), of which this thesis research was a part. The project as a whole aimed to provide a methodology for designing and evaluating different function allocations, or configurations, of human and automation system roles in the rail industry. The main research question is: can human and system performance be predicted from a task model that defines the role of the engineer and conductor under various automation conditions? The aspect of this question addressed in this thesis relates to the workload of the engineer. Specifically, the technical aim of this thesis asks, can the workload of the locomotive
engineer, under different automation conditions, be predicted from the task model developed in the GE/MIT project? While this technical aim deals with the computational prediction aspect of the task model, a second aim of this thesis is to look at the utility of the qualitative information contained in the task model on the role of the engineer and conductor and their interactions with automation. Specifically, is the qualitative information contained in the task model useful for policy discussions? And if so, what steps need to be taken to get the information into an accessible format for use in the policy sphere?

In order to answer these questions, following this introduction, the second chapter of this thesis outlines the task modeling work. The final task model is a combination of an Abstraction Hierarchy and a Hierarchical Task Analysis, both common human factors analysis techniques that will be introduced in the chapter, and details the tasks performed by the locomotive engineer and the conductor in typical scenarios encountered while operating a train. From Kim’s work mentioned above, the metric of workload was applied to evaluate the function allocations assigned in the task model.

In order to validate the task model and see if it appropriately predicted the workload of the engineer, a human-in-the-loop experiment was conducted at a locomotive simulator at the Volpe Transportation Systems Center in Cambridge, MA. This experiment, detailed in Chapter 3, used a secondary task to measure the workload of the locomotive engineer under three different task allocations. These allocations manifested in the form of three different automation conditions. In manual condition, the engineer subjects drove the train as they would in normal operation. The second condition used the automation system Trip Optimizer (TO), a General Electric product in wide use throughout the freight rail industry. The third condition used an automation system called Trip Optimizer Plus (TO+), a prototype system designed specifically for this project. Regression analysis was then used to compare the workload prediction from the task model with the secondary task data from the experiment.

Finally, Chapter 4 of this thesis circles back to the task model introduced in Chapter 2 and addresses the policy implications of the qualitative information on the role of the engineer, conductor, and automation in rail operations. This chapter establishes the relevance of the task model information and provides some of the potential benefits of using this task information in the policy sphere. Additionally, it highlights the steps necessary to get the task model
information into an accessible form that can be used by various stakeholders involved in the rail automation discussion.
2. The Task Model

The development of a task model that describes the role of the engineer and conductor in rail operations and their interactions with automation is the foundational work set forth in this chapter. The goal of developing this model was to use it to predict changes in human and system performance as the agents (engineer, conductor, and automation) are assigned different tasks during operation. Again, this thesis focuses on one aspect of human performance: changes in the engineer’s workload.

The first step in the modeling work was to distill the previous rail domain work from several Cognitive Task Analyses (CTA) into concept maps to visually represent the tasks and their linkages. Information gaps were identified and added to the concept maps through expert interviews. From there, the techniques of Abstraction Hierarchy and Hierarchical Task Analysis were used in conjunction to formally capture the goals, purposes, and low-level tasks of driving a train. As it stands, the model details the tasks performed by the engineer and the conductor in four different operational scenarios: 1) Managing signal response, 2) Managing temporary speed restrictions, 3) Managing maintenance of way interactions, and 4) Managing interactions at grade crossings. Based on the information inputs and outputs required for each task, a workload value can be assigned. In order to relate the four operational scenarios to a particular trip driven by an engineer, heuristics were developed that identify when the engineer is engaged in particular tasks. This allows for a workload prediction in tenth of a mile segments throughout a trip. Finally, the model was used to allocate various tasks to an automation system and predict the changes in workload experienced by the engineer and conductor as their roles change.

2.1. Background on Task Analysis Techniques

Before delving into the specifics of the task model developed for this project, it is necessary to define some of the techniques available to model tasks in the Human Factors domain. While this is by no means an exhaustive list, the techniques covered here are: Hierarchical Task Analysis, Cognitive Task Analysis, and Abstraction Hierarchy. First, Hierarchical Task Analysis (HTA) “involves describing the activity under analysis in terms of a hierarchy of goals, sub-goals, operations, and plans” resulting in “an exhaustive description of the task activity” (Stanton 2013, 40). HTA focuses on the goals and the physical tasks required in
the activity to achieve those goals. The form of an HTA is typically a straightforward hierarchy or can be represented in tabular format, but for large and complex systems, both the hierarchy and tabular format can become difficult to read and comprehend (Stanton 2013). Another disadvantage of HTA is in reliability as different analysts can create entirely different hierarchies for the same task.

While HTA focuses primarily on physical tasks that are part of an activity, Cognitive Task Analysis (CTA) is “used to determine and describe the cognitive processes used by agents” (Stanton 2013, 69) and focuses on knowledge and thought processes. Similar to HTA, CTA is goal-based and attempts to identify the tasks that an individual performs to accomplish a stated goal. However, the purpose of CTA is to better understand the cognitive role individuals play in the system.

Abstraction Hierarchy (AH) is a framework typically used within CTA to get “a context-independent description of the domain” (Stanton 2013, 74). AH helps identify the important high level system functions within a work domain and the constraints and dependencies within the functions. Typical abstraction hierarchies consist of the following levels: 1) high level functional goals and purposes, 2) abstract functions or the priorities, constraints, and values that influence the ways in which the high level goals are achieved, 3) generalized functions that are executed within the realm of the priorities and constraints, and 4) the specific physical functions and features in the domain that help realize the generalized functions (Vincente 1999). At the top three levels, an AH is not specific to any particular activity or task within a domain, but rather offers a description of the “overall objectives of the domain.” The form of an abstraction hierarchy is often a complex hierarchical network, which can be difficult to comprehend for non-experts in the domain.

Each of these techniques—HTA, CTA, and AH—were employed in some form in the course of the GE/MIT project and the following sections detail their use.

2.2. Synthesizing Information from CTA Reports Through Concept Maps

Three CTA reports commissioned by the Federal Railroad Administration (FRA) were identified as the prior work in the rail domain: 1) A CTA on technology implications for locomotive engineers (Roth and Multer 2007), 2) A CTA on the cognitive and collaborative demands on conductors (Rosenhand et al. 2012), and 3) A CTA on how dispatchers work (Roth
et al. 2001). These CTA reports were reviewed in detail in order to synthesize and better understand the roles and responsibilities of the locomotive engineers, conductors, and dispatchers; and identify some of the knowledge gaps in the CTA literature.

While the CTA reports were fairly comprehensive, the written rather than graphical format made it more difficult to immediately see the linkages and relationships between each agent's tasks and how the agents interacted with each other. So, the technique of concept mapping was used to visually represent, record, and synthesize information on the roles and responsibilities of the train crew as discussed in the CTA reports. Concept maps are flexible, free-flowing charts that help to organize knowledge and information to describe concepts within a particular domain, and have been used successfully by research administrations such as NASA to capture knowledge within complex work domains (Coffey and Carnot 2003). In a concept map, the concepts are represented through nodes, and the relationships between concepts are linked through arrows. One of the advantages of concept mapping is that it is an easy and effective way to visually conceptualize knowledge, and understand the missing "links", relationships, and the details that refine a particular sub-concept. It is also a simple and cost-effective way to visually organize information from narratives and reports, so that key ideas and insights are quickly communicated and internalized within a team (Crandall et al. 2006). The tool used to construct these concept maps was IHMC CmapTools, a software tool developed by the Institute of Human and Machine Cognition. Using the results and information reviewed in the three CTA reports previously mentioned, concept maps detailing the roles and responsibilities of a locomotive engineer, conductor, and dispatcher were created (Cañas et al. 2004).

Members of the GE/MIT project team created the concept maps independently and then completed several iterations of comparing the maps and discussing the similarities, differences, and necessary additions. All the information was then combined into a single concept map for each operator and independently checked for accuracy and completeness of information. Finally, the concept maps were sent out to the rest of the team, including one industry expert (a retired locomotive engineer) to gather feedback. Expert feedback helped both refine and validate the developed concept maps. Based on the feedback and a brief review of some supplemental reports, several concepts were added to complete the map. Figure 1 shows the Concept Map for the locomotive engineer and to clarify the notion, Figure 2 shows a detailed view of a portion of this map. The concept nodes with a white background are those from the CTA reports, blues
nodes are additions from an industry expert, and gray nodes are concepts from supplemental reports. This map extends to include the macro-cognitive processes and tasks that the engineer needs to engage in, and the specific skills, strategies, and tools needed to accomplish the roles and execute the cognitive processes. Therefore, the concepts at the top part of the map mostly detail the roles and cognitive processes, while the bottom part of the map break down into the detailed cues, tools, and strategies used. See Appendix A for the conductor and dispatcher concept maps developed.

Figure 1. Concept map for the locomotive engineer's roles and responsibilities.
The level of detail in each of the concept maps helped to identify some significant gaps in the CTA reports as it illuminated missing links or areas that needed further clarification. For the locomotive engineer, details on the specific strategies used to slow down and brake the train were not included in the CTA report. Likewise, the CTA report did not detail the specific attentional strategies (visual, auditory, etc.) that engineers use to balance their monitoring demands. The engineer and conductor are often referred to as a “joint cognitive system” in the CTA reports. However, the reports lack detailed information on the collaborative strategies and assistance that each provides the other during key operating scenarios. The final gap that the concept map exercise illuminated was that although driving the train efficiently was identified as an important goal within the rail industry, it is not necessarily part an individual engineer’s goal or motivation. The motivations and goals of the engineer, as distinguished from management, are another set of details absent from the CTA report.

Just as with the engineer map, the conductor concept map helped illuminate the gaps in the CTA report. For instance, the CTA stated that conductors use information about the train consist to understand train dynamics and thus monitor engineer performance. However, the CTA did not give any further details about how or when this monitoring takes place. The CTA also did
not provide specific example situations to illustrate this process. Additionally, the existing CTA had few details about what the conductor specifically needs to do in various unexpected situations. Finally, the CTA did not explicitly state the work conditions for the conductor in the cab. A single gap in the dispatcher CTA was also identified. The CTA report discusses the different ways in which Data Link technology and sophisticated and detailed visualizations on computer displays can help dispatchers plan and project into the future. However, the CTA report did not specify the different kinds of communication and proactive suggestions the dispatchers could provide to the engineers in order to help with the engineer’s goal to anticipate and plan ahead.

This exercise illustrated a key difference between the locomotive engineer’s role and that of the conductor. While the engineer’s roles and responsibilities are all intricately interconnected, the conductor has more specific, discrete tasks to complete. Many of the concepts on the engineer concept map network are all interconnected and contribute to the overall goal of driving the train. However, on the conductor concept map, there are numerous discrete tasks that branch out directly from the main conductor node. This is an interesting difference to note between the role of the engineer and conductor in rail operations because this relates to how tasks are currently distributed between the two, and the difference in the nature of the tasks might suggest imbalance in workload and the opportunities to introduce automation.

In summary, concept maps can be a useful way to capture and represent knowledge. They are one tool within the conceptual framework repository to organize and synthesize information. They served well in their purpose of visually representing and organizing knowledge in a quick and efficient manner by distilling content from lengthy reports. However, one of the limitations with representing information in a concept map is that it does not represent the decisions that the conductors/engineers/dispatchers have to make and the actions that result from those decisions. The free-flowing nature of the concept maps makes formal representation of sequential, temporal, or conditional information and statements (that deal with if-then-else clauses) difficult to represent. Concept maps also do not lend themselves to indicate associated workload demands for the cognitive processes and tasks represented, nor do they capture those tasks/functions that need to be performed in parallel or separately.
2.3. Task Model on the Role of the Engineer and Conductor in Rail Operations

In order to address some of the limitations with concept maps and traditional CTAs and more formally capture the goals and purposes of the rail domain as a whole, an abstraction hierarchy was created with the ultimate goal of predicting the workload of the engineer. To reiterate from Section 2.1, the abstraction hierarchy framework helps identify the important high level system functions within a work domain and the constraints and dependencies with the functions. It is not meant to be a description of specific tasks in the domain, but rather to offer the overall objectives.

The first level of the abstraction hierarchy, the high level functional goals and purposes of the rail domain, was constructed directly from the main responsibilities of the engineer found through the concept map exercise. These are 1) safe operation of the train and 2) efficient operation of the train. The concept maps also revealed the main priorities and constraints that form the second level of the abstraction hierarchy. These include maintaining awareness of:

- Train state
- Train rules and regulations
- Train location and extent
- Track conditions
- Topography
- Long term/intermediate train handling plans
- New slow orders
- Overall integrity of the system
- Safe speed
- Nearby traffic
- State and limitation of the automation system

The generalized functions that form the third level of the abstraction hierarchy are the functions that help satisfy the priorities outlined the level above it. So in order to maintain awareness of abstract functions in the second level, the follow aspects need to be managed:

- Train Speed
- Braking Systems
- Train Forces
- Signal Violation
- Signal Uncertainty
- Dispatcher Communication
- Crew Agreement
- Stability of Work Environment
- Train Location Uncertainty
- Work Location Uncertainty
- Track Encroachment Uncertainty
- Mode of Automation System
- Verbal Records
- Warning Systems
- Track Defects Detection
- Maintenance Crew Communication

Typically, at this point in an abstraction hierarchy, the fourth level lists the physical functions or objects that contribute to realizing the generalized functions in the level above. However, the goal of the model in the context of this project is to get down to the low level physical tasks that an engineer performs in certain scenarios in order to come up with a workload prediction. To generate these low level tasks, the method of hierarchical task analysis was employed. Again, this method involves decomposing a task pertaining to a particular operational scenario into a hierarchy of sub-tasks. The analysis focused on the following four operational scenarios as context for generating the lowest level tasks:

- M1. Managing Signal Response
- M2. Managing Temporary Speed Restrictions (TSR)
- M3. Managing Maintenance of Way (MOW) interaction
- M4. Managing Interaction at Grade Crossings

Figure 3 shows a portion of the abstraction hierarchy with some of the lowest level tasks identified from the hierarchical task analysis for the operational scenario of Managing Signal Response. Graphical representation of the model became quite difficult, so all the information on the different levels and their mapping to levels above and below are contained in tables. These tables also include information on which agent (engineer, conductor, or automation) is involved in a particular task. The abstraction hierarchy levels, in combination with the low level tasks
delineated through the hierarchical task analysis, constitute the "task model" as it will be called throughout this thesis.

The last step before obtaining a workload prediction was to detail the inputs (information needs) and outputs (products) for each task. For example, under the operating scenario of Managing Signal Response, if the task performed is to preemptively slow the train in order to better see the upcoming signal, the input or information need to this task is knowledge of the upcoming signal and the output of this task is to notch down and/or apply the brake. These inputs and outputs were counted for each task so in this example the input has a value of 1 and the output has a value of 2 because the engineer interacts with the throttle/brake on the control stand and the display screen in front. These inputs and outputs were then used to calculate the static metric of workload, static because it can be computed offline and in advance, without having to measure it dynamically based on unpredictable events occurring during an experiment.

The following figures provide an example of how this static workload is computed from the function allocation task model. Figure 4 shows a sample function allocation with four levels.

Figure 3. Graphical representation of the Abstraction Hierarchy of Managing Signal Response. Only some generalized functions and corresponding tasks are displayed to aid readability.
The first two levels, the abstract functions of agents and the generalized functions, would be from the abstraction hierarchy aspect of the model. The second two layers, leaf tasks and joint tasks that have been broken down into the parts performed by each agent, would be taken from the hierarchical task analysis exercise.

![Diagram](image)

**Figure 4. Example function allocation.**

The final two levels of the example function allocation were marked with the inputs required at each node, shown in lowercase letters in Figure 5. Then the static workload can be calculated for a particular agent by counting the unique inputs, adding one to that\(^1\), counting the unique outputs, adding one to that as well, and finally multiplying those two numbers together. For the engineer in this example, there are 11 unique inputs and the number of outputs (6) is the number of leaf tasks assigned to the engineer. This yields a total workload of 84 for the engineer. This value is not calibrated in any way and is intended more for comparison with other function allocations. The workload for the conductor and the automation in this particular configuration are shown in Figure 5 as well.

---

\(^1\) The addition of one to both the input and output count was made to boost the low input/output workload combinations to better relate them. This way the workload calculation for single input single output combinations are not deemed quite so low.
2.4. Developing Heuristics to Relate Static Workload Prediction to Tasks Along a Route

This section explains the method used to relate the four task model operational scenarios (M1-M4 above) to a particular trip an engineer might drive. The examples given are specific to the route used in the experiment that is described in Chapter 3 of this thesis, but the heuristics developed can be generically applied to any trip. First, all the trip events and their locations were delineated in a spreadsheet in tenth of a mile segments. These events included all the speed restriction areas, crossing gate locations, signal locations, MOW zones, and dispatcher interactions. With the consultation of a domain expert and knowledge of what generally happened in the experiment itself, each tenth of a mile segment was marked with an ‘x’ if the engineer was likely to be engaged in one of the four task model scenarios. The low level tasks identified in the hierarchical task analysis also contributed to the identification of locations when the engineer was engaged in particular tasks. An example of this mapping is shown in Table 1 for two of the task model operational scenarios.
Table 1. Example of mapping model tasks to the experiment scenarios. The start of the TO+ trip is shown with the mapping to two operational scenarios (M1 and M2).

<table>
<thead>
<tr>
<th>Milepost</th>
<th>Event</th>
<th>M1 (Managing Signal Response)</th>
<th>M2 (Managing Speed Restrictions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>77.9</td>
<td>Start of trip, Signal DIVERGING CLEAR, Turnout to Main 2</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.5</td>
<td>Grade crossing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>Dispatcher call to inform of en route TSR at MP89</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.1</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.2</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.3</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.4</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.5</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79.8</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>79.9</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>80.1</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>80.2</td>
<td>Signal, Warning Board 40mph</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>80.3</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>80.4</td>
<td></td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

The x’s were assigned according to certain guidelines about typical engineer behavior and these guidelines were used throughout the mapping exercise to maintain consistency. For the scenario of Managing Signal Response, the engineer begins performing the task 0.4 miles in advance of all the signals. This task continues until the locomotive passes the signal and the
engineer can no longer see the signal. For non-clear signals (anything other than a green signal), the engineer is engaged in the task from the location they see the signal to the location they confirm and either pass or stop at the next signal. This is because the engineer must approach the next signal at a certain speed or be prepared to stop if the previous signal was not clear. The distance that the engineer begins performing this task does depend on the sight distance. Generally, it is reasonable that the engineer can see that there is an upcoming signal 0.4 miles in advance in the simulator, but there are exceptions to this if the signal is around a corner, for example. In this mapping all the signals were treated the same.

For the scenario of managing speed restrictions, the permanent speed restrictions differed slightly from the slow orders received en-route. First, for the permanent speed restrictions, at the warning board (two miles before the start of the speed restriction) the engineer begins performing the task 0.1 miles in advance of the warning board and continues until 0.2 miles after the warning board, processing the information and planning ahead. One mile before the start of the speed restriction, the engineer engages in the task again, for 0.1 miles, in order to check speed and location. Half a mile before the start of the speed restriction the engineer engages in task and continues until 0.2 miles into the restricted area.

Second, for the en-route temporary speed restrictions, the engineer listens to the dispatcher call (and is thus engaged in the task) for 0.5 miles. Then the engineer engages in the task two miles in advance of the speed restriction, adjusting the speed and confirming the location at each milepost. Half a mile before the start of the speed restriction the engineer again engages in the task and continues 0.2 miles into the speed restriction zone.

This mapping is clearly far from a complete representation of the engineer’s behavior and driving tendencies. For this static model scenario an argument could be made that the engineer is doing the task of managing the speed restriction throughout the entire speed restriction area, not just at the start. Also, for the en-route temporary speed restriction, the mapping doesn’t very well represent the frequency that the engineer may be refreshing their memory on the location of the upcoming speed restriction. As it is, they receive the information and then “forget” it until 2 miles in advance of the restriction.

The third operational scenario of Managing Maintenance of Way Interaction utilized the following rules for mapping onto the events in the trip. The engineer asks the conductor to call the MOW foreman about 5 miles in advance of the known MOW area. This task takes 0.2 miles.
Then the engineer looks over the information for 0.5 miles. The engineer checks the speed and location of the locomotive, and makes appropriate adjustments, every mile in advance of the upcoming MOW area. Then the engineer is engaged in the task for the entire mile before the start of the MOW area, very aware of speed, location, and the distance counter. This continues until the head end makes it through the first quarter mile of the zone, as were the instructions given by the foreman.

Finally, for the fourth operational scenario of Managing Interactions at Grade Crossings, the mapping includes the engineer beginning the task 0.3 miles in advance of the crossing and continuing until the locomotive is through the crossing. When there is a faulty gate, the engineer continues to engage in the task for 0.2 miles after the crossing to ask the conductor to call the problem in to the dispatcher.

2.5. Using the Task Model for Automation System Prototype Development

Aside from using the task model to predict the workload for a given trip driven by an engineer, the task model was also used to develop features for a prototype automation system. For each of the four operational scenarios (M1-M4 listed above), the tasks were allocated among the agents as they are in typical operation today and a static metric of workload was calculated for the engineer and the conductor. This value represents the workload of the given operator over the entire scenario (e.g. the workload of the engineer for the Managing Signal Response scenario). This representation of the model then allowed the team to change the allocation of various tasks in each automation scenario and see the effect on workload in that static scenario. We used this method to assign tasks to a prototype automation system for use in the experiment, described in detail in the next chapter. The idea was to see how allocating a particular task or group of tasks to automation would change the workload of the engineer. In allocating tasks differently from normal operation, the model predicts a certain change in static workload. The goal then is to design an automation system with capabilities that can be tested in an experiment to validate the model and see if the model predicts the correct changes. For example, if the model predicts that allocating a certain task to automation decreases the workload of the engineer, that workload can be measured in an experiment and be compared to the model’s prediction. The following three automation capabilities were proposed and eventually incorporated into the
prototype automation system used in the experiment, and are described in greater detail in the following chapter.

The first capability was in-cab-signaling which involves displaying the upcoming signal information to the engineer in the cab before they can see it out the window. Generally, the static measure for workload makes conversations between the conductor and the engineer expensive. Based on this judgment, we noted that in-cab signals would remove the need to get agreement on the signal aspect as well as the need to predict when and where an upcoming signal may be visible, both distinct tasks in the model.

The second capability designed to change the workload of the engineer was called "pacing". Responding to non-clear signals is a large source of workload for the engineer. Pacing is a proposed new TO automation feature that would work to automatically slow down the train based on traffic ahead so that engineer always sees green signals.

The final capability incorporated was "autohorn". Without automation, the engineer must remember to blow the locomotive horn in a specific warning pattern (two long blasts, one short blast, and another long blast) when approaching a grade crossing. For the operational scenario of Managing Interactions at Grade Crossings, actuating the horn is a main source of workload for the engineer. Incorporating an automatic horn could decrease the workload of the engineer accordingly and provide another change in workload to detect in the experiment.
3. Validation of Task Model Workload Prediction Through a Simulator Experiment

A human-in-the-loop experiment was conducted in the locomotive simulator at the Volpe Transportation Systems Center in order to validate the model of the locomotive engineer’s tasks. The goal of this experiment was to measure human and system performance, particularly the workload of the engineer, under different task allocations in the cab and then compare the measured workload from the experiment with the predicted workload from the task model described in the previous chapter. These different allocations manifested themselves in three different automation conditions: 1) Manual operation with no automation, 2) Trip Optimizer- a current automation system used widely in the industry to optimize for fuel efficiency, and 3) Trip Optimizer Plus- a prototype automation system developed specifically for this experiment that adds features to Trip Optimizer. The workload during the experiment was measured using a secondary task. The response time from the secondary task was then compared to the workload prediction from the task model using regression analysis.

3.1. Previous Rail Human Factors Studies

While there are not many human-in-the-loop simulator studies pertaining to rail human factors, three particular studies pertaining to automation in the rail industry helped inform the design of the experiment. However, the motivation for the experiment differed from past experiments in that the goal here was to design the experiment to see how function allocation affected performance and workload in order to validate the model of engineer tasks. This is a slight, but important, difference to note when comparing the experiment to the past studies that looked specifically at how particular automation systems or levels of automation control affected the engineer’s performance and workload.

The first study, by Lanzilotta and Sheridan, and published in January 2004, looked at three different levels of supervisory control in the locomotive cab. They were interested in how supervisory control impacted the workload of the engineer and by extension, how the increased or decreased workload impacted their performance driving the train particularly for high speed passenger rail applications. The dependent variable in this experiment was response time and response accuracy to different emergency situations they encountered on the route. The three
emergency situations encountered were: brake failure, traction motor failure, and a grade crossing obstruction. Their results suggest that in high automation modes, the engineer tends to monitor the external environment more than in manual mode (Lanzilotta and Sheridan 2005).

The second study dealing with automation was done by Marinakos, Sheridan, and Multer published in July 2005. This experiment, also done with high speed rail applications in mind, was concerned with attention allocation strategies employed by engineers when using automation systems and the effect on vigilance and situation awareness. They also used three different supervisory control levels (manual, partial supervisory, and full supervisory) and had a series of events participants had to detect and respond to. Vigilance was measured by looking at the detection rate of equipment failures and obstructions built into the scenario. Situation awareness was measured using the Situation Awareness Global Assessment Technique (SAGAT) where the simulation is stopped, all screens are blanked, and the experimenter asks the subject questions about the current internal and external environment. In addition, after the completion of all three trips, they asked the subjects questions about their perceived workload using the different automation levels. The results of this experiment indicated that automation improved vigilance and situation awareness, but the authors acknowledge the complexities in attention allocation strategies that this experiment could not actually tease out (Marinakos et al. 2005).

The final, and most recent study, that informed the experiment design was a dissertation done at the University of New South Wales completed in 2011. Spring and colleagues conducted three experiments on levels of automation and its impact on workload and vigilance. Like the previous two studies, this one was also focused on passenger rail applications. In the first experiment, student participants drove in four different automation conditions and had to watch for dark signals and track obstructions. Vigilance was formally measured through a detection distance, the distance between the object and the train at the moment the emergency brake was applied. Spring found that driver vigilance was negatively impacted at high levels of automation and hypothesized that this was due to under-loading on the engineer. Spring tested this hypothesis in his second experiment by adding a secondary task to increase the task load. He tested two different types of secondary tasks. First, a sensory task which required the participant to respond to a visual alert periodically throughout the trip by pressing a button. This was a fairly low workload secondary task, whereas the other type of secondary task was a cognitive task that
increased the mental workload of the participant. In the cognitive task, the participant had to do various arithmetic problems incorporating their location and speed. Spring found that with the cognitive secondary task, at high automation levels the participants were not under-loaded. This result was then taken into account in a third experiment that went back to measuring vigilance, but this time with the both types of secondary tasks to see if performance improved. This final experiment suggested that the negative impact of high automation levels on vigilance can be counteracted by a secondary task that increases task demands, like the arithmetic task (Spring 2011).

These studies, while differing in overall motivation and conclusions, provided a great introduction to the challenges of performing human factors experiments using a locomotive simulator. It was important to get an idea of how to design scenarios that both realistically represent events that a locomotive engineer might encounter and provide value in answering the research questions. It was also important to begin to understand how workload, situation awareness, and vigilance are related in the rail application and how they can be measured effectively in a human-in-the-loop experiment.

3.2. Methods

This section provides an overview of the locomotive simulator, study participants, and details the capabilities of the automation systems used in the experiment. It also gives information on the secondary task used to measure workload, the specific driving scenarios, and tasks that subjects encountered. The training procedure and schedule for the day of the experiment, experiment design, and statistical methods are included in this section as well.

3.2.1. Locomotive Simulator

This experiment utilized the DOT-FRA Cab Technology Integration Laboratory (CTIL) locomotive simulator located at the Volpe Transportation Systems Center adjacent to the MIT campus in Cambridge, MA. The Corys “Full Scale Rail Simulator” (Corys, Inc., Jacksonville FL) is a computer based locomotive training simulator and the software simulates the dynamic train handling characteristics of various train configurations. The version in the DOT CTIL was modified for research purposes, and mounted inside a simulated locomotive cab shell.
Figure 6 shows the inside of the locomotive cab. The engineer subject sat on the right side beside a conventional AAR-105 control stand with realistic throttle, brake, bell, and radio controls. Two conventional locomotive operating displays in front of the engineer displayed speed, brake pressure, throttle/brake notch position etc. Four color LED monitors mounted immediately outside the simulator cab shell (two in the front and one on either side) rendered a realistic out-the-window view of the track and the surrounding area. An experimenter confederate sat facing forward on the left side of the cab and played the role of the conductor. The simulator operator in an adjacent room controlled track signals and played the role of the dispatcher. Instructions or information was communicated from the dispatcher to the cab through an intercom system that functioned as a typical radio in rail operations.

Figure 6. Inside of the CTIL locomotive simulator. The conductor station is on the left and the engineer sits on the right.
3.2.2. Human Subjects

A total of 11 subjects from four Class I freight railroad companies participated in the experiment. All of the subjects were familiar with and had operated freight trains equipped with Trip Optimizer. The average age of the subjects was 44.2 years (37-55 years). They averaged 9.9 years of experience as an engineer with a range from 3-22 years. The subjects as a group averaged 7 years of experience in management with a range from 0-13 years. Three subjects had 2 years or less experience as a manager, while the remaining subjects had at least 5 years of experience. None of the subjects were familiar with the territory that was used for the experiment trips. All subject costs (e.g., travel, lodging) were paid by their respective companies. The experimental protocol was reviewed and approved by the MIT Committee on the Use of Humans as Experimental Subjects.

3.2.3. Automation Conditions

The three automation conditions used in the experiment were Manual- no automation, Trip Optimizer (TO)—the automation currently in use in most of the Class 1 railroads, and Trip Optimizer Plus (TO+)—the prototype automation system designed specifically for this experiment.
Manual operation required the engineer to drive the train, adjusting the throttle and brake notch, without any automation assistance. This mode mimicked real world manual operation with one exception. The rolling map display from the TO automation system was included on the main display for the Manual trip. The reason for the inclusion of the rolling map in the Manual condition was to help the engineers drive on unfamiliar territory. The rolling map shows the mileposts, permanent speed restrictions, and a course graphical indication of grade on the track. In the real world, engineers qualify on the routes they drive and are intimately familiar with all the route features. Since all the subjects were unfamiliar with the experimental route, the rolling map provided a surrogate for this route knowledge. Also, it was important to prevent the presence of the rolling map in the TO and the TO+ conditions from artificially influencing performance over the Manual condition.

Trip Optimizer is a GE product widely used in the rail industry. It is an automation system that optimizes for fuel efficiency and includes the rolling map to promote situation awareness. Engineers can engage TO and it controls the throttle and dynamic brake. To engage TO, the engineer must go through several set-up screens before beginning the trip to indicate the track, the type of locomotive, and the load. Then during the trip, the engineer selects TO from the display screen and the system prompts them to move the throttle into notch 8 (normally the full power position). Figure 8 shows the running screen for TO, including the rolling map. The rolling map display also displays the TO planned “optimal speed” which is the speed trajectory that TO will use if it is engaged and which minimizes the use of braking. If, while TO is engaged, conditions change from those originally planned, and additional air braking is needed beyond the dynamic engine braking available by reversing the throttle, the system will visually prompt the engineer to use the air brake. Information about the mode and air brake prompting is displayed above the rolling map. In addition, the system displays the automation notch position.
"Trip Optimizer Plus" is the automation system concept conceived specifically for this experiment. The task model was used as a guide for brainstorming features that allocate tasks to the automation system that are typically done by the engineer. The features included in the final version of TO+ were chosen based on the following criteria: ease of implementation in the simulator, impact on the overall workload of the engineer, and likelihood of adoption of the feature in a future product. In the end, TO+ added three additional features to the normal TO system which were incorporated onto a single display screen added to the simulator cab (see Figure 9).
First, TO+ included in-cab signals. This shows the engineer the distance, in miles, to the next signal. Once the train is within one mile of the upcoming signal, the display shows the aspect of that signal. The display updates the signal aspect in real time so if the aspect changes, that change is reflected on the display in the cab as well.

The second feature of the TO+ system is an automatic horn shown in Figure 10. This feature shows the distance, in miles, to the next crossing. Once the train is 15-20 seconds away from the crossing, the horn sequence sounds automatically. When the horn sequence is active, the display shows “Horn Active” and if the crossing is in a quiet zone, the display shows “Quiet Zone.”
The final feature of the TO+ system is automatic pacing. This feature makes the system knowledgeable about train traffic ahead. If a train ahead is going slower, pacing limits the power on the subject train when automatic mode is engaged. This slows the subject train down to the point where the engineer continues to see clear or approach medium signals, indications that the engineer is permitted to continue into the next block of track. The goal is to prevent the subject train from catching up to the train ahead and hitting a red signal that requires a full stop, since stopping and starting a long freight train consumes a significant amount of fuel. On the display, if the system detects a train ahead and employs pacing, it shows what notch the system is limiting (see Figure 11). For example, if the system is limiting the power to notch 5 or lower, the display shows “N5 Limit.” If pacing is not active, the display shows “No Limit.”
3.2.4. Secondary Task Workload Measurement

During the experiment subjects executed a visual secondary task as a method for quantifying a subject’s spare attention (or task workload). Measuring the response time to a visual secondary task, in addition to the primary task of driving the train, essentially measures the “spare visual attentional capacity” of a subject. In order for the secondary visual task to be an objective (albeit proxy) measure for spare attention and mental workload, the primary task should also be visually demanding. For example, if the primary task requires monitoring displays, a notification on a display that requires action could be a good secondary task. Use of secondary task response time as a quantitative, but proxy, measure of spare visual attention and mental workload is widely accepted (O’Donnell and Eggemeir 1986). Subjective methods for mental workload assessment exist as well, typically in the form of surveys and rating scales administered during the task or directly afterward, that focus on the operator’s feelings about the workload (Casner and Gore 2010).

The secondary task in this experiment used a small display screen placed in the upper right hand corner of the cab and a small keypad on the window sill to the right of the engineer’s
seat. A red stimulus (Figure 12) appeared randomly on the screen every 2 to 10 seconds and remained visible for 10 seconds or until the subject responded by pressing a key on the keypad. Subjects were instructed to focus on the primary task of driving the train and respond to the secondary task only when they felt capable of performing the task without compromising train handling performance, which helps ensure that changes in primary task workload appear in the response to the secondary task (O’Donnell and Eggemeir 1986). In general, if the primary task workload is high at a particular time, the subject should respond less frequently to the secondary task stimulus and the response times should be greater than during periods of low workload. The response time and response rate were recorded for the workload analysis.

Figure 12. Secondary task display.

3.2.5. Experiment Scenarios

All participants drove a route from BNSF’s Aurora Subdivision from milepost 72 (Steward) to milepost 143 (Plum River) as shown in Figure 13. This route featured portions of double-main and single-main track and various grade changes to add to the difficulty of the trip. None of the participants were qualified on this territory, so the morning training session was key for them to gain familiarity. Each participant drove the same train (MIT Train #7239) comprised
of three head end locomotives, 25 empty freight cars in the middle, and 20 loaded freight cars at the rear (Figure 14). In total, the train was 3,144 ft and 4,296 tons.

![Milepost Diagram](image)

Figure 13. Route on BNSF’s Aurora Subdivision used in the experiment.

![Train Profile](image)

Figure 14. Train profile with three head-end locomotive, 25 empty cars, and 20 loaded cars.

Three scenarios were developed for the subjects to drive in the afternoon portion of the experiment day, one scenario each for the Manual, TO, and TO+ automation conditions. While the scenarios stayed the same, the order in which each subject did the three trips varied to try to
mitigate any potential automation condition order effects. Each scenario featured five permanent speed restrictions and two quiet zones, which were in the same location for each scenario. Information on these restrictions and quiet zones, as well as general track information, was given to the participants in a timetable at the start of the experiment (see Appendix B). Several additional events were included in the scenarios, but varied in location in order to mitigate learning effects as the subject became familiar with the territory. The specific locations of these events are detailed in Appendix C, but each scenario contained:

- Two temporary speed restrictions. The subject was notified of these restrictions while en-route in a call from the dispatcher.

- One Maintenance of Way (MOW) zone. The subject was given a Form B with the location of this zone before the start of each trip.

- Five faulty gates. The crossing gates were set to the ‘up’ position for these crossings and the subject was expected to report gate malfunctions to the dispatcher. (See example in Figure 15.)

- Trespassers. The dispatcher notified the subject of possible trespassers once during each trip.

Figure 15. Faulty gate as seen in the simulator (see yellow arrow).
In this experiment, the engineer participants interacted with the dispatcher and conductor (both experiment confederates) as they would in normal operation, except the conductor would not initiate any actions, only execute an action as directed by the engineer. For example, the majority of the subjects had the conductor write down temporary speed restriction information as it came over the radio from the dispatcher. Several subjects also had the conductor report faulty gate locations to the dispatcher. The dispatcher sat in the experiment control room and used a script (see Appendix D) with the locations of specific dialogue in order to interact consistently during each trip and with each subject.

In addition to typical dispatcher interactions, the dispatcher also asked the subject situation awareness questions during the course of the trip using the Situation Present Assessment Method (SPAM) (Durso and Dattel 2004). At locations specified in the dispatcher script, the dispatcher prompted the subject with a radio call “Dispatch to MIT 7239, I have a SPAM question for you. Let me know when you are ready to answer.” The engineer responded when they were ready to answer a SPAM question. This response time, recorded with a stopwatch, provided a measure of workload for that moment. The dispatcher then followed the engineer’s reply with a question about their current state or upcoming events. The accuracy of the subject’s answer and the time it took then to answer provided two measures of situation awareness. The following questions were posed to the subject as SPAM questions:

- What is the current grade?
- Are you in a quiet zone?
- How far are you from the next speed restriction and when will you take action to comply?
- How long do you estimate it will take you to get to milepost xx?
- At what milepost is the end of train?
- Over the last five minutes have you been catching up to traffic ahead of you?
- What is TO planning to do in the next 5 miles? (Only in the TO trip.)
- Is the behavior of TO+ accounting for traffic ahead? (Only in the TO+ trip.)

The dispatcher also provided some simulated radio chatter during the trip which added realism to the scenarios and simulated the mental workload of attending and responding to the dispatcher calls.

3.2.6. CTIL Simulator Cab Configuration

Figure 16 shows the full implementation of the systems in the locomotive simulator at CTIL. The TO rolling map is on the left main display in front of the engineer as it is in typical
operation, although the subject could choose to have TO on either of the two main screens. The TO+ display added for this experiment was placed on the left side of the cab, above the brake and throttle controls. Finally, the secondary task display was placed in the upper right corner of the cab.

Figure 16. Full setup of the experiment systems in the locomotive cab.

3.2.7. Subject Training and Experiment Schedule

Subjects arrived at the facility around 8 AM and were escorted to the CTIL simulator where they reviewed, then signed, the Informed Consent form. After that, subjects participated in a briefing session to introduce them to the goals of the study, the test scenarios, the CTIL controls, and the new TO+ features. The secondary task was also shown and explained to the subjects.

Following this briefing session, subjects were brought to the CTIL to begin their two training trips. The first trip covered the entire length of the route to familiarize the engineer with the territory. The route was set to show only clear signals and the engineer drove the train in the standard manual control mode and with the TO rolling map showing upcoming grades and curves. In addition, the subject was instructed to respond to the secondary task during the
training, but the response times would not be used for any analysis. After a short 10-minute break, the subject did another training trip in the simulator, this time using TO. After driving for about 15 miles in TO mode, the simulation was paused and the features of TO+ were reviewed on the display screen of the cab. The subjects drove for another 15 miles in TO+ mode. During these training runs, a member of the team was available in the cab to answer questions about the route and the automation systems.

The briefing session and the two training runs completed the morning training portion of the experiment. After a break for lunch, the subjects completed three runs in the simulator in Manual, TO, and TO+. The order the subjects completed these runs varied. Following each trial run, the subjects completed a brief verbal survey on how they handled various aspects of the trip. Then at the end of all three runs a final survey was conducted to glean the subjects’ thoughts on the automation systems and their general performance.

3.2.8. Experiment Design

A within subject experiment was conducted with 11 subjects who were tested three times under three different automation conditions. (Due to the length of the route, and time constraints on subject availability, it was not practical to perform repeated testing within each automation condition.) With the goal of validating the static workload prediction from the task model, the dependent variable was the response time to the visual secondary task. The independent variables were subject, in this case 11 randomly chosen individuals, and the three different automation conditions used in the experiment.

3.2.9. Statistical Analysis Methods

The statistical analysis was performed using Systat 13 Version 13.00.05. In order to establish sample independence in the secondary task response time data, as well as the static workload prediction data, the autocorrelation was checked for both of these data sets (using the Matlab autocorrelation function). Hierarchical mixed regression\(^2\) was used to compare the secondary task response time data with the static workload prediction from the task model. The

\(^2\) In Systat the following series of menus were used for the Hierarchical Mixed Regression: Analyze>Regression>Mixed>Hierarchical Data

44
Kolmogorov-Smirnov (KS test) was used to check for normality in the residuals of the regression result and Levine’s test was used to check for heteroscedasticity. Both tests were done in Systat.

3.3. Experiment Results

This section describes the analysis performed on the secondary task response time data collected in the experiment. The goal of this analysis was to provide an initial relation between the workload predicted from the task model and the workload measured in the experiment with the secondary task. Regression analysis was used to analyze their ratio against the effects of subject, condition, and distance traversed at the time of measurement. Visual comparisons of the static workload prediction and the secondary task response time from the regression model suggest that they are correlated for certain sections of the trip. The coefficient in the regression model for automation condition, however, indicates that the static workload prediction requires some adjustment to best describe the actual workload experienced by the engineer.

3.3.1. Data Processing

As a first step in data processing and cleaning, the static workload prediction, calculated for every tenth of a mile, was linearly interpolated to match up with the locations of the subject responses to the secondary task. This gave a response time from the experiment and a corresponding workload value from the task model for each time the engineer responded to the secondary task. In total there are 33 sets of data, from 11 subjects each doing three trips in the simulator.

The second step in data processing consisted of removing certain points from the data. In Matlab, the following points were removed from the raw data of secondary task response time and static model workload prediction values:

- Missed stimulus. The secondary task record showed the times when the subject did not respond to the stimulus on the display. Those points were omitted.

- Preemptive keypad presses. The instances where the subject pressed a button on the keypad without a red stimulus on the display screen were omitted.

- Response times less than 100ms. Some of the subjects had a tendency to rest their finger on the keypad during the trips, often unknowingly depressing a key. This resulted in very fast response times because as soon as the stimulus appeared, a depressed key would register a response time. These ultra-fast response times were removed to preserve intended responses and omit inadvertent responses.
- **Response times greater than 10s.** The red stimulus remained on the screen for 10 seconds. Any response times greater than 10 seconds are responses after the stimulus has disappeared and should be recorded as a missed response.

- **Static model workload values less than 10.** The static model includes only the four scenarios mentioned above, so there are parts of the trip where there is not a prediction for the workload because the engineer may be engaged in tasks that are not included in the task model. The goal is to validate the model so it is necessary to examine the responses wherever there is a value for the workload from the model. In addition, because of the linear interpolation of workload prediction values, there were some uncharacteristically small values in the data. In order not to skew static model values, small values less than 10 were removed.

- **Manual portions of the trip for TO and TO+ conditions.** For the automation conditions, the static model assumes the engineer is in automatic mode through the entire trip. In reality of course, the engineer transitions from auto mode to manual mode several times in each of the auto trips. As a result, for the TO and TO+ trips, the data where the engineer is in manual mode were not included.

These selections and exclusions from the data promote a reliable comparison between the experimental workload and the static model workload.

### 3.3.2. Establishing Sample Independence

In order to do hypothesis testing based on the stepwise regression analysis, the pairs of data samples used in the regression must be statistically independent. To establish the independence of samples in the data set, plots of the autocorrelation function were considered for each of the 33 runs on both the response time data and the static model workload prediction. The autocorrelation looks at how statistically correlated successive data points are as a function of time lag. High autocorrelation indicates that the data points are not statistically independent of one another. On a plot of the autocorrelation, the first point always shows high autocorrelation because it is being compared to itself. The goal was to see how many intervening samples are required for the autocorrelation function to drop below a 95% confidence interval indicating a low autocorrelation. Some of the response time data sets showed no significant autocorrelation out to the lag 20 as shown in Figure 17 for the TO+ condition of Subject 6. For others, the response time data fell below the 95% confidence line somewhere between 5 and 10 points (or lags) as shown in Figure 18 for the TO condition of Subject 10. However, 8 of the 33 functions (24%) still showed significant persisting autocorrelation between lag 10 and lag 20 as shown in the example in Figure 19 for the TO condition of Subject 4.
Figure 17. No significant autocorrelation is seen in the secondary task response time data for the TO+ condition of Subject 6.

Figure 18. Autocorrelation in the response time data falls and remains below the 95% confidence interval for Subject 10 at lag = 4 in the TO condition.
The static workload prediction data displayed more periodicity than the response time data, with 22 of the 33 sets displaying significant autocorrelation between lag 10 and lag 20. As a result, it was necessary to down-sample the data in order to minimize autocorrelation and ensure independence of samples. Since the majority of functions from the response time data displayed diminished autocorrelation by lag 10, the data was down-sampled by selecting every 10th point. This reduced the autocorrelation such that only 15% of the response time data sets and 15% of the static model workload data sets displayed any significant autocorrelation.

3.3.3. Hierarchical Mixed Regression Result

The goal of the data analysis was to determine how well the static task model predicts the workload of the engineer under the manual, TO, and TO+ conditions. The following figures (Figure 20- Figure 22) the distributions of static workload prediction data, the secondary task response time data, and the ratio of the secondary task response time to the static workload prediction, respectively. Since the static workload prediction and the secondary task response time are both skewed to the right, the data was transformed using a logarithmic transformation.
Figure 20. Distribution of static workload prediction data.

Figure 21. Distribution of secondary task response time data.
A hierarchical mixed regression was used to find a relationship, in the form below, between the response time to the secondary task and the static model workload prediction.

\[
\log \left( \frac{\text{response time}}{\text{static model}} \right) = \text{subject effect} + \text{condition effect} + \text{distance effect} + \text{error}
\]

Choosing a logarithmic transformation on the response time and the static model workload prediction helped to normalize the skewed data. Representing the dependent variable as a ratio of the two measures being compared helped to reduce some of the variability in the data. The model can be re-written as:

\[
\log(\text{response time}) = \log(\text{static model}) + \text{subject effect} + \text{condition effect} + \text{distance effect} + \text{error}
\]

This form allows for a clearer understanding of the relationship between the secondary task response time and the static model workload prediction. Note that the “static model” data
includes the main automation effect, so the "condition effect" describes automation effects that are not included in the task model.

Before running the Systat stepwise regression analysis, some further trimming on the down-sampled data was deemed appropriate. Even with the above mentioned exclusions from the data set before down-sampling, it was helpful to omit more points at the extreme ends to improve the fit. The five points below were excluded from the regression analysis:

- Static model workload value of 13. This value is less than half of the value of the next point and its response time:static workload ratio was extreme.

- Static model workload value of 3,039. This value was much greater than the next point below it at 2,646.

- With the log(response time/static model) values in ascending order and looking at the more positive extreme, two more points were removed (in addition to the static model workload value of 13) that were close in value. These values are shown in the table below.

- With the log(response time/static model) values in ascending order and looking at the more negative extreme, one additional point was removed. This had a value of -9.304 whereas the next point was at -8.805. These values are shown in Table 2 below.

<table>
<thead>
<tr>
<th>Response Time (s)</th>
<th>Static Model Workload</th>
<th>Log(response time/static model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.167</td>
<td>1,834</td>
<td>-9.304</td>
</tr>
<tr>
<td>0.169</td>
<td>1,125</td>
<td>-8.805</td>
</tr>
<tr>
<td>0.241</td>
<td>1,188</td>
<td>-8.503</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>3.218</td>
<td>51</td>
<td>-2.758</td>
</tr>
<tr>
<td>7.018</td>
<td>103</td>
<td>-2.685</td>
</tr>
<tr>
<td>7.617</td>
<td>63</td>
<td>-2.118</td>
</tr>
<tr>
<td>3.650</td>
<td>29</td>
<td>-2.704</td>
</tr>
<tr>
<td>3.242</td>
<td>13</td>
<td>-1.378</td>
</tr>
</tbody>
</table>

The total number of raw data points for all subjects and conditions was 8774. Table 3 shows the number of samples for each subject and condition after down-sampling and removing the additional five points enumerated above.
Table 3. Number of data points for each subject and condition after down-sampling and removing five additional outliers.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Manual Condition</th>
<th>TO Condition</th>
<th>TO Plus Condition</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>38</td>
<td>17</td>
<td>26</td>
<td>81</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>13</td>
<td>24</td>
<td>73</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>14</td>
<td>27</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>16</td>
<td>16</td>
<td>67</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>16</td>
<td>26</td>
<td>82</td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>17</td>
<td>21</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>36</td>
<td>14</td>
<td>29</td>
<td>79</td>
</tr>
<tr>
<td>8</td>
<td>33</td>
<td>12</td>
<td>27</td>
<td>72</td>
</tr>
<tr>
<td>9</td>
<td>49</td>
<td>22</td>
<td>28</td>
<td>99</td>
</tr>
<tr>
<td>10</td>
<td>36</td>
<td>21</td>
<td>27</td>
<td>84</td>
</tr>
<tr>
<td>11</td>
<td>37</td>
<td>16</td>
<td>24</td>
<td>77</td>
</tr>
</tbody>
</table>

Finally, with the data trimmed, a significant, coherent regression result was obtained. Table 4 identifies the coefficients associated with the subject effects. And Table 5 shows the effects of automation condition and distance.

Table 4. Effect coefficients for each subject.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-5.811</td>
</tr>
<tr>
<td>2</td>
<td>-5.932</td>
</tr>
<tr>
<td>3</td>
<td>-5.946</td>
</tr>
<tr>
<td>4</td>
<td>-6.236</td>
</tr>
<tr>
<td>5</td>
<td>-6.150</td>
</tr>
<tr>
<td>6</td>
<td>-5.907</td>
</tr>
<tr>
<td>7</td>
<td>-5.989</td>
</tr>
<tr>
<td>8</td>
<td>-5.564</td>
</tr>
<tr>
<td>9</td>
<td>-6.037</td>
</tr>
<tr>
<td>10</td>
<td>-5.964</td>
</tr>
<tr>
<td>11</td>
<td>-5.825</td>
</tr>
</tbody>
</table>
Table 5. Coefficient, standard error, and p-value for all the effects in the regression model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>-5.942</td>
<td>0.080</td>
<td>0.0005</td>
</tr>
<tr>
<td>Fixed Effect- Manual</td>
<td>0.054</td>
<td>0.042</td>
<td>0.196</td>
</tr>
<tr>
<td>Fixed Effect- Trip Optimizer</td>
<td>-0.306</td>
<td>0.052</td>
<td>0.0005</td>
</tr>
<tr>
<td>Fixed Effect- Trip Optimizer Plus</td>
<td>0.252</td>
<td>0.045</td>
<td>0.0005</td>
</tr>
<tr>
<td>Distance</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

All of the effects are significant (p<0.05), except for the fixed effect of the manual condition. The effect of the manual condition on the ratio is not significantly different (p = 0.196) from the average effect of all three conditions. The magnitude of the subject effect can be considered a scaling factor that scales the static workload prediction to the response time. The sign on the automation condition effects indicates that the TO or TO+ predictions either overestimated or underestimated the workload. Inserting these numbers into the regression model above yields, by subject, then by condition:

Equation 3

$$\log\left(\frac{\text{response time}}{\text{static model}}\right) = \left\{\begin{array}{c}
-5.811 \\
-5.932 \\
-5.946 \\
-6.236 \\
-6.150 \\
-5.907 + (0.054 + (0.005 \times \text{distance}))
\end{array}\right\} + \text{error}$$

The residuals for the regression fit are shown below in Figure 23. The variance in the residuals is 0.816. Compared to the variance of 0.907 in the log(response time/static model) data, the variance in the residuals indicates that the regression reduces the overall variance in the data by about 10 percent. The one-sample KS test was used to test the regression model for normally distributed residuals. Levene’s test was used to test for stable variances across different slices of the data. The null hypothesis of the one-sample KS test is that the residuals on the log(response time/static model) are normally distributed. It is rejected if p < 0.05. The one-sample KS test gave p= 0.048, barely rejected. Realistically, since that value is sensitive to the omission of
outliers, we can consider that the residuals are marginally normally distributed and accept the analysis as coherent.

Levene's Test does not reject the null hypothesis that the variances are stable over four equal intervals of predicted values. The null hypothesis for Levene's test is that the data does have equal variance. Based on the mean, \( p = 0.315 \) supports stable variances.

![Figure 23. The residuals of the regression fit and a histogram of the residuals plotted against the predicted value of the log(response time/static model).](image)

Table 6. Results of the tests for normally distributed residuals and equal variance across the data.

<table>
<thead>
<tr>
<th>Test</th>
<th>p-Value</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-sample KS test</td>
<td>0.048</td>
<td>Reject (barely) the null hypothesis that residuals are normally distributed.</td>
</tr>
<tr>
<td>Levene's Test (based on mean)</td>
<td>0.315</td>
<td>Cannot reject the null hypothesis that the variances are stable.</td>
</tr>
</tbody>
</table>

### 3.4. Discussion

To illustrate the relationship between the regression model result and the actual response time to the secondary task obtained during the experiment, the data for one subject is shown
below in Figure 24 - Figure 26. These plots use the data from Subject 9 and show the response time predicted from the regression model in the following form:

**Equation 4**

\[
\text{Response Time (s)} = e^{(-6.037 + \text{cond} + 0.05 \times \text{dist} + \log (\text{static model}))}
\]

In this equation, -6.037 is the subject effect corresponding to subject 9, cond is the condition effect adjustment inserted depending on the manual, TO, and TO+ conditions, dist is the distance in miles from the start of the trip, and the static model is the value of the static model workload prediction corresponding to a particular distance.

It is important to remember that (as noted earlier) the data sets in the regression model for the TO and TO+ conditions only included response time and predicted workload pairs for the portions of the route when the engineer drove in an automatic mode. In general, each subject had TO or TO+ engaged from about milepost 13 to milepost 30. Rather than displaying the entire trip, this is the section shown in the figures below. Subject 9 had TO+ engaged through the entirety of this section, however in the TO condition, the subject disengaged auto twice as seen by the Auto/Manual line on Figure 25.

![Subject 9- Manual- Comparison of Response Time from Experiment and Regression Model Prediction](image)

**Figure 24.** For the Manual condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment.
Figure 25. For the TO condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment. Also shown at the top are the time intervals in auto mode.

Figure 26. For the TO+ condition, a comparison of the response time from the regression model to the actual response time to the secondary task during the experiment.
These three plots demonstrate that the response time, as predicted from the regression model, does follow the general trend of the actual response time for Subject 9. Particularly, the location of the peaks in response time are generally the same for both the regression model and the experiment.

However, the peak magnitude is often not well captured in the regression model. In Figure 24, the Manual case, the peaks in response time to the secondary task from the experiment between mile 22 and 30 do not show up in the regression model. Likewise, the large peak in response time to the secondary task around mile 20 in the TO+ condition, is also not captured in the regression model. This could be due to the fact that the regression model utilized down-sampled data and may not have included the full magnitude of the jumps in response time. This results in a similar trend of peaks and valleys in the regression model, but does not represent the magnitude of the peaks very well. These extremes in response time to the secondary task however, indicate the areas of high workload for the engineer and may be the most interesting sections of the data to consider. Thus, revising the regression model by including more data may be a helpful in the next phase of analysis. Additionally, there are a number of other effects that were not included in the regression model that contribute to the large residual noise in the data. Effects like test order, fatigue, distraction, changes in driving strategy by individual people, all have an effect on the secondary task response time during the trip. While the regression predictions rise and fall in about the same way as the experimental data, these other factors indicate that there is variance that is not explained by the regression model.

For a full visual comparison, the static model workload prediction is included in the following Figure 27 - Figure 29. These show the same section of the trip, mile 13 to mile 30, for the three conditions. The scale of the static model workload prediction is indicated by the y-axis on the right side.
Figure 27. Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the Manual condition.

Figure 28. Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the TO condition.
Figure 29. Comparison of the response time during the experiment, the regression model response time, and the static model workload prediction (shown on the right y-axis), for the TO+ condition.

These figures with the static model workload prediction, like the three before it, show that the static model workload prediction does follow the general trend of the actual response time to the secondary task for Subject 9. The regression model though, establishes a relationship between the response time to the secondary task in the experiment and the static model workload prediction that scales the static model down to better match the actual response time numbers.

While the visual comparison gives some indication of the correlation between the static workload prediction and the response time to the secondary task, the coefficients on the regression model provide further quantitative information. The regression model, especially in the form with log(response time) on the left side of the equation and log(static model) on the right side of the equation (Equation 2), shows the values that need to be added to the log(static model) in order to get the log(response time). The static model and the response time to the secondary task do not directly match. The subject effect, condition effect, and distance effect are all modifiers that help to correspond the static model workload value to the secondary task response time.
Since the coefficients on the subject effect are the largest, it is clear that the subject effect has the largest influence when comparing the static model workload to the secondary task response time. This subject effect is due to the variability in driving style including priority given to the secondary task. Every subject had different driving strategies or tendencies, variability that is encompassed in the subject coefficient.

The inclusion of the automation condition term in the regression model indicates that the static model of engineer workload did not appropriately account for all the changes in engineer tasks when automation is included. Theoretically, the workload prediction takes into account the automation condition. So, if the static model perfectly accounted for all the changes in engineer tasks when using automation, then there would not be a condition effect. Instead, the condition effect acts as an adjustment term that broadly indicates issues with the static model. More specifically, the directionality of the condition coefficient shows that the static model either underestimates or overestimates the actual workload. Since the coefficient for TO is negative, and the effect is significant, then the static model for TO overestimated the actual workload. Similarly, since the coefficient for TO+ is positive, and it is a significant effect, then the static model for TO+ was an underestimate of the actual workload.

Finally, the distance effect indicates that the static model workload prediction may need to include a distance factor that was not initially considered in the prediction. One explanation for this effect is that it represents the subject’s fatigue through the course of a trip. The static model does not consider changes in workload due to fatigue, but a fatigue factor may be necessary to compare the static model to the secondary task response time. Another explanation of this effect is that the static model does not include responses to non-clear signals. In the experiment, the majority of the signals at the beginning of the trip are clear and the static model reflects that accurately. As the trip continues though, the subject is more likely to encounter non-clear signals and there is certainly workload associated with responding to non-clear signals. The static model does not include this workload and so nearing the end of the trip the static model is underestimating the actual workload.
4. Utilizing Human Factors Work in the Policy Sphere

The workload analysis described in the previous chapter is a first attempt at validating the workload prediction aspect of the task model. While further work and iterations on the model are necessary to confidently predict the changing workload of the engineer under different automation conditions, the model contains valuable qualitative information on the job of the engineer and the conductor and the ways they interact with automation. This chapter poses two questions. First, can the information obtained during task model development and incorporated in the model be useful in policy discussion and decisions? Answering this question in the affirmative and providing the benefits of using the information from the model in the policy sphere leads directly to a second question: What steps need to be taken to get the information from the task model into an accessible form?

Before answering these two questions, it is necessary to situate the task model and the work of this thesis in the greater context and discussion of automation in the rail industry. Numerous events over the past decade have influenced the direction of both technological advancement of automation and the policies that regulate it. These key events highlight the relevance of the modeling and experiment work described in the previous chapters and the potential usefulness of information developed by the modeling effort in policy discussions.

4.1. Key Events in Rail Automation Discussion

The timeline of events that contribute to the discussion of rail automation involves several key players, or stakeholders. The following stakeholders each have a strong voice and interest in the discussion.

(1) Railroads. There are seven Class 1 railroads in North America: BNSF, CSX, Union Pacific, Kansas City Southern, Norfolk Southern, Canadian National, and Canadian Pacific.

(2) Federal Railroad Administration (FRA). Established in 1966 with the creation of the Department of Transportation, the FRA is the regulating body. It has the authority, through the Railroad Safety Act of 1970, to regulate the rail industry on any and all safety issues (McDonald 1993).

(3) National Transportation Safety Board (NTSB). The NTSB is an organization that conducts safety studies and performs accident investigations in the areas of aviation, rail, highway, and marine. They provide recommendations to the appropriate industries on
how to improve safety, however these are only recommendations. The NTSB does not have the authority to mandate that an industry adopt their recommendations. Each year, the NTSB publishes a “Most Wanted” list of the critical improvements that need to be made in various industries.

(4) Labor Unions. Several unions such as the Brotherhood of Locomotive Engineers and Trainmen (BLET), the International Association of Sheet Metal, Air, Rail and Transportation Workers (SMART), and the United Transportation Union (UTU) represent the interests of the engineers, conductors, and other employees in the industry.

(5) General Electric (GE). GE’s Transportation Division is one of two U.S. companies that manufacture locomotives. Many freight locomotives in use today are also equipped with GE’s Trip Optimizer, an automation system that controls the throttle and dynamic brake of a locomotive optimizing for fuel efficiency. As evidenced by the GE/MIT project, GE is actively involved in the research and development of new automation systems for locomotive control, including human factors engineering issues.

A key event in this timeline is a major train collision that occurred between a Southern California Metrolink passenger train and a Union Pacific freight train in September 2008. The accident resulted in 25 fatalities and about $12 million in damages. The NTSB found that the collision was caused when the driver of the Metrolink train failed to stop at a red signal because he was distracted text messaging. In the accident report, the NTSB notes that the collision would have been prevented had Positive Train Control been installed in the locomotive as it would have stopped the train before the red signal (National Transportation Safety Board 2010).

For a bit of background, Positive Train Control (PTC) is a collision prevention automation system that uses GPS to locate the train precisely on the track and communicate with other trains and dispatchers about appropriate movement authority. Ultimately, PTC systems can slow down or stop a train in the event of unauthorized movement, rule violations, or other incidents before an accident occurs. This was conceived as a backup safety system that should work seamlessly in the background of normal operations. It is the locomotive engineers’ responsibility to stop at red signals, communicate with dispatchers about the location of other trains, and comply with all movement authorizations they are given. However, unanticipated circumstances do occur and human error does play a role in both minor accidents and major collisions. Having a redundant system in place that activates in these unanticipated circumstances or cases of locomotive engineer mistake or inattention adds a significant level of safety to rail operations as a whole.
PTC had been on the NTSB's “Most Wanted” list, since 1990. But the Metrolink collision was the impetus that led the FRA to mandate, through the Railroad Safety Improvement Act of 2008, that Class 1 railroads install PTC systems on all passenger routes and all tracks that carry toxic-by-inhalation materials by the end of 2015 (Association of American Railroads). Despite the fact that the FRA issued a Final Rule on PTC systems in 2012 that made 10,000 of the 230,000 miles of US rail track exempt from the mandate (Federal Railroad Administration), the 2015 deadline was unachievable and Congress granted an extension to the railroads until 2018 (US House 2015). Currently, the railroads are continuing to upgrade and implement PTC technology to meet the new deadline (Association of American Railroads).

With PTC implementation in the works in the years following the congressional mandate, one of the biggest railroads in the country, BNSF, made a bold move. In the summer of 2014, BNSF proposed a one-person crew configuration, eliminating the conductor on routes and trains equipped with PTC (Wilner 2014). While the US freight industry has financially done well in recent years, rail companies, like most businesses, are looking for ways to cut costs and increase profits in the face of high fuel and labor costs. The opportunity to do this comes with the technological advance of automation systems. The use of Trip Optimizer is one approach to the issue of high fuel cost as it optimizes to save fuel. BNSF took it a step further by exploring the possibility of reallocating some of the conductor functions to automation as well. The proposal was negotiated with union officers from SMART, which represents conductors, and created a new position of "master conductor" who would be on-call to drive by car to trains that need a conductor to be physically present e.g. to couple/uncouple cars, check siding clearance etc. Union members strongly opposed the proposal and argued firmly that having two individuals in the cab is necessary for safe operation of a train. As a result, members voted down the proposal (Bradbury 2014). Despite the proposal’s failure, the idea of eliminating the conductor from the cab sent a shock wave through the industry. As technology like PTC and other automation systems continues to develop, the discussion of what is necessary for safe operation of a train will continue as well.

The next relevant event in the discussion of locomotive automation is the May 2015 Amtrak train crash in Philadelphia. This crash resulted in the death of 8 people and many more injuries. The passenger train derailed when it entered a curve at 106 mph, more than twice the 50 mph speed allowed on this particular section of track. According to the NTSB accident report,
released a full year after the accident took place, the engineer was distracted by radio discussion of an emergency on another train in the area and as a result lost awareness of his location. Thinking he was past the curve already, he accelerated the train to meet the 110 mph allowed on the track following the curve. The NTSB also concluded that PTC, which was not yet implemented on this particular section of track at the time of the accident, would have prevented the fatal error and slowed the train down prior to the curve (National Transportation Safety Board 2016). This unfortunate crash highlighted the importance of systems like PTC and brought the previously primarily technical discussion of automation on trains to the press and public’s attention.

Finally, in early 2016 two of the prominent unions representing freight railroad employees (SMART and BLET) wrote separate letters to the FRA to express their concern about reliability, crew reliance, and potential deskilling resulting from use of cab automation designed to assist the engineer in reducing fuel consumption, specifically the use of GE’s Trip Optimizer and a similar WABTEC, Inc. system LEADER (Locomotive Engineer Assist Display Event Recorder). The letter from the BLET on February 4:

“BLET joins SMART-ID in requesting that [the FRA] issue an emergency order prohibiting the use of these technologies until they are further examined, to ensure that they do not pose risks to the safe operation of freight railroads. At a minimum FRA should examine to what extent new regulatory requirements are necessary to ensure safe operations, testing and maintenance and actual performance of these systems in the field” (Pierce 2016, 1).

In these letters, the unions acknowledge that the addition of automation systems in the cab changes the role of the engineer from being a “primary operator” to a “passive monitor of automation” (4). They are concerned that the FRA is not adequately considering the safety implications of this changing role. Currently, TO is classified as a “non-vital business” system, a different classification from safety systems like PTC. Since the use of TO and LEADER have been so widely adopted by the railroads, the unions are advocating that the FRA consider regulating its use and adoption in the industry.

All these events, from the congressional mandate of PTC to the union letters petitioning the FRA to evaluate certain automation systems further, indicate the progression of the discussion around automation and the changing role of the engineer and conductor in rail operations. While the future is a bit unknown, the unions and the railroads are keenly aware of
the technological advances in automation systems and the changes in rail operations made possible by these systems. The FRA then is aware of the new regulatory needs prompted by these systems.

4.2. Can the qualitative information from task modeling be useful in the policy sphere?

The outline of events in the rail automation discussion provided above establishes the relevance of further research on human-automation interaction in the locomotive cab. Given this relevance, there is certainly a need and a use for the qualitative information contained in the task model. Specifically, the task model provides both high level information about the goals and priorities of the rail domain and low level details about the job of the engineer and conductor and their interactions with automation. From the high level goals of safety and efficiency, to the aspects the engineer and conductor need to maintain awareness of and manage, to the lower level tasks involved in various operational scenarios—using this kind of information and knowledge about rail operations in the policy sphere has numerous benefits.

First, the model can be used to bring stakeholders and other interested people up to speed on the job of an engineer and conductor and the human-automation interactions that are part of their current roles. Before the task model created for the GE/MIT project was developed, individuals interested in rail operations had access to the aforementioned cognitive task analysis reports on the engineer, conductor, and dispatcher. While these reports are certainly useful, and their ethnological task descriptions are very readable, they are not always conducive to helping non-experts appreciate the complex interactions between all the operators and the automation systems. The model developed for this project built upon the cognitive task analysis reports by identifying the knowledge gaps in the reports, updating them to reflect current rail operations, and adding information from expert knowledge. Thus significant effort was put into understanding rail operations and expanding the foundation of knowledge in the domain. The information contained in the resulting task model then has the potential to educate people who need to know or have an interest.

The identified goals and detailed task definitions in the task model can also help establish a shared language between stakeholders. This is key to facilitating discussions that will inevitably involve the changing role of human operators. Having a baseline understanding of the
current job of an engineer and a shared language to talk about the tasks that they do, can make conversations more efficient and effective.

Another benefit of using the task model information on the role of the engineer and conductor in rail operations is that using it will promote knowledge translation between academia and the FRA. Baumbusch and her colleagues write about the “gap between research and practice” in the medical setting and the need to “transfer research findings into practice” (Baumbusch et al. 2008, 131). In the case of this project, the FRA, as the policymaker, can be classified as the practitioner, much like doctors and nurses are practitioners in the clinical setting. Baumbusch et al. highlight the challenges of translating knowledge gained through research into practice and similarly, the knowledge gained from the modeling work of this project has the potential to stay confined within academia and industry. Instead, actively using the task model in the policy sphere cultivates an environment of information exchange and helps extend the audience beyond the individuals directly involved with the project.

Overall, using the qualitative information from the task model in policy discussions helps ensure that the knowledge gained through the modeling work, the understanding of the tasks of the engineer and conductor and automation, will not remain in academia or industry, but rather will reach a wider audience. Research and experiments are costly and time consuming. A significant amount of effort went into creating this model and validating the computational workload prediction. Given this effort, it is worthwhile and relevant to use the qualitative findings to help bring people up to speed on rail operations, create a shared language for all stakeholders involved in rail automation discussions, and further the knowledge translation between researchers and policymakers.

**4.3. What steps need to be taken to make the task model information accessible to policy makers?**

While using the information in the task model can be beneficial in policy and rail automation discussions, unfortunately the model is not in an accessible format for immediate use by the FRA or other stakeholders. This is typical of many of the formal models that come out of human factors or task analysis. Recall from Chapter 2.1, the form that most abstraction hierarchies or hierarchical task analyses take is a complex, often cumbersome, hierarchy or an extensive series of tables describing tasks and their relation to one another. This is certainly true
of the model developed as part of this project. The hybrid abstraction hierarchy/ hierarchical task analysis is contained in tables that make reading and cross-referencing difficult. In order to compute the workload prediction, the information in the task model is also contained in the program Lisp and accessible only to a small subset of the project team.

The remainder of this section offers several suggestions for both improving the model and getting it into a format that is usable and accessible to non-experts who could benefit from the information. First, it would be helpful to add information on the role of the dispatcher in rail operations to the task model. The concept map and a cognitive task analysis of the dispatcher’s job was synthesized as part of this project for gap analysis, but was not formally or completely included in the task model. In future rail automation discussions, the role of the dispatcher, in addition to the engineer and conductor, could change and evolve as automation systems develop. It is important then that the qualitative information and linkages in the model reflect the whole system. Including the dispatcher tasks will achieve this.

Second, the information from the model on the engineer, conductor, dispatcher, and automation interactions should be transformed into a visual and interactive tool. This could be a graphical interface that shows the hierarchy of goals, priorities, and low-level tasks that each operator does, similar to the hierarchy that is hidden within the current Lisp language version of the task model. The ability to highlight different tasks or full branches and see more detailed descriptions of the tasks would be helpful in bringing people up to speed and creating a shared language. There should also be a process for editing and adding tasks or further descriptions to the model. Another aspect that could be incorporated into this visual tool is the temporal nature of the engineer and conductor tasks. Hierarchies can often hide the timeline of events that occur, but a kind of side-by-side timeline as tasks occur in different scenarios could enhance the information further. Overall, this visual and interactive tool will make the model information most usable and could help illuminate safety concerns or operational issues as different automation configurations are considered by stakeholders.

The final suggestion for getting the task model information into an accessible form for use in policy discussions is to involve all stakeholders. Since the overall goal of making the information from the model accessible is to engage stakeholders in discussion and create a shared language then it is vital to start involving these individuals in the development process as well. Various stakeholders can provide input on creating the interactive tool, check the
information and descriptions of tasks, ensure the language is understandable and accurate, and identify further gaps in knowledge that could be filled. To be clear, a number of stakeholders were involved in various aspects of the GE/MIT project from the beginning. For example, individuals primarily from GE participated in a survey to glean information about the future of freight rail automation. In addition, several of the major railroads supplied subjects for the experiments and as a result gave feedback on the prototype TO+ system. However, the unions as a group were not involved in any part of the project up to this point. Arguably, the unions have considerable expert knowledge about the job of the engineer and conductor and their interactions with automation systems currently in use. This is the kind of information and input that would improve the visual tool and help make it accessible. Also, as evidenced from the letter written to the FRA in early 2016, the unions are concerned about the future of rail operations. Including them directly in the process of making the qualitative information from the task model accessible and useful could increase acceptance and credibility of the model in the future.

Mostashari and Sussman advocate for the inclusion of stakeholders in developing visual models for use in environmental policy discussions. They write of the benefits of stakeholder involvement, benefits which could translate directly into the rail policy sphere as well:

“The involvement of stakeholders from early on can ideally help the expert in constructing a model with realistic assumptions and relationships. It can also improve stakeholders’ understanding of the behavior of the larger system in question, thus enabling them to appreciate the rationale behind different policies. Another advantage of early stakeholder involvement is that it will allow for technical, economic, social, and political feasibility of policies to be considered simultaneously, potentially leading to improved policy recommendations” (Mostashari and Sussman 2005, 360).

In the case of rail operations, involving the stakeholders, especially the unions, will help produce a visual tool that most accurately depicts the role of the engineer, conductor, dispatcher, and automation. With multiple groups’ input and iterations on both the information and the format in which it is presented, the tool produced and used could help everyone involved to understand the impact of changes in rail technology. Mostashari and Sussman conclude their proposal for involving stakeholders in model development by saying:

“It can lead to improved models that take into account stakeholders’ local knowledge and experience, more inclusive policy process leading to higher acceptability of recommendations, better adherence to implementation due to process ownership, better understanding of stakeholder concerns in direct modeling dialogue, and better understanding of institutional issues” (384).
This appropriately summarizes the benefits stakeholders could see in the policy sphere as well if they are involved in making the qualitative information from the task model of rail operations accessible.

4.4. Limitations and Conclusion

The above proposal to use the qualitative information from the task model in policy discussions about the future of rail automation systems is not without its limitations. First, it is likely that an expert—someone knowledgeable about current rail operations, with an understanding of the task model developed as part of the GE/MIT project, and ideally with knowledge of human factors modeling techniques more generally—would be needed to facilitate the process of transforming the information from the task model into the interactive tool described above. Because of the funding structure of research, it may be difficult to allocate the time and money necessary to do the facilitating work. Then even with an expert helping with the process, it will be a challenge to truly ensure that the tool is easy to work with and the information on the role of the engineer, conductor, and automation is accessible. Without the input from numerous stakeholders with varying degrees of knowledge about rail operations, the information from the task model could remain jargon-filled and inaccessible.

Involving stakeholders is key to the success of this proposal, but recruiting stakeholders to participate may be a challenge. While the process of transforming the qualitative information from the task model into an interactive tool and then using that tool in policy discussions is meant to be collaborative and bring people to the table, it may in fact have the opposite effect. Some stakeholders might see this effort as a potential threat and worry that their participation will imply an endorsement of future rail automation outcomes. The key to prevent this kind of opposition is to emphasize the education and shared language potential of using the information from the task model. The purpose of this effort is not to select automation systems or show that roles can be eliminated, rather the goal is to further elicit the expert knowledge from union members, educate policy makers in the specifics of rail operations, and create a tool that can help bring everyone together for the tough discussions about the future of rail automation.

The proposal offered here is the beginning of a conversation about using the information and knowledge that emerges from human factors analysis and experiments in the policy sphere. Translating the qualitative information from a task model into an accessible, interactive, and
visual tool that engages stakeholders in the development and the end-use is one way to begin to move the rail human factors work out of the research setting and into policy discussions.
5. Conclusion

This thesis has two distinct, yet related, aims. The technical aim of this work was to see if the workload of the locomotive engineer, under different automation conditions, could be predicted from a task model that details the role of the engineer and conductor in rail operations and their interactions with automation. This task model incorporated a static workload computation which relied on a series of heuristics that identified when the engineer was likely engaged in a particular task during a trip. Following a human-in-the-loop experiment in a locomotive simulator, which measured the engineer’s workload through a secondary task, the prediction of the workload changes over the course of a trip was compared to the actual changes in response time to a secondary task using regression analysis. Visual comparisons of the static workload prediction and the regression model result show that the two values appear correlated for certain sections of the trip. However, the presence of a coefficient on the automation condition term in the hierarchical mixed regression result indicate that some adjustment is necessary in the static workload prediction in order to best represent the workload changes experienced by the engineer. In addition, while the regression model shows significant effects, the large residuals indicate that the regression model only predicts a fraction of the total variance in the data.

Given the regression result and the adjustment term of the automation condition coefficient, one logical next step for future work would be to iterate on the task model. Going back to the visual comparisons of the secondary task response time and the static workload prediction would illuminate locations of the trips where the correlation between the two was particularly poor. Figuring out what the engineer was doing at those points in the trip and adjusting the task model to better reflect the engineer’s true tasks would help fine-tune the static workload prediction. In addition, the heuristics that identify where the engineer is engaged in a particular operational scenario could be assigned differently to better reflect what the engineer is doing and the static workload computation itself could be changed. Once changes are made to the task model, heuristics, or static workload computation, the regression analysis can be performed again. Evidence of improvement would be a reduction or elimination of the coefficient on the automation condition.

Aside from iterating on the task model to fine-tune the prediction of changes in the engineer’s workload, future work could involve expanding the task model to include other
operating scenarios. The four scenarios in the current task model are all things that happen en-route, while the train is moving. The beginning and end of trips could be another aspect of the locomotive operator's job to include in the task model and consider the ways in which automation plays a role. Finally, the secondary task response time data from the experiment was used to validate the task model workload prediction for the engineer. However, the conductor tasks were modeled as well. An experiment to validate the predicted changes in the conductor’s workload and performance could be interesting future work in this area.

While the first aim of this thesis was technical and dealt with the computational aspect of the task model, the policy aim looked at the utility of the qualitative information contained in the task model. Several benefits were listed in response to the question of whether or not the qualitative information from the task model could be useful in policy discussions about the future of rail automation. Namely, the information from the model can help educate people on rail operations, provide a shared language among stakeholders for use in discussions, and assist in knowledge translation between researchers and policymakers. With these benefits in mind, a proposal was put forth for getting the information into an accessible form. First, it would be helpful to include the dispatcher interactions in the task model in order to represent the whole system—the interactions between the engineer, conductor, dispatcher, and automation. Second, the information from the task model should be transformed into a visual and interactive tool that can be used by stakeholders. And finally, stakeholders should be involved in the development of the tool in order to increase potential buy-in, credibility, and end-use accessibility. While this proposal is not without its implementation challenges, it is the beginning of a conversation about expanding the audience of the task model and utilizing the knowledge gained through human factors research endeavors in the policy sphere.
References


Appendix A. Concept Maps for the Conductor and Dispatcher
### Appendix B. Route Timetable

#### BNSF Aurora SUB – Steward to Savanna

<table>
<thead>
<tr>
<th>STATIONS</th>
<th>MP</th>
<th>TYPE OF OPER.</th>
<th>X’S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steward</td>
<td>77.9</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td>Lease</td>
<td>80.2</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td>Rochelle</td>
<td>83.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(UP) NX Xing</td>
<td>83.7</td>
<td>2 MT</td>
<td></td>
</tr>
<tr>
<td>CP 844</td>
<td>84.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flag Center</td>
<td>86.3</td>
<td>Q Z</td>
<td></td>
</tr>
<tr>
<td>Chana</td>
<td>92.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>98.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stratford</td>
<td>107.4</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td>Polo</td>
<td>112.4</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td>Carter</td>
<td>116.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milledgeville</td>
<td>122.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chadwick</td>
<td>129.4</td>
<td>Q Z</td>
<td></td>
</tr>
<tr>
<td>Burke</td>
<td>138.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plum River</td>
<td>142.3</td>
<td>CTC</td>
<td></td>
</tr>
<tr>
<td>Savanna</td>
<td>143.7</td>
<td>2 MT</td>
<td></td>
</tr>
<tr>
<td>IMPLR Xing</td>
<td>144.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP 1462</td>
<td>146.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**GCOR RULES IN EFFECT ON BNSF Aurora – Steward to Savanna.**

BNSF Aurora Subdivision, Steward to Savanna, is controlled by the Aurora sub train dispatcher.
Maximum Authorized Speed is 60 MPH for freight trains on the Aurora Sub- Steward to Savanna, unless noted below.

PERMANENT SPEED RESTRICTIONS:
MP 82.2 TO MP 83.7    40 MPH
MP 83.7 TO MP 83.9    35 MPH
MP 83.9 TO MP 84.4    45 MPH
MP 95.8 TO MP 102.3   45 MPH
MP 142.0 TO MP 144.5  35 MPH

Maximum Authorized Speed through all crossover turnouts and siding turnouts is 35 mph.

Quiet Zones at highway/pedestrian rail grade crossings from MP 86 TO MP 94 and from MP 128.5 TO MP 1
Appendix C. Scenario Summaries

In all scenarios:
1) MP 77.7 Before grade crossing (signal only)
2) MP 77.9 Turnout to MAIN 2, Signal shows DIVERGING CLEAR aspect
3) MP 84.3 Signal shows APPROACH aspect
4) MP 86.15 Signal shows DIVERGING CLEAR aspect
5) MP 86.3 Crossover to MAIN 1
6) MP 142.4 Signal shows APPROACH aspect
7) MP 143.2 Signal shows STOP aspect

Scenario 1 (TO+) Specific:
1) MP 79 Dispatcher call to inform of TSR at MP 89
2) MP 81.8 Faulty gate
3) MP 89 to MP 90 Temporary speed restriction 45mph
4) MP 93.3 Faulty gate
5) MP 104 Faulty gate
6) MP 110 Signal board for MOW
7) MP 111.5 Dispatcher call to inform of TSR at MP 119.5
8) MP 112 Start of MOW
9) MP 114 End of MOW
10) MP 115 Dispatcher call asking for estimated arrival at MP 126
11) MP 119.5 to MP 120.5 Temporary speed restriction 45mph
12) MP 123 Faulty gate
13) MP 129.9 Faulty gate
14) MP 139.5 Dispatcher call reporting intruder at MP 142

Scenario 2 (TO) Specific:
1) MP 78.5 Faulty gate
2) MP 88 Signal board for MOW
3) MP 90 Start of MOW
4) MP 92 End of MOW
5) MP 96.1 Faulty gate
6) MP 98 Dispatcher call to inform of TSR at MP 113
7) MP 106 Dispatcher call asking for estimated arrival at MP 117
8) MP 113 to MP 114 Temporary speed restriction 45mph
9) MP 114.75 Faulty gate
10) MP 116 Dispatcher call reporting intruder at MP 119
11) MP 121 Dispatcher call to inform of TSR at MP 131
12) MP 122.5 Faulty gate
13) MP 131 to MP 132 Temporary speed restriction 45mph
14) MP 132.3 Faulty gate
Scenario 3 (Manual) Specific:
1) MP 81.1 Faulty gate
2) MP 91 Dispatcher call to inform of TSR at MP 106
3) MP 92.3 Faulty gate
4) MP 100 Dispatcher call asking for estimate arrival at MP 109
5) MP 105.6 Faulty gate
6) MP 106 to MP 107 Temporary speed restriction 45mph
7) MP 108 Dispatcher call reporting intruder at MP 111.2
8) MP 116.5 Signal board for MOW
9) MP 118.5 Start of MOW
10) MP 120.5 End of MOW
11) MP 127 Dispatcher call to inform of TSR at MP 137
12) MP 128.8 Faulty gate
13) MP 138 to MP 139 Temporary speed restriction 45mph
14) MP 139.4 Faulty gate
## Appendix D. Conductor and Dispatcher Script for Manual Scenario

<table>
<thead>
<tr>
<th>Scenario Manual</th>
<th>MP</th>
<th>Who</th>
<th>Dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>77.7</td>
<td>Engineer</td>
<td>This is MIT 7239 to the Aurora sub dispatcher. Requesting permission to depart.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dispatcher</td>
<td>Ok, MIT 7239 comply with signal indication.</td>
</tr>
<tr>
<td></td>
<td>81.1</td>
<td>Engineer</td>
<td>Report to dispatch, there is a faulty gate at 81.1. The barriers are up.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dispatcher</td>
<td>Alright, MIT 7239. I understand that we have a crossing gate failure at crossing at MP 81.1. I will get the signal maintainer out there. Thank you.</td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>Dispatcher</td>
<td>Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dispatcher</td>
<td>Are you in a quiet zone?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Engineer</td>
<td>MIT 7239, this is the Aurora Sub Dispatcher. I have a temporary speed restriction for you ahead, let me know when you’re ready to copy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conductor</td>
<td>We are ready to copy. Go ahead.</td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>Dispatcher</td>
<td>This is Track Bulletin Form A No. 6531 on the Aurora Subdivision to MIT 7239. Line No. 1. Between MP 106 and MP 107 speed is limited to 45 mph on all tracks. Flags at milepost is none. I'm ready for a repeat.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Engineer</td>
<td>This is Track Bulletin No. 6531. On the Aurora sub, I understand that between MP 106 and MP 107 speed is limited to 45mph on all tracks. No flags displayed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conductor</td>
<td>That is correct 7239. Let's make that complete at (time). (dispatcher initials)</td>
</tr>
<tr>
<td></td>
<td>92.3</td>
<td>Engineer</td>
<td>Report to dispatch, there is a faulty gate at 92.3. The barriers are up.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dispatcher</td>
<td>Alright, MIT 7239. I understand that we have a crossing gate failure at crossing at MP 92.3. I will get the signal maintainer out there. Thank you.</td>
</tr>
</tbody>
</table>
Dispatcher

Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.

Response Time ________________ seconds.

Dispatcher

How far are you from the next speed restriction and when will you take action to comply?

Engineer Response

Dispatcher

Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.

Response Time ________________ seconds.

Dispatcher

How long do you estimate it will take to arrive at MP 109?

Engineer Response

100

 Dispatcher

Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.

Response Time ________________ seconds.

Dispatcher

Report to dispatch, there is a faulty gate at 92.3. The barriers are up.

Alright, MIT 7239. I understand that we have a crossing gate failure at crossing at MP 92.3. I will get the signal maintainer out there. Thank you.

108

Dispatcher

Dispatch to MIT 7239 crew over.

Engineer

Go ahead dispatch.

Dispatcher

This is the Aurora sub dispatcher to MIT 7239. We have reports of trespassers near the private crossing at MP 111.2. Please whistle freely and keep an eye out.

Engineer

Ok, dispatch. I understand there are reports of trespassers near the private crossing at MP 111.2 I will whistle and watch out.

Dispatcher

That is correct. Dispatch out.

109

Dispatcher

Record Time ________________

113.5

Engineer

(Contacts foreman. Gives train location, train number, track.)

Foreman

This is Foreman _______ using Track Bulletin Form B No. 6822, on the Aurora Subdivision to MIT 7239. Line No. 1. between MP 118.5 and MP 120.5. MIT 7239 may pass red flag at MP 118.5 without stopping and proceed at maximum authorized speed, but do not exceed 25mph between MP 118.5 and MP 118.75, head end restriction only.
Foreman ______ using Track Bulletin Form B No. 6822, on the Aurora Subdivision. Line No. 1 between MP 118.5 and MP 120.5. I understand we can pass the red flag at MP 118.5 without stopping. But do not exceed 25mph between MP 118.5 and MP 118.75, head end restriction only.

Foreman That is correct 7239. Foreman ______ out.

117 Dispatcher Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.
Response Time _______________ seconds.

Dispatcher What is the current grade?
Engineer Response ________________________________

121 Dispatcher Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.
Response Time _______________ seconds.

Dispatcher What MP is the end of train at?
Engineer Response ________________________________

127 Dispatcher MIT 7239, this is the Aurora Sub Dispatcher. I have a temporary speed restriction for you ahead, let me know when you're ready to copy.

Engineer/Conductor We are ready to copy. Go ahead.

Dispatcher This is Track Bulletin Form A No. 6532 on the Aurora Subdivision to MIT 7239. Line No. 1. Between MP 137 and MP 138 speed is limited to 45 mph on all tracks. Flags at milepost is none. I'm ready for a repeat.

Engineer/Conductor This is Track Bulletin No. 6532. On the Aurora sub, I understand that between MP 137 and MP 138 speed is limited to 45mph on all tracks. No flags displayed.

Dispatcher That is correct 7239. Let's make that complete at (time). (dispatcher initials)

128.8 Engineer Dispatcher Report to dispatch, there is a faulty gate at 128.8. The barriers are up.

Alright, MIT 7239. I understand that we have a crossing gate failure at crossing at MP 128.8. I will get the signal maintainer out there. Thank you.

132 Dispatcher Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.
Response Time _______________ seconds.
Dispatcher: How far are you from the next speed restriction and when will you take action to comply?

Engineer Response

139.4

Dispatcher: Report to dispatch, there is a faulty gate at 139.4. The barriers are up.

Engineer Response: Alright, MIT 7239. I understand that we have a crossing gate failure at crossing at MP 139.4. I will get the signal maintainer out there. Thank you.

Dispatcher: Dispatch to MIT 7239. I have a SPAM question for you. Let me know when you are ready to answer.

Engineer Response: Response Time _____________ seconds.

Dispatcher: Over the last 5 minutes have you been catching up to traffic ahead of you?

Engineer Response: ___________