A Workload Model for Designing & Staffing Future Transportation Network Operations

by

Victoria Chibuogu Nneji

Department of Mechanical Engineering & Materials Science
Duke University

Date:_______________________
Approved:

____
Mary (Missy) L. Cummings, Supervisor

____
Alfredo A. Garcia

____
Jeffrey T. Glass

____
Michael M. Zavlanos

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Mechanical Engineering & Materials Science in the Graduate School
of Duke University

2019

ABSTRACT

A Workload Model for Designing & Staffing Future Transportation Network Operations

by

Victoria Chibuogu Nneji

Department of Mechanical Engineering & Materials Science
Duke University

Date:_______________________
Approved:

____
Mary (Missy) L. Cummings, Supervisor

____
Alfredo A. Garcia

____
Jeffrey T. Glass

____
Michael M. Zavlanos

An abstract of a dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Mechanical Engineering & Materials Science in the Graduate School
of Duke University

2019

Abstract

Across multiple industries (e.g., railroads, airlines, on-demand air taxi services), there are growing investments in future automated transportation systems. Even with these investments, there are still significant human-systems engineering challenges that require deeper investigation and planning. Specifically, fleets that include new levels of automation may require new concepts of how to design and staff network operations centers. Network operations centers have existed for over a century in the railroad and airline industries, where dispatchers have played a central role in safely and efficiently managing networks of railroads and flights. With operators in such safety-critical and time-sensitive positions, workload is the key indicator of their performance in terms of accuracy and efficiency. Yet, there are few tools available for decision-makers in these industries to explore how increasing levels of automation in fleets and operations centers may ultimately affect dispatcher workload.

Thus, this thesis presents a model of dispatcher workload. While automation may be the most pressing change in transportation industries, 10 variables related to configurations of the fleet and the operations center and how those variables interact to influence dispatcher workload were defined. These ten variables come from fleet conditions, strategic design factors, tactical staffing factors, and operational factors. A discrete event simulation was developed to computationally model dispatcher workload

with over 10^{18} possible configurations of these variables. Additionally, using time-based metrics and integrating results from a prior human reliability assessment, the simulation predicts human error on tasks.

A multi-level validation strategy was developed to build internal, external, and general confidence in using the dispatcher workload model across different domains with data from freight railroad, commuter railroad, and airline operations. In the process of developing and validating the workload model, several other research contributions were made to the field. Eighty-five probability density functions of dispatcher task inter-arrival and service time distributions were generated in the three domains. A data collection tool, Dispatcher's Rough Assessment of Workload-Over Usual Times (DRAW-OUT), was designed to gather empirical dispatcher-generated estimates of utilization, the proxy for workload, throughout their shifts.

Using the model, experiments were conducted to analyze the sensitivity of dispatcher workload and performance to changes in different parameters. The size of the fleet a dispatcher managed was found to be the most significant factor out of all the other internal parameters. On the other hand, shift schedule, environmental conditions, and operator strategy were the parameters found to have the smallest influence on dispatcher performance. The model was also used to investigate future scenarios that managers could not previously explore due to limitations of time and resources. Results show that the general model is applicable for use in simulating dispatcher workload in both freight and commuter railroad operations as well as airline operations, including short- and long-haul flights, in present-day and future cases.

General confidence was built in the workload model and the Simulator of Humans & Automation in Dispatch Operations (SHADO) was developed as an online platform to provide open access to the underlying discrete event simulation. SHADO is a novel tool that allows stakeholders, including operational managers, to rapidly prototype dispatch operations and investigate human performance in any transportation system. With several theoretical and practical contributions, this work establishes the foundation for future research in the growing field of advanced transportation network operations.

Dedication

This work is in honor of my dear mother, Monica Chibuogu Nneji. As she would always say, to God be the glory! This journey has not been easy, but I have found each moment and opportunity I have had to learn and grow in the process to be a blessing. The journey continues.

Contents

List of Tables

List of Figures

Acknowledgements

I thank God for this moment.

I also wish to thank everyone who has supported me on this journey.

This dissertation really started with my adviser, Professor Missy Cummings, inspiring me. Meeting her changed my trajectory for the better. Through her, I got to ride on my first freight train in my first month of my PhD! And this experience introduced me to the world of transportation systems.

I could not have done this great work without the people who welcomed me into the industry. From Canada to California, to DC. In particular, Chief Orlando Quiroz in Fort Worth and Mr. Mike Timpe in Portland, your hours and weeks devoted to helping me succeed has meant so much. I have great joy to have built something that your organizations value in return. I sincerely appreciate the new friends I have made in the railroad, airline, logistics, and flying car business. Thank you all.

I also value the investment I received from the Federal Railroad Administration via Mr. Mike Jones and from NASA Langley Research Center via Mr. Ken Goodrich.

The ground work on this PhD was not done alone. I have big hugs for all the students and associates who joined my "SHADO-AV HQ" enthusiastically! You all have been incredible research assistants and I wish to support you in your endeavors as well.

I thank my comrades in the Humans & Autonomy Lab and Duke Robotics. I also thank writing expert Rene Caputo who encouraged me even in the roughest of drafts.

And, to my dissertation committee-at-large, thank you for being available to listen to my progress and direct my path toward great research. Professors Glass, Garcia, and Zavlanos, each gave me timely advice and support. I also appreciated the strategic perspective of Professor Rich Burton. Dean Valerie Ashby, you didn't have to do anything, but you have done so much to encourage me in this challenging period. You have been a remarkable example of faith and perseverance. Thank you.

Mommy, my life began with you and I was so sad to lose you during my final semester of my PhD. You were the greatest motivator of my academic pursuits from Day 1. I love you and what you did for me on this Earth. You gave me siblings to lighten the load. To my siblings: I love you all. I am especially blessed by my sisters. Thank you, sisters, for joining my lab meetings, for taking the best notes in my classes, for making all the Duke parties fun, for reading almost everything I wrote, for listening to me practice my presentations, and giving the best feedback! I cannot express my gratitude to my mother enough for all of this, but I do thank God for the strength, courage, and wisdom she passed on to me. Mommy, our journey continues to infinity and beyond!

A special note to future students who may come across this work: you can totally do this! I wish to thank the many students who came before me and laid great templates of writing a dissertation. I hope my work will be of help to those who come after me too.

1. Introduction

1.1 Motivation

Railroad and airline operations are two major modes of transportation in the United States (Bureau of Transportation Statistics, 2016). Today, the standard is that trains have two crewmembers onboard operating the locomotive and planes also have two crewmembers onboard controlling the cockpit. Yet, nationwide, these onboard operators must interface with a traffic control agent to ensure safe separation from other vehicles in their vicinities. For commercial operations, the vehicles only move with the support of a dispatcher remotely managing the fleet (Berry & Pace, 2011).

The mission of a dispatcher in both settings is to maximize business efficiency while ensuring the safety of crews, passengers, and neighboring communities for their entire transportation network. In railroad operations, dispatchers also serve as the traffic controllers. Though dispatchers in railroad and airline operations may manage different types of fleets, they tend to use the same set of tools with computer aided dispatch systems, multiple displays and communication consoles. Dispatchers in the two domains experience similar work environments in terms of communications with personnel within and beyond their organization, including their chief dispatchers, crewmembers, and emergency officials.

The dispatch operations center is often thought of as the central nervous system of a transportation network. Yet, as the railroad and airline industries move to integrate new technology into their fleets, with increasing levels of automation and decreasing onboard crew size, little work has been done to understand how these changes may interact with dispatch operations.

The U.S. railroad industry has a long history of introducing new technologies to meet evolving demands. Innovations such as advanced signaling technology and automated methods for tracking and distributing cars on complex networks of tracks have helped to improve safety by deterring many accidents. They have also facilitated smaller crew sizes and increased workload on crews (Martland, 1982). However, human factors are the leading causes of train accidents in the United States (FRA, 2016). To improve safety, Congress mandated major railroads to integrate automation in the form of positive train control (PTC) by 2015 but the deadline has since extended into 2020 (110th US Congress, 2008). PTC acts as a back-up system for the onboard crewmembers by monitoring travel conditions and initiating emergency braking when the human operator fails to do so. At the same time, companies like GE Transportation have been promoting cruise control systems to become the industry standard for energy management (Brecher & Shurland, 2015).

Combining such systems could provide a form of autopilot in trains like those of self-driving cars. In fact, Asian-Pacific and European nations lead the world with over 75% of the fully automated metro lines today (International Association of Public Transport, 2016). Companies like Rio Tinto in Australia are investing millions of dollars into implementing fully autonomous railway systems to transport heavy freight over long distances (Peters, 2016).

In the airline industry, since Uber Technologies' 2016 announcement that the company would be getting into the air taxi business (Holden & Goel, 2016), a surge of interest, investment, and development has ensued. To date, the primary focus has been on the design of the electric vertical takeoff and landing aircraft, popularly referred to as "flying cars." While the vehicle design is a critical element of such futuristic systems, fleet management is another significant component that needs to be developed. Particularly, operations centers will be required to remotely manage fleets of air taxis. This creates a new demand for the dispatcher role.

On-demand air taxi operations will likely introduce increasing numbers of flights over shorter distances and times that could overburden pilots and ATC (Mueller, Kopardekar, & Goodrich, 2017) unless supported by newly defined dispatchers. How such operations should be supported by dispatchers has only recently been studied (Nneji, Cummings, Stimpson, & Goodrich, 2018). Revolutionary concepts of operations may require not just a drastic shift in customer expectations but also in the tasks and responsibilities currently allocated to airline dispatchers who remotely monitor flights and communicate with the pilots flying the aircraft. Moreover, as onboard autonomy increases and trained human resources onboard decrease, customer interfacing requirements present a new breed of functions that will draw inspiration from call center operators, flight attendants, and captains in how to communicate with passengers directly.

Previous research in human-automation interaction has demonstrated that additional automation does not necessarily guarantee increased system effectiveness or safety (Cummings & Ryan, 2013; Parasuraman & Riley, 1997). Often, automating a task within a larger system modifies the task by transferring the operator's workload from one physical or cognitive resource to another, thereby changing the task rather than eliminating it (Sebok, Wickens, & Laux, 2015). Poorly designed automation can contribute to errors and reduce system effectiveness due to implementations that increase cognitive workload.

Workload is of interest in dispatch operations as the job of dispatchers in transportation increasingly involves changing demands in cognition and decision making as low-level control functions are reallocated to machines (Singleton, 1989). Cognitive workload is the "level of attentional resources required to meet both objective and subjective performance criteria, which may be mediated by task demands, external support, and past experience" (Young & Stanton, 2005). While the importance of workload is clear, numerous challenges remain in objectively measuring workload (Wierwille & Eggemeier, 1993).

To this end, this research presents a task- and time-based workload modeling approach applied to both rail and airline dispatch operations. Developed using objective task-time data collected during observations of dispatchers and analytical data collected from interviews with subject matter experts (SMEs), the Simulator of Humans & Automation in Dispatch Operators represents dispatcher workload under numerous operating conditions. While simulation models have been routinely developed to explore such questions in aviation for pilot workload in the cockpit, there has been little work into extending such objective and quantitative approaches to the railroad industry and generally in dispatch operations. SHADO extends previous efforts to better understand how workload is affected by the introduction of automated technologies, as well as how these advanced technologies could or should inform operations design and staffing configurations. The ultimate objective of SHADO is to provide stakeholders with a tool that can be used to explore the human factors risks and opportunities in planning the future dispatch workforce and innovative system designs for different modes of transportation.

1.2 Research Questions

The way an organization is staffed and designed can affect how it performs. Modeling is a proven method of describing how organizations are staffed and designed today and experimenting with how changes to this structure might affect key performance indicators (Burton & Obel, 2018). In dispatch operations centers, the dispatcher role is paramount to performance.

Therefore, human factors requirements are needed to inform the design and staffing of dispatch operations centers that will manage fleets in railroad, airline, and future transportation systems. Concept development is necessary to engineer requirements and human-systems integration is a critical aspect (The MITRE Corporation, 2014). The systems engineering lifecycle begins with concept development and continues in oftentimes iterative loops through to maintenance and transition of systems (Figure 1).

Figure 1: Concept development, where this research effort fits in the systems engineering lifecycle framework.

Concept development entails performing an operational needs assessment, documenting concepts of operations, mapping out operational requirements, and defining the concept at a high-level. Modeling is a concept development method that allows stakeholders to identify and quantify the components, processes, and resources required to meet their human-system objectives. The work presented here provides stakeholders in dispatch operations across the transportation industry with a tailorable model to simulate dispatcher workload in networked transportation operations centers by answering three research questions:

1. How can we develop a workload model of dispatchers managing fleets in railroad and airline operations centers so that stakeholders can explore future concepts of operations?
2. What parameters in the model are most influential to dispatcher workload in these dispatch operations centers?
3. What are the limitations of the model and how generalizable is it for fleets in railroad, airline, and future transportation systems?

1.3 Thesis Organization

In the process of answering the three research questions in this dissertation, the Simulator of Humans & Automation in Dispatch Operations (SHADO) was developed. SHADO is a novel tool for rapidly prototyping fleet network operations to explore how different concepts of operations may influence the workload of dispatchers. This first chapter presented a motivation for this research found in paradigm shifts across major modes of transportation all in need of a method for exploring concepts of operations with consideration of human factors. In the next chapter, background into workload as it relates to dispatchers, and how it has been modeled in previous work is presented.

In Chapter 3, SHADO and the underlying mathematical model that allows us to computationally simulate operational settings are introduced. Then, in Chapter 4, the process of verifying and validating SHADO for use in real-world railroad and airline operations is explained. Through the process of internal verification, data, open- and closed-box validation, confidence was built in using SHADO to simulate multiple settings of two companies in industry. Sensitivity analyses were conducted to describe relationships between inputs, internal design and staffing parameters, and key performance outputs. This further validated SHADO as generalizable for applicability to multiple present-day and future transportation network operations.

In Chapter 5, SHADO is shown to be useful to railroad and airline stakeholders exploring questions about the future of their operations. Questions about the potential impacts of 1) automation installed in dispatch operations, 2) growing fleet size, 3) different training approaches to attention allocation, and 4) changing the way flights are staffed on dispatcher workload and performance were investigated using SHADO. Chapter 6 discusses the generalizability and limitations of SHADO and closes by reviewing the key contributions of this thesis research.

2. Background: Dispatcher Workload Models

In this chapter, the history of dispatcher roles and their current functions in both railroad and airline operations are presented. The present-day functions of dispatchers align in many ways between rail and air operations, but the dispatcher role may change in each industry as new levels of automation are integrated. A literature review was conducted on published work investigating rail and air dispatchers and how new functionalities of automation may affect the cognitive workload of dispatchers. Through this investigation, cognitive task analyses of the dispatcher roles regarding goals, work contents, interaction parties, and communication tools were identified and refined. By analyzing commonalities and differences between rail and air dispatch operations, key internal parameters of dispatch operations that could be generalizable across rail, air, and other operations were found. Findings from this research contribute to strategies that key stakeholders can employ to strategize for future human-system integration requirements leading to equivalent or better levels of safety in transportation.

2.1 The Dispatcher Role

2.1.1 Railroad Dispatch Operations

Railroad dispatchers have been around since at least 1851 when it was reported that a railroad manager issued a telegram using American Morse code on his telegraph to control the movement of trains in his territory (Hungerford, 1946). Since then, railroad dispatchers have taken on other modes of communication for supervisory control of railroad traffic as new technologies like the phone and radio developed. Even as the cab crew size has been reduced from about seven people to just two (Martland, 1982), typically a locomotive engineer and a train conductor, the dispatcher's role has remained paramount.

The Federal Railroad Administration (FRA) commissioned a report by Roth, Malsch, and Multer that was published in 2001 (Roth, Malsch, & Multer, 2001). They reported a thorough cognitive task analysis (CTA) on the railroad dispatcher role. The work included field observations at dispatch centers and SME interviews. Their results provided insight into the functions and demands of a dispatch operator and they concluded with recommendations on potential solutions to alleviate challenges that dispatchers revealed.

Dispatchers maintain a line of communication spurred from reports or requests of the cab crew. A single dispatcher may be accountable to more than one train at an instant and their actions ultimately impact the safety and efficiency of railroad operations. Roth et al. (2001) listed sources of input for a dispatch operator and found that the locomotive crew and fellow dispatchers were the primary source of up-to-the-minute changing information.

According to Roth et al. (2001), there are challenges that dispatchers face in satisfying multiple demands from maintenance-of-way (MOW) crew. Compared to the locomotive engineer who drives the train, the dispatcher must maintain a larger memory bank of tasks that must be revisited to handle an unanticipated request from the system that requires more immediacy. Dispatchers have a set of responsibilities to maintain supervisory control of their railroad track territory:

(1) Safe mainline train operations
    a. Adhering to operating rules
    b. Monitoring traffic to avoid conflicts
    c. Alerting train crews in cases of emergency
(2) Efficient routing for timely transit of passenger trains
(3) Routing all other trains passing through territory
(4) Safe scheduling of MOW crews

During irregular operations, particularly in dark territory, MOW, emergencies, or other exception-handling, there is a sequence of potential tasks that a dispatcher could follow. Dark territory are regions of railroads that do not have train detection systems or switch systems for dispatchers to remotely monitor and control traffic through their computer. In dark territory, dispatchers must vocally communicate with locomotive engineers, manually block tracks to authorize train movement, complete paperwork on movement permissions, vocally communicate movement permissions, and listen in for locomotive engineers onboard to confirm completion of train movement.

MOW workers are railroad construction crews. Railroad dispatchers vocally communicate with MOW workers are as well to permit track usage, block tracks to authorize their work, complete paperwork on work permission, vocally communicate the permission and again listen in for completion of work. Emergencies include unexpected events that disrupt regular operations that could be major, time-sensitive, safety-critical and require external services. In these events, dispatchers would be the point-of-contact for first responders and may director conversations with locomotive engineers and the railroad's troubleshooting desk. So, although there is a generally established schedule at the start of each shift, there is significant uncertainty in the schedule that must be managed by dispatchers.

2.1.2 Airline Dispatch Operations

Railroad dispatchers share some similar history and functionalities with airline dispatchers. The first instance of an aircraft dispatcher was recorded in 1920, when the United States Post Office required air mail service (Airline Dispatchers Federation, 2014). With increasing demands for safety as well as on-time performance, the challenges in air transportation safety were so great that Congress passed an Act in 1938 that established the dispatcher as a new official airman role. In 1964, the US Code of Federal Regulations, Part 121.533 defined the joint responsibility that dispatchers share with each pilot in command (PIC) of the aircraft (FAA, 1964). Their initial responsibility of coordinating aircraft navigation now expanded to include the legal right to be involved at all stages of flights, from preflight planning, to determining delays, and releases of flights, communicating with pilots in the sky, and re-planning to alternative landing locations. Aircraft dispatchers must monitor progress, broadcast required safety information, and decide the viability of each flight in their operational control.

Along with ensuring safety and regulatory compliance, airline dispatchers also consider business factors that influence customer experience and airline efficiency. During preflight planning, dispatchers must optimize fuel to maximize options for holding (often due to air traffic control ground programs at busy airports) and alternate landing as well payload limits for passengers and cargo while minimizing the total cost of fuel on that route. The dispatcher releases each flight plan for their pilots-in-command to review and it is ultimately recorded with air traffic control to ensure safe separation from other aircraft. Once the plane is approaching takeoff, the dispatcher monitors the flight using tools for tracking until the plane reaches its destination.

Monitoring the progress of flights is commonly categorized as flight following. Flight following accounts for the bulk of a dispatcher time spent on each flight. It is required for each phase of each flight by the FAA. This function is critical to rapidly responding and supporting flights when operations do not go as planned due to passenger medical issues, weather, mechanical failure, and major disruptions to airlines.

2.1.3 Comparison of Dispatch Operations

The dispatcher role in both rail and air operations generally involves 1) receiving directives from prior-shift dispatchers, chief dispatchers, and traffic control; 2) applying that information to the scheduling of vehicles and allocation of resources in order to meet customer objectives; 3) communicating with onboard operators to confirm trip conditions; and 4) coordinating with internal and external operational stakeholders to quickly resolve supply-and-demand and safety risks in unplanned situations (Federal Aviation Administration, 2013; Roth, Malsch, Multer, & Coplen, 1999).

Figure 2 shows that operations begin and end with an operations briefing to ensure a shared understanding of vehicles currently in service, relevant weather projections, and business demands. Then, each operator has a set of X vehicles they are responsible for monitoring and scheduling. Throughout their shift, they also allocate resources to meet any supply-demand needs in their network under management. Dispatchers are ready to quickly communicate and solve challenges that may arise in operations (Carrel, Mishalani, Wilson, & Attanucci, 2013; Devoe, 1974; Rahimi, Dessouky, Gounaris, Placencia, & Weidner, 2000).

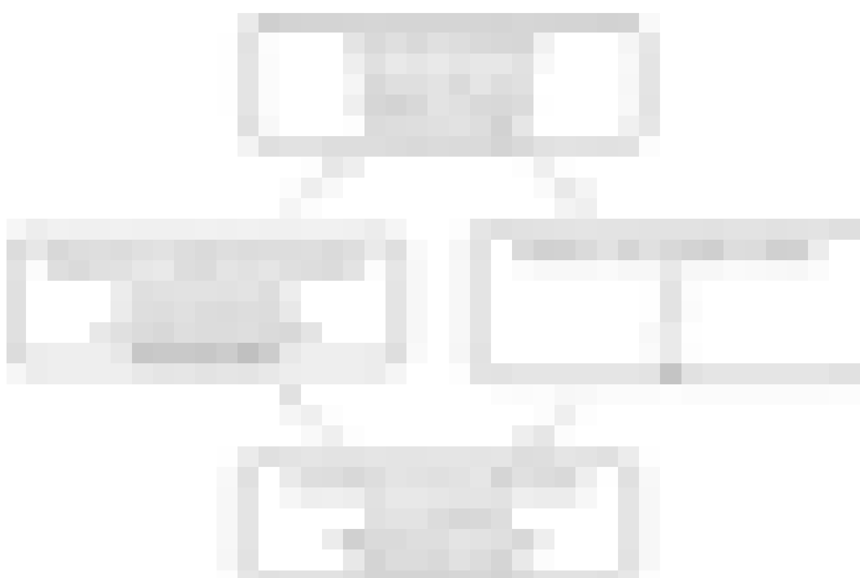


Figure 2: Schematic of dispatch center operations.

More similarities exist in how dispatchers in rail and air operations work when one considers differences in passenger and freight services**.** Passenger services adhere to stricter preplanned trip schedules. Companies know that not doing so may lead to higher workload in the form of phone calls to customer service from passengers inquiring. Therefore, in coordinating resources, such dispatchers keep time as an important metric of performance. On the other hand, freight services are often on-demand, leading to less predictable schedules.

Another shared distinction is on long- versus short-haul trips. The longer the trip, the more workload a dispatcher can expect over time due to changes in regulatory requirements and notices and bulletins for each region. Shorter trips at great frequencies may come with their own high task loads and this is largely due to the traffic for the route and any congestions at the ports (i.e. terminal yards/stations for railroads or airports for airlines) (FRA, 2016; Rosenhand, Roth, & Multer, 2012). The beginning and end of trips in transportation systems have been shown to yield higher chances of error in operations.

Differences between the two domains lie in access to information. Freight rail dispatchers often rely on generic weather reports. In dark territory, defined earlier, where there are no sensor-signal systems embedded on the railroad, dispatchers have low certainty of train locations. Conversely, passenger airline dispatchers have flight-specific weather projections and precise digital updates of flight status and location.

One environmental similarity of dispatchers at railroad and airline operations centers is that they are all surrounded by displays. This can be seen in the sample photos in Figure 3 and Figure 4. Over a decade since Roth et al.'s (2001) report, displays have expanded and upgraded. These have been reported in more recent observations of these technologies in railroad (Huang, Cummings, & Nneji, 2018) and airline operations (Nneji et al., 2018). "Changing technology" was found in an audit by the Federal Railroad Administration (1995) to be one of the sources of dispatcher stress. Stress has been associated with workload in the operational domain (Popkin, Gertler, & Reinach, 2001). Workload is a key consideration as to where and when to implement new technology into the work environment of dispatchers. In the next section, we will learn what workload means and how it can be measured.

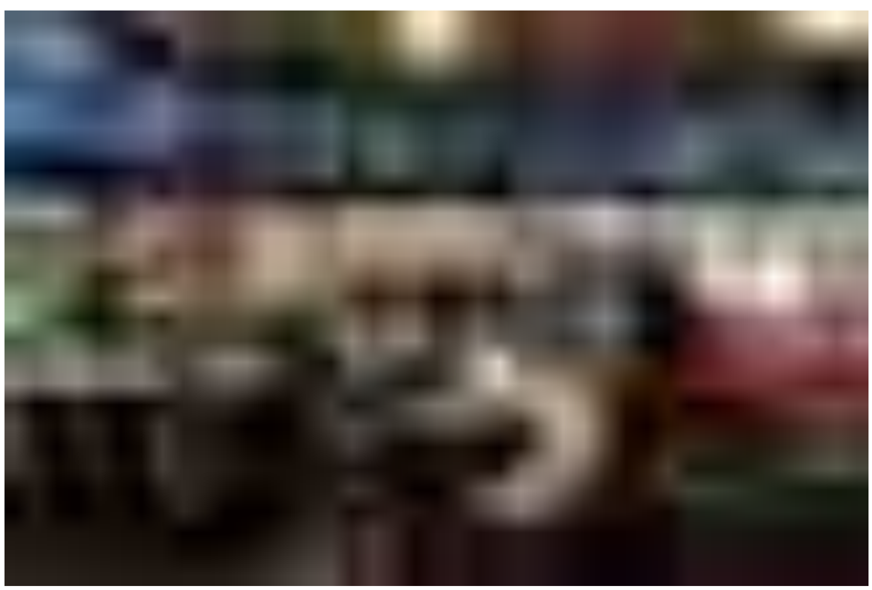


Figure 3: A railroad dispatch operations center

Figure 4: An airline dispatch operations center

2.2 Human Operator Workload

The workload of dispatchers in operations centers today is a key determinant of performance. Gertler and Nash (2004) defined dispatcher workload in their study by wait time, number of calls into dispatch, and number of official complaints. In this context, we define workload as the amount of cognitive resources employed for an operator to perform tasks (Senders, 1964; Wickens, 1979). Dispatcher workload is influenced by task load, system design, staffing levels, and operator behavior (Johannsen, 1979; Young & Stanton, 2005).

Human performance can be influenced by workload in that high workload can lower performance when the demands of tasks are beyond the cognitive resources available to respond. Performance may also be negatively impacted by low workload if tasks are too basic for an operator to be adequately engaged at work. Spending too much time with too much to do or too little to do can lead to decrements in performance. Similarly, experiencing sharp switches in workload from a 'valley' to a 'peak' into irregular operations abruptly can make an operator's immediate response to be inefficient and/or incorrect.

Task load is the demand required by the work environment. In dispatch operations, this demand often comes from the network include fleet and environmental factors. The FRA (1995) identified fleet size, task interarrival times, and fleet heterogeneity to be contributors to dispatcher workload. System design relates to how functions are allocated across components of humans and machines. Staffing levels determine which operators can address the tasks as they arrive into the system. Finally, operator behavior may express individual differences in one's approach to tasks.

Workload can be measured in qualitative and quantitative methods (Wickens & Hollands, 2000; Wierwille & Eggemeier, 1993) from questionnaires to biometrics. In a distributed lab of four universities in a space automation and robotics consortium with NASA's Johnson Space Center, a platform to quantitatively study effects of various system components and technologies on overall telerobotic task performance was designed. Neuromotor workload measures, as a function of time and parameterizations, were derived from the manual controller channels on the platform.

Workload is often thought of as a multidimensional concept (Lysaght et al., 1989; Rusnock, Borghetti, & McQuaid, 2015) which can be measured analytically/empirically and objectively/subjectively. The two dimensions are determined by the sources of information, such as whether objective data or expert opinions were used, and whether the measures were derived from opinion or objective analysis of tasks. Analytical measures infer workload per the researcher's understanding of the operational domain. Empirical workload measurements come from experience as data is gather through direct observations. Objective workload measurements are derived from factual data. Two objective approaches to measuring workload include computing the amount of work to do or computing the time to do the work. Subjective workload measures come from individual opinions.

Collecting operator opinions via self-reported questionnaires is a method that uses both subjective and empirical measurements of workload. The NASA-TLX (Hart & Staveland, 1988) is a well-known tool for subjectively gathering empirical data. Subject matter experts provide subjective and analytical estimates of workload. These methods are particularly useful in developing new concepts of operations in which operator performance on tasks is not yet known.

Physiological measures are object-empirical approaches to estimating operator workload by using biological feedback such as heart rate. As sensitivity as these approaches may be to cognitive workload, their requirements are too intrusive for dispatch operations at this time. Objective-analytical measures incorporate task, environment, and individual characteristics as inputs into mathematical models to quantify estimated workload. A mathematical model would provide use with a tool to consistently calculate human operator workload across different conditions and time scales (Rusnock et al., 2015) that are applicable to dispatch operations.

Work in human supervisory control (Cummings & Nehme, 2010), human-machine interaction (Rouse, 1983), air traffic control (Schmidt, 1978) have supported time-on-task as a reliable estimation of workload. Considering the safety-critical nature of transportation operations, in which network demands fluctuate over periods of time, a time-based measure is most appropriate (Donmez, Nehme, & Cummings, 2010). Wierwille et al. (1993) suggests that time estimation can reveal extreme regions of workload that may be missed in other assessments.

Cummings & Guerlain (Cummings & Guerlain, 2007) used a utilization metric defined as percentage of time an operator is busy out of total operation time as an objective workload measure. Results consistently showed that at utilization levels above 70%, operator performance declines. Some suggest similar results at utilization levels below 30%. These studies support the follow the Yerkes-Dodson theory of the impact of operator workload on system performance (Cummings, Gao, & Thornburg, 2016; Cummings, Mastracchio, Thornburg, & Mkrtchyan, 2013; Cummings & Nehme, 2010; Rouse, 1983; Yerkes & Dodson, 1908). The parabolic performance curve as a function of workload hypothesizes that when operators experience too much or too little workload, performance decreases (Figure 5).

Figure 5: The Yerkes-Dodson inverted-U theory.

Dispatchers may experience high cognitive workload when they meet unplanned demands and must dynamically re-plan around issues. Therefore, conditions and factors that influence workload were studied and four areas identified below (presented in Figure 6), fleet conditions, strategic design factors, tactical staffing factors, and operational factors were found to be potentially the most relevant to dispatcher workload in dispatch operations centers.

Figure 6: Conditions and factors that may influence remote operator workload.

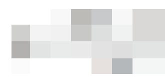
In dispatch operations, task load can be represented by fleet and environmental factors. System design can be represented by dispatcher decision support systems available to assist dispatchers. Staffing levels can be represented by the team structure and schedule. Finally, operator attention allocation strategies are engrained in the operator behavior module. The key elements of dispatch operations can be categorized in these four areas, discussed in detail below.

2.2.1 Task Load

Task load is the demands required by the system. Task load can be measured by the number of tasks over time. In some studies (Cummings, Bertucelli, Macbeth, & Surana, 2014), task load was represented by number of vehicles. Several experiments have been undertaken to study just how many unmanned vehicles (UVs) a single operator can manage (Cummings & Mitchell, 2008; Cummings, Nehme, Crandall, & Mitchell, 2007; Dixon, Wickens, & Chang, 2005; Ruff, Narayanan, & Draper, 2002). One found that in an experimental military missile mission setting, an operator can supervise up to 12 homogenous UVs (Cummings & Guerlain, 2007). From the perspective of remote operators, managing larger fleets may slow their performance. Studies of air traffic controller workload came to similar conclusions that the total time remote operators spent on tasks were increased as the number of aircraft in their sector increased (Majumdar & Ochieng, 2002; Mogford, Guttman, Morrow, & Kopardekar, 1995).

What differentiates task load from workload is the operator’s capacity to respond to demands. Researchers have applied the fan-out principle from digital electronics, as shown in Equation 1, to modeling the capacity number (N) of robots a single operator can manage simultaneously (Cummings & Mitchell, 2005; Olsen & Goodrich, 2003; Olsen & Wood, 2004) before reaching their limits of cognitive resources.

Equation 1: Basic fan-out theoretical basis



In the fan-out equation, the level of robot autonomy is represented by its neglect tolerance, or time, (NT) and interaction time (IT) (Crandall, Goodrich, Nielsen, & Olsen, 2005; Olsen & Goodrich, 2003; Olsen & Wood, 2004; Steinfeld, Fong, & Kaber, 2006). Robots with low levels of autonomy require more attention and interaction time from operators so an operator would have a low fan-out. The more the robots can operate on their own and be neglected while maintaining performance, the higher the fan-out, or number of robots an operator can manage. However, this fan-out equation fails to consider the impact of operator situation awareness and robot heterogeneity. Wait time due to loss of situational awareness (WTSA) is an important factor since humans, unlike machines, may not be 100% attentive and ready to respond due to natural tendencies like boredom, distraction, or being occupied by other tasks. When WTSA was considered, results showed that operator capacity was originally over-estimated by one- to two-thirds (Cummings & Mitchell, 2008).

Another fan-out formula was proposed to account for the case of managing robots with heterogeneous levels of autonomy, NTs, or a single robot with heterogeneous operator ITs (Crandall et al., 2005; Goodrich, Quigley, & Cosenzo, 2005). If the neglect time for any robot is greater than or equal to the sum of interaction time with the other robots, then it is feasible for a human to manage that team. Otherwise, it is infeasible per the formula. In addition to heterogeneity across vehicles, we can represent the heterogeneity in task types within vehicles too (Mekdeci & Cummings, 2009; Nehme, 2009; Nehme, Mekdeci, Crandall, & Cummings, 2009). Pfleiderer (2005) suggests that the mix of the aircraft is a contributing factor to air traffic controller workload.

Along with fleet size, levels of vehicle autonomy and heterogeneity, the level of network autonomy is another important characteristic that may influence remote operator workload. Network autonomy relates to how well vehicles can collaborate with one another (Cummings, 2004; Parker, 2008). Many researchers have presented the benefits of reduced reliance on a central planner when vehicles collaborate and are able to perform decentralized task allocation (Alighanbari & How, 2006; Choi, Brunet, & How, 2009; Pavone, Bisnik, Frazzoli, & Isler, 2007). Robots sharing information with each other can improve network capabilities and efficiencies (Burgard, Moors, Fox, Simmons, & Thrun, 2000; Howard, Mataric, & Sukhatme, 2002). In fact, lack of inter-vehicle communication has been shown to be detrimental to system performance in a tactical mission setting with unmanned aerial vehicles (Chandler, 2004). Wang, Wang, and Lewis (2008) found that operations with high network autonomy reduced demand on dispatchers and increased performance. The percent of time an operator allocates toward sharing information with other robots (occupied time, OT) out of the total available time in between controlling a lead robot (NT) is coordination demand, or task load required for managing the network of robots. With unmanned aerial vehicles (UAVs), workload has been shown to be less when managing decentralized networks with high autonomy than centralized or with low autonomy, particularly when humans and machines are collaborative (Cummings, 2015).

Higher level autonomous vehicles could share data amongst each other in the network while lower level vehicles with onboard operators would rely on limited network radio channels (e.g. pilot-to-pilot frequencies) or communications from the dispatcher to support their operational decision-making (Fries, 2008). In SHADO, this may manifest as reduced frequency of tasks of a type to the dispatcher from semi- to fully autonomous vehicles if any vehicle in the network has had that type of task addressed by a dispatcher in the operations center. In this way, vehicles can communicate issues and solutions to other vehicles to improve efficiency.

Finally, the environment is a key contributor of exogenous events that the remote operators have no control over but must respond to for their system to perform as required. One component of the environment is weather, which has been found to impact remote operator workload (Mogford, Murphy, & Guttman, 1994), by increasing the task load with frequency of tasks arriving to be handled (Koetse & Rietveld, 2009).

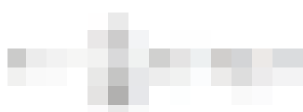
Weather can cause delays in operations and lead to higher rates of issues to be resolved by remote operators when vehicles become inoperable. Therefore, it is an important factor to be considered here. Task load largely comes from factors beyond the dispatch operations center including fleet size, fleet heterogeneity, fleet autonomy, and environmental factors. However, there are other factors that may influence dispatcher workload from within the operations center. These factors form the design and staffing parameters.

2.2.2 System Design

Two key elements of system design which affect workload in dispatch operations may be the levels of team coordination and artificial intelligence support. Coordination is the process of supporting the different efforts of others toward a common higher-level goal (Mohammed & Dumville, 2001). It can be represented by communication processes. To coordinate, teams use different modes of communication including face-to-face, call, and text messaging (Bowers, Jentsch, Salas, & Braun, 1998; Gorman, Cooke, & Winner, 2006; Langan-Fox, Code, & Langfield-Smith, 2000).

Depending on the mode of communication, the cognitive demand of transmission and receipt may vary (Salas, Diazgranados, & Lazzara, 2009). For example, the information shared from a call is likely to be ephemeral in memory and must be responded to within a short period or the call must be reattempted until the operator is available or able to disrupt a lower priority task. On the other hand, a text message is more evergreen in memory and can wait in queue for the receiving operator's availability to act on the information (Endsley, 1995; Posner, Nissen, & Klein, 1976). There is some cost to communication since operators are taking time away from other tasks but it may be a great investment by adding value to the efficiency and quality of service after using the shared information (Driskell & Salas, 1992; Wilson, Salas, Priest, & Andrews, 2007).

Although dispatchers may have an upfront cost in increased workload from time spent on communication tasks within the team, their performance in terms of speed on future tasks should be improved. Neth, Khemlani, Oppermann, and Gray (2006) identified a relationship between time-on-tasks and entropy (Equation 2). Through experiments, they found that operators who prioritized tasks at a higher degree, and therefore exhibited lower entropy, were more likely to perform better and to have had more experience in the synthetic task environment.



Equation 2: Entropy (H) as a function of time (t) on task (i) for n tasks.

Artificially intelligent (AI) support can also come in different forms (Sycara & Lewis, 2004). By design, AI could function as (1) an equal member completing the same set of tasks as humans on the team, (2) a supporting assistant to a single operator's task load, or (3) a supporting assistant to a team's task load. How information is presented from the AI agent and how the interaction interface is designed facilitates human-to-agent communication. In an experimental setting, multi-AI agents functioning as supporting assistants to teams was found to potentially mitigate states of cognitive overload attributed to team coordination and to improve team performance by increasing SA of fleet conditions (Sukthankar, Sycara, Giampapa, & Burnett, 2009).

2.2.3 Staffing Levels

Other elements in dispatch operations are based on decisions by operational managers which may help mitigate workload. Shift schedules, team size, and team expertise may all influence how dispatchers experience the task load. Operator shift schedule is such an important factor that the US Congress gave authority to the US Department of Transportation (DOT) to set new rules to establish limitations on hours-of-service (Federal Railroad Administration, 2012). The DOT went further to require that railroad companies use models to support their scheduling decisions. Gertler and Nash (2004) found that railroad dispatchers, particularly those who worked irregular or night shifts suffered from fatigue which influenced their workload and led to negative performance. The authors defined workload by wait time, number of calls into dispatch, and number of official complaints. Operators have shown to spend more time planning and unnecessarily repeating or prolonging tasks when experiencing mental fatigue (Linden, Frese, & Meijman, 2006). Fatigue is an important human factor that influences workload and has been shown to deplete the level of attentional resource available for operators to attend to tasks (Roach, Fletcher, & Dawson, 2004).

In a transportation setting, lapses in attention can be especially costly (National Transportation Safety Board, 1999). Experiencing fatigue can lead operators into dangerous states prone to human error from delaying, neglecting, or incorrectly performing time-sensitive tasks (Dorrian, Roach, Fletcher, & Dawson, 2007; Graeber, 1988; Pollard, Sussman, & Stearns, 1990). Hursh, Redmond, Johnson, Thorne et al. (2004) found that cognitive performance capacity decays over consecutive hours of wakefulness.

Working in a team may alleviate the burden on individual operators. Research has shown that larger teams lead to faster performance but the more people required for consensus slows performance (Bergum & Lehr, 1962). Foundational experiments have included up to five operators (Waag & Halcomb, 1972; Wiener, 1964). Additional teammates must be contacted and depending on the mode of communication, there may be some lost information with each degree of separation from the original source of information. How operators share tasks is just as important.

The way tasks are distributed in a team is an important aspect of team composition (Cooke et al., 2003; Gao & Cummings, 2014; Lewis et al., 2011; Naylor & Dickinson, 1969). Tasks may be distributed in a range from organically to mechanistically based on team expertise (Burns & Stalker, 1994). Organic teams are ones in which every operator has a generalist background to address most types of tasks that come into the system. Mechanistic teams are the opposite in that they are composed of individuals who each have a specialist role to play in the organization. Air traffic controllers operate as mechanistic teams since they have sectored regions that each controller oversees (Gao, 2013; Lewis, Polvichai, Sycara, & Scerri, 2006).

Some studies found that mechanistic teams operate more efficiently than organic teams (National Research Council (U.S.), 1993). However, organic teams may be more resilient during irregular conditions since they are able to reallocate tasks to balance workload and minimize task wait times (Barnes et al., 2008; Burns & Stalker, 1994; Porter, Gogus, & Yu, 2010). In the case of mechanistic teams, there is a chance that one operator is overloaded when certain types of tasks arrive in the system that only s/he is skilled to handle (Mekdeci & Cummings, 2009). So, understanding the demand-and-supply of remote operators is key to deciding how to organize teams (Macmillan, Entin, & Serfaty, 2004) in complex dispatch operations centers.

2.2.4 Operator Behavior

Operator behavior is the final key element in dispatch operations. Operators may follow an attention allocation strategy (Crandall & Cummings, 2007b; Mekdeci & Cummings, 2009). This is an individual human factor that is a controllable behavior and trainable skill (Gopher, 1980), unlike other important factors such as age (Kirby & Nettelbeck, 1991; Verwey, 2000). A set of rules define how s/he prioritizes tasks. Some operators handle tasks in chronological order (first-in-first-out, or FIFO) or based on how critical the incoming tasks are to the system (Pinedo, 2012).

Depending on the strategy, the wait time of tasks in the operator's queue and how frequently the operator switches tasks are impacted (Crandall & Cummings, 2007a, 2007b; Cummings & Mitchell, 2005; Goodrich et al., 2005; Nehme, 2009; Squire, Trafton, & Parasuraman, 2006; Wang & Lewis, 2007). Therefore, testing out the implications of the micro-strategies under different conditions can yield positive returns in macro-performance (Sheridan & Tulga, 1978).

2.3 Modeling Workload

Modeling workload may help operations center planners with human-system integration requirements for designing and staffing centers that perform well as the fleets change. To predict dispatcher workload, particularly in presence of new technologies, a model would be useful for rapid prototyping of concepts of operations. The model could provide stakeholders with a tool that supports staffing and design decision-making with strategic consideration of human factors.

Researchers have used modeling to study how conditions and factors influence operators. However, the prevailing modeling techniques differ by operational context and there are few models that holistically model the human-system interaction with multiple modules in operations. In railroad and airline operations, today, we can find call centers, traffic management and dispatch centers staffed with customer service advisors, traffic controllers, and transportation dispatchers, respectively. This research considers how remote operators are modeled in each of these contexts to select the most fitting modeling technique for current and future operations. Queueing analytic and discrete event simulation are two notable techniques for modeling.

Queuing analytic (QA) models prevail in call center research (Aksin, Armony, & Mehrotra, 2007) for phone service modeling with goals typically to increase call operator speed and relevance of response to customer requests. A queuing system is composed of arrivals and service entities, assumed to follow some distribution. There are different categories of queuing models: single server, multiple server. Within each, there can be a finite or infinite queue length. The latter assumes that there is no restriction on the number of callers that are waiting. Some call centers with the automatic call distributor that limit the number of calls to the number of trunks and ask callers if they would like to be called back once an advisor is available. Additionally, models may be differentiated by whether they have an infinite customer population. In the case of fleet management, the number of vehicles under management is constant and therefore finite however, there may be a situation where each vehicle may present infinite tasks. Balking, reneging, and jockeying are commonly associated with queuing models where customers can observe and decide to discontinue in the line they would have otherwise waited in for service.

As shown in Figure 7, some callers abandon, others call back, others wait in a queue if up to K spots are available while N operators are busy, or they get assigned immediately to one of up to N available operators (Garnett, Mandelbaum, & Reiman, 2002). These calls are communication tasks which generate additional tasks for the operator as they serve customers (Green, Kolesar, & Whitt, 2007).

Figure 7: Schematic of call center operations

The queueing analytic model can be represented as *M/M/s*. It assumes an *M* (Markovian, memoryless property) Poisson process of arrival rate, exponential service distribution times, and *s* servers. QA models can be solved using closed-form mathematical formulas (Hall, 1990; Lawrence & Pasternack, 2002). They are used to answer questions like what the probability is of a certain number of calls waiting in the system at steady-state. Additionally, for performance metrics, the length of the system, length of the queue, waiting time in system, and waiting time in queue are values to compute. However, QA models rely on several assumptions that would not hold in real ROC settings.

One assumption, that arrivals must always follow a Poisson distribution, would mean that the model would not be able to reflect how call arrival rates may vary with time or how one event may trigger another to arrive in the system (Green, Kolesar, & Svoronos, 1991). The analytic method would also fail to model events that arrive on a schedule. Unlike customer service advisors who primarily respond to issues that disrupt independent operations, traffic controllers and transportation dispatchers also determine the flow of somewhat dependent operations.

The traffic controller role involves monitoring positioning of vehicles and re-routing them as needed to maintain safe separation based on sectors and stages of trips. The dispatcher role generally involves receiving directives from prior-shift dispatchers, chief dispatchers, and traffic control; applying that information to the scheduling of vehicles and allocation of resources for vehicles to complete their individual missions of meeting customer objectives; communicating with onboard operators to confirm trip conditions; and coordinating with internal and external operational stakeholders to quickly resolve supply-and-demand and safety risks in unplanned situations (Federal Aviation Administration, 2013; Roth et al., 1999).

Call advisors are less dependent on network behavior than traffic controllers and dispatchers who rely on it to determine the best plan of action for each individual vehicle under their management. Additionally, once an issue is resolved per a call, an advisor may never have to interact with that vehicle again and a customer retrial may route to a new advisor who must re-assess the situation. Dispatch, traffic control, and call centers exist today, and some aspects of these conceptual models can be applied when designing and staffing future ROCs for transportation systems.

Whereas air traffic controllers control via voice communications with onboard human pilots, Sheridan (1989) published an extensive review on the field of tele-robotics and defined human supervisory control of machines which dates as far back as the 1960s. An opened control loop allows for a vehicle or fleet of vehicles to request service from remote operator and allows that operator to respond with high-level commands to redirect the vehicle within certain parameters. This maximizes the value of the remote human as an experienced advisor in responding to new conditions and the value of the low-level vehicle sensor-actuator system in performing motion tasks.

Laughery (1999) found that much of the performance of a complex system is quite dependent on the performance of humans operating the system. Humans are not as predictable as machines and are prone to unanticipated error due to the vast amounts of data that humans' sense and process, not just of the present task at hand, but of irrelevant life experiences and history. Usability heuristics that attempt to guide systems designers into designing for humans are not so valuable in quantifying the impact of design changes to the overall system performance. Performing empirical experiments to collect this data is risky and costly in safety-critical and time-sensitive environments like dispatch operations centers. Laughery discussed methods to computationally model human performance which align with modeling other components of the system. Laughery presented discrete event simulation as a viable method in the form task network modeling.

Discrete event simulation (DES) models have been successfully applied to model operator workload in supervisory control domains including air traffic control (Humphreys, 1998; Loft, Sanderson, Neal, & Mooij, 2007; Majumdar & Polak, 2001; Schmidt, 1978; Tewes, 1999) and single operator control of multiple unmanned vehicles (Donmez et al., 2010; Nehme, 2009; Nehme et al., 2009). They have also been used in call center planning (Lam & Lau, 2004; Mazzuchi & Wallace, 2004). Recently, DES has been extended to teams of operators in military command and control settings (Gao & Cummings, 2012). A DES was developed and used to analyze how automation and crew size could impact the workload and performance of onboard train crew (Nneji, Cummings, & Stimpson, 2019).

Using DES, the effects of team coordination can be positive and modeled by a decrease in service time and chance of error on future tasks of the same type (Gao, Cummings, & Solovey, 2014). For example, the expected frequency and duration of communication tasks a dispatcher may handle would be drawn with a probability from a distribution based on patterns observed in several related real-world conditions.

Mekdeci and Cummings (2009) used DES to model multiple human operators in supervisory control of heterogeneous UVs for a search-and-rescue mission. They found that although mechanistic teams, as described in Section 2.2.3, had overall better performance than organic teams, the organic teams could maintain workload levels during emergent irregular conditions.

Another set of researchers focused on the issue of workload and crew size per shift in nuclear plants. Plott, Scott-nash, Hallbert, and Sebok (1995) approached this issue with task network modeling using a tool for discrete event simulation, Micro SAINT. The systems engineer could use this DES tool by inputting parameters, running the simulation, and receiving a final snapshot output file of information as shown in Figure 8 below.

Figure 8: Abstracted computer model of operating crew to aid in crew size analysis

The task network model extends a sequence of tasks and creates closed loop representation of the human-machine interaction. Designers can use this model to estimate how changing certain processes may impact the time on task. Each task can have associated means and standard deviations of timing, system state requirements and effects of completion. Depending on the sensory requirement of the task, some may be handled simultaneously by a single operator (Wickens, 2008).

Drews, Laughery, Kramme, and Archer (1985) used DES to determine the helicopter crew staffing requirement for the Army. The organization wanted less people on board with heightened technological automation. However, the analysis of crew workload led to the conclusion that reducing crew size would not yet be safely possible. Time and again, others (Fontenelle & Laughery (1988); Plott et al. (1995); Allender et al. (1995)) contributed to developing a method for estimating individual and crew workload by building workload estimation metrics into the DES model. Embedding a representation of a theory of workload (Wickens, Sandry, & Vidulich, 1983) within the task network, researchers were able to find areas of extreme workload which may cause delay or error. This method has been used to study where tasks may be reallocated to other operators, or to show where systems have too high of an expectation for humans.

2.4 Thesis Research Area

Workload modeling allows for operational system variables and effects of changes to those variables to be compared. There are different methods to model workload, some with higher fidelity than others and some application-specific while others robust enough to be generalizable. Discrete event simulation (DES) is one such method that represents humans operating in the loop of a complex process on a high-, system-level frame. It has been used for years in modeling manufacturing and other time- and safety-critical services—allowing for designers to realize areas that cause undesired delays or bottlenecks in the system.

DES has been used in some studies tangentially related to dispatch operations. When a discrete event simulation was implemented to model an operator supervising multiple heterogeneous unmanned vehicles (UVs) (Nehme et al., 2009), the simulation revealed significant and opposite effects of fleet size and autonomy level on system performance. Much of the workload modeling efforts in this area have been in a military context. However, as demands increase for remote human operators in transportation domains, the industry needs a rapid prototyping tool for considering human factors in staffing and design decisions. The focus of this thesis research is to design the first tool of its kind for modeling dispatcher workload that can be generalizable to different operational settings.

Discrete event simulation models involve queuing-based constructs including events, arrival processes, service processes, and queuing policies to model the dispatch operator as a serial processor of tasks. For this thesis research, the level of task performance is measured and modeled at the level of functional allocation. The input variables are primarily the timing of various dispatcher tasks (set up as arrival rates) and the distributions of task durations (set up as service times).

This chapter covered a background of the role of dispatchers in related domains, the meaning and measures of workload, as well as methods for modeling workload. Key elements of remote operator workload from task load, system design, staffing levels, and operator behavior were presented. In Chapter 3, these key elements will be connected as internal parameters in the conceptual and computational discrete event simulation model of dispatcher operations.

3. Model and Simulation of Humans & Automation in Dispatch Operations

In Chapter 2, the key elements of dispatch operations centers were identified as they pertain to the importance of operator workload. In this chapter, a conceptual model of humans and automation in dispatch operations is introduced that considers these key elements to describe the workload of a dispatcher and test hypotheses of factors that influence workload under numerous operational conditions. A discrete event simulation (DES), the Simulator of Humans & Automation in Dispatch Operations (SHADO), is specified to computationally model the workload of dispatchers. SHADO was designed to represent dispatch centers that exist today as well as serve as a platform to explore future concepts of operations.

In the next section, the key elements of dispatch are abstracted and connected into a structure of internal parameters that affect workload. Then, the DES method and input-output architecture are introduced. The flow of processes in SHADO is presented. Finally, the usefulness and limitations of simulation are discussed.

3.1 Conceptual Model

Dispatcher workload impacts an operations center's ability to respond to emergent network behavior. Therefore, the conditions and factors that were found to influence workload in dispatch operations were categorized to design a general model of dispatcher workload as a basis for SHADO. The four categories are fleet conditions, strategic design factors, tactical staffing factors, and operational factors.

Fleet conditions include the fleet size, fleet heterogeneity, fleet autonomy, and environment. These conditions all influence the frequency and volume of tasks that arrive into dispatchers' queues to which to respond. The more vehicles in a fleet, the more tasks one can expect from the fleet. Fleet heterogeneity, the diversity in types of vehicles under management, can mean that some vehicles require more attention than others due to the frequency of their requests to dispatch. On the other hand, greater fleet autonomy results in fewer overall requests to dispatch due to vehicle-to-vehicle communications. Environmental conditions are uncontrollable but can cause irregular operations that require dispatchers to rapidly respond to contingencies.

Within an operations center, decisions made on a strategic, tactical, and operational level can ultimately influence how efficient dispatchers can work. Strategic decisions are long-term investments made by center directors in office design to facilitate team coordination and installation of new technologies for artificial intelligence support. Coordinating work in a team means that each dispatcher may spend more time on internal communications tasks but less time on other tasks due to the assistance of teammates. In the future, support may also come in the form of artificial intelligence. Artificially intelligent decision aids can serve as equal teammates, task assistants, and/or team coordination assistants.

Tactical decisions are short-term plans prepared by operations managers on shift schedules, team size, and team expertise. The longer dispatchers are on duty, the more time it may take them to complete tasks due to increasing fatigue. However, the inefficiencies that come from spending more time on each task can be better managed if dispatchers are staffed to share work with each other, such that no task ever waits for too long before it is handled by a qualified dispatcher. Managers can decide to staff teams of generalists or specialists, or some mix of expertise. Generalist teams are made up of dispatchers that can handle any type of task. In specialist teams, dispatchers are rigidly assigned to handle specific types of tasks based on their expertise.

Finally, operational decisions are made instantaneously by each dispatcher about how they allocate their attention when faced with multiple demands. Some dispatchers may use a first-in-first-out approach while others prioritize work in some way. This factor is one that is trainable.

3.2 Computational Model

Whether these ten factors and conditions increase or decrease dispatcher workload depends on the input parameters of the operational settings. To investigate the impact of such changes in the complex dispatch operations center system on dispatcher workload, a computational method is required. In this work, discrete event simulation was selected as the computational method for developing a workload model of a remote dispatcher.

In this research, the level of task performance measurement is important to ensure that workload is modeled at a granularity that stakeholders can make operational decisions with confidence. The level of task performance was decided by identifying key high-level functional requirements of dispatchers. For example, airline dispatchers' three main functions are to plan flights, monitor flights, and be ready to respond to contingencies in irregular operations. Task performance is measured and modeled at the level of functional allocation. The input variables are primarily the timing of various dispatcher tasks (set up as arrival rates) and the distributions of task durations (set up as service times). For example, the expected frequency and duration of communication tasks a remote dispatcher may handle would be drawn with a probability from a distribution based on patterns observed in several related real-world conditions.

The DES model framework of SHADO is represented in Figure 9. In this schematic, each vehicle in a fleet has one event stream and may have an event active in the system at any instant. Additionally, the environment presents exogenous events such as poor weather. Each event that arrives waits in the queue for up to T time. The event is serviced by a remote operator (dispatcher) or expires from the queue if it has not been started by T. Event arrivals are represented by a λ_i process, typically random, for each event stream i. Environment and fleet settings impact those inter-arrival times.

Dispatchers handle events with service process μ, typically random, that may be influenced by shift and team settings. Team expertise influences how tasks are assigned to dispatchers, and attention allocation influences which task a dispatcher handles next. Figure 9 is a notional descriptive model and depending on operational settings, there could be one or many more dispatchers at work with different team structures beyond the example pictured here. For example, one desk of fleets could be staffed by three operators who share workload. On the other hand, a single operator may receive assistance from artificial intelligence that takes the role as an equal operator that can process the same tasks humans can, perhaps at a faster speed and to a greater accuracy. How these design and staffing parameters interact with the logical flow of events is further detailed below.

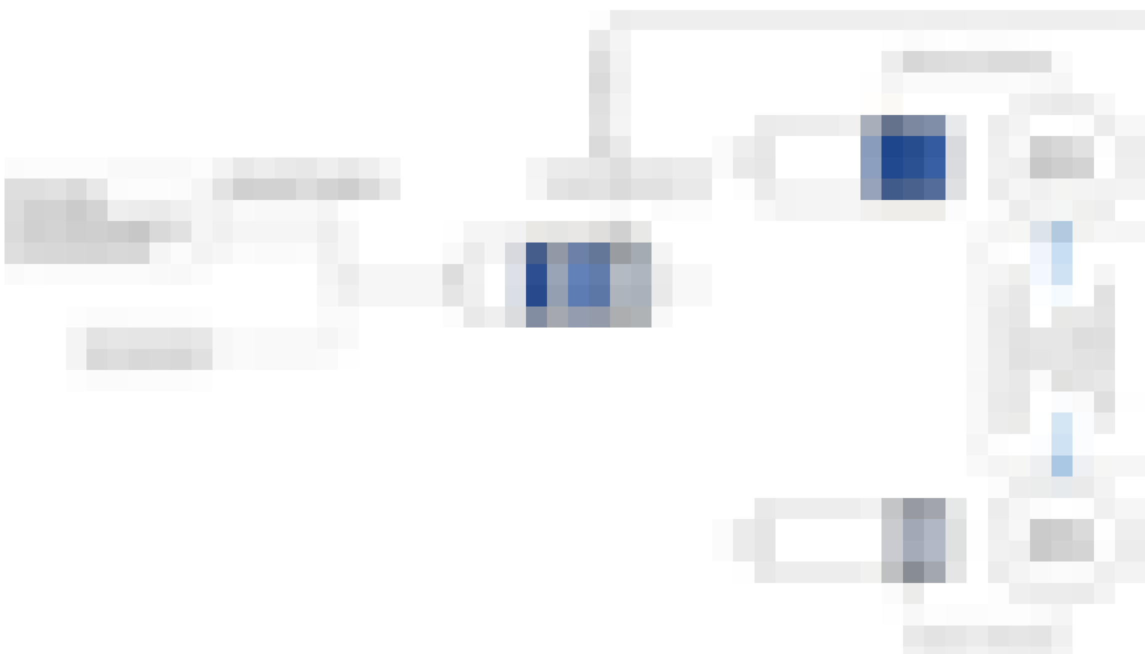


Figure 9: Descriptive model of how the internal and input parameters fit in the discrete event simulation of SHADO

In SHADO, a simulation of dispatch operations initializes with the user-defined input parameters as mapped in Figure 10, including the number of replications. Over the next few pages, we will follow the process of a simulated shift including processes of the tasks that arrive from the system and ultimately influence operators' workload. A replication represents a different day that takes random draws from the same shift parameter distributions. A higher number of replications allows for the user to view a wider slate of possible outcomes based on probabilities. Each new day may have some variation in the timing of transfer-of-duty periods, timing of tasks as well as any extreme conditions a user simulates. For example, if a transfer-of-duty period is set to be complete by dispatchers on a uniform probability density function between 5 and 10 minutes, the first day it could take 9 minutes but on the second day it could only take 6 minutes.

As a day's shift begins in Figure 10, a task enters the system given an associated probabilistic distribution and then is assigned to a remote operator's queue. The task awaits dispatcher availability and then leaves the queue to be processed once the dispatcher begins the task. While being processed, the task may be interrupted by another task of higher priority and thus returned to wait in queue until the operator is available. Finally, the task exits the system. At any point in this process, the task may expire before completion, at which point it exits the system prematurely. The task may also go unfinished if the last transfer-of-duty begins or the shift ends. This expiration feature can be used to model certain types of phone calls which practically do not ring endlessly.

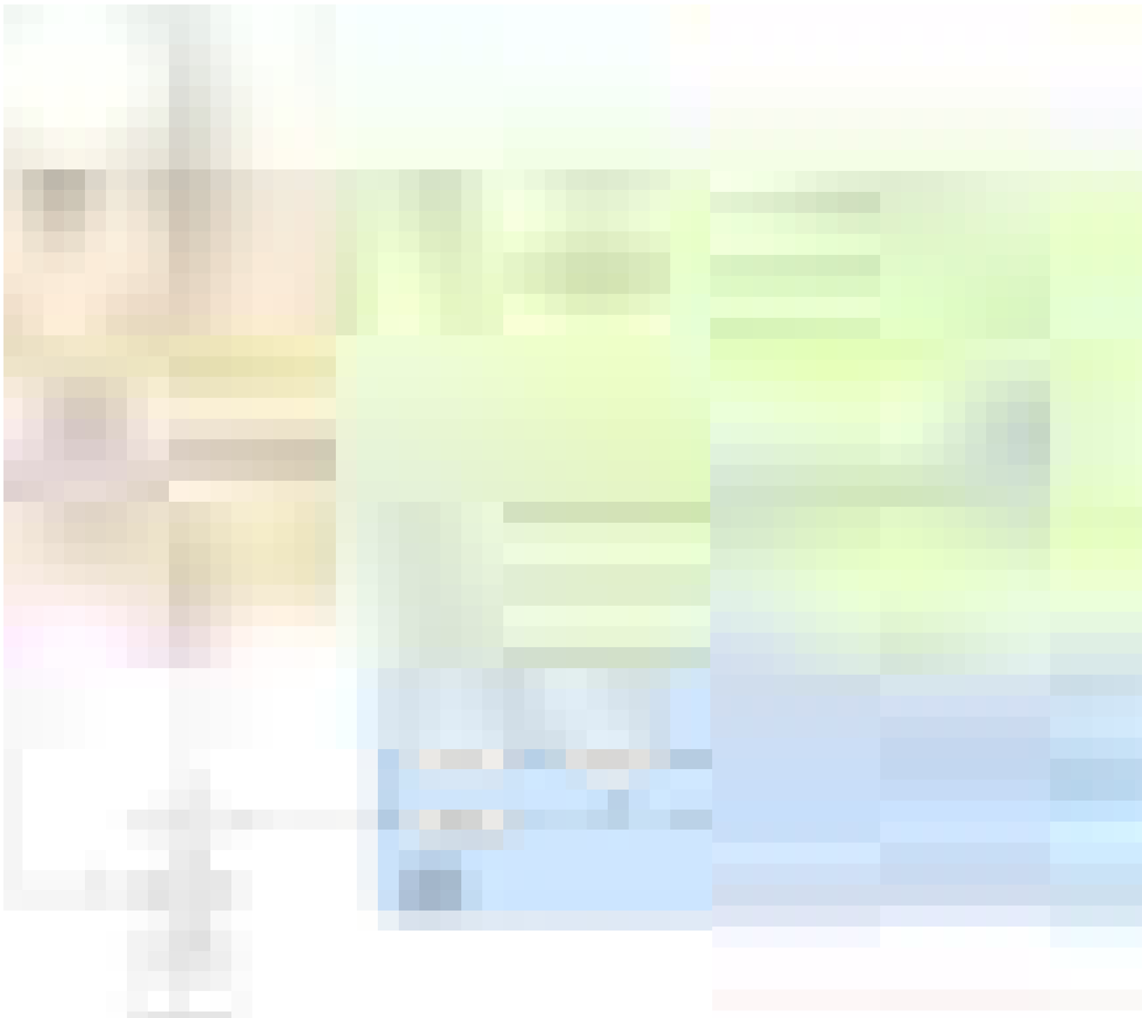


Figure 10: Flowchart of Simulation Runs in SHADO

The model starts with one replication of one shift schedule. Users of SHADO can adjust many initial settings of SHADO, related to the shift as well as other submodules like team and artificial intelligence features. The shift schedule may include roughly three phases: beginning transfer-of-duty, formal operation, and ending transfer-of-duty with the next shift's dispatchers. Within each phase, there is a staff of at least one dispatcher team with at least one dispatcher managing at least one fleet of at least one vehicle. As depicted in Figure 9 and logically described in Figure 10, each fleet generates tasks that arrive into the dispatch operations center system. Factors like fleet size or heterogeneity may affect this arrival rate. When a new task (Task A) arrives into the system, it is in an instantaneous global holding queue for the model to check if there is a dispatcher trained to handle that type of task. In some cases, when Task A arrives into the system, it may be a 'lead' task. This means that there are some 'follow-up' tasks that do not arrive independently but rather with interarrival times relative to Task A's real arrival time.

If there is more than one dispatcher, as pictured in Figure 11, then the model checks the length of each dispatch team's queue. The team with the shortest queue gets to take on this new task onto their local queue. If there are both humans and *equal operator artificially intelligent decision aids* present on the team, then the model will only direct the task to the *equal operator* if no human dispatcher is presently available to handle the task. For example, if a dispatcher is busy responding to a high priority task, then the *equal operator AIDA* would process the task with a processing time and probability of error that the user decides relative to the human operator's average speed and accuracy.

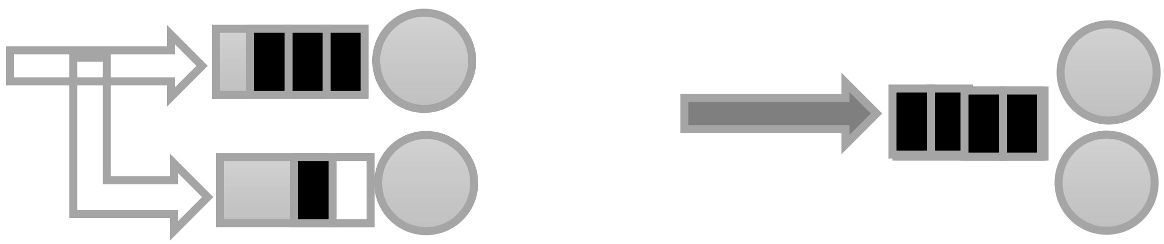


Figure 11: Schematic of queuing structures for team work

Operators working in two separate teams with different expertise would have their own expertise queue. However, operators and equal operator artificially intelligent decision aids would have a shared queue for tasks they can respond to as generalists. If there is a human dispatcher available when the task arrives in the queue then it is immediately processed in a time that is randomly drawn from the service time probability density function, which is included in the dispatcher's measure of utilization. If there are no *equal operator AIDA* present, but rather *task AI decision aid* designed to assist with this type of task, then the human dispatcher's true service time would be reduced by 30% or 70%, depending on whether it was *some* or *full* level of TAIDA. Additionally, the human error probability would also be reduced by 30% or 70% depending on the level of task assistance. The 30% and 70% markers are the thresholds the researchers assumed to be significant enough per lack of real-world data on such points.

If there were neither BAIDA nor TAIDA, but rather a team coordination AIDA (TCAIDA) and the team was composed of more than a single human dispatcher and there existed at least *some* level of team coordination, then the TCAIDA would similarly reduce the service time on team coordination tasks and human error probability on all team tasks both by 30%, or 70% with *full* TCAIDA. For example, if a team communicated using a system that automatically re-assigned tasks as needed or alerted each operator of issues any of the team is facing that could affect their own work, then using that could improve how quickly they share information and how accurately they perform tasks. The team can have a combination of all types of AIDA or none. With no AIDA, the task would wait in the team's queue until a dispatcher became available to handle it. If the task has an expiration time (e.g. a phone call that stops ringing after 30 seconds), then it may prematurely exit the system before it is completed if its total wait time and service time are greater than its expiration time.

Task A may expire if another task (Task B) enters in the queue and shifts its position to later in the queue. If the team is fully occupied with Task A when Task B arrives into their queue, then the model checks if Task B is an essential task. If it is essential and Task A is an interruptible task, then Task B would interrupt Task A. Task A would be placed back onto the queue with the portion of true service time remaining recorded. If a team operates with a user-selected attention allocation scheme other than First-In-First-Out (FIFO), the model checks for certain attributes of Task A and Task B. A user can select a Shortest Task First (STF) scheme. In this case, the model would check for any difference in length of true service times and reorder the tasks waiting in the queue accordingly with shorter tasks to be completed earlier. For a Priority scheme, the model would check for any difference in priority ranking of the tasks and reorder the tasks waiting in the queue with higher priority tasks to be completed earlier. If neither task was defined by the user as essential but Task A was interruptible, and Task B was a higher priority, then the same would occur. However, if Task A were essential too or simply uninterruptible, then Task B would have waited in the queue until a dispatcher on the team became available.

A similar logic is employed in this model for a team with the STF scheme. In this case, though, if the task is defined as interruptible, the true service time remaining is compared with the time required by each task waiting to pre-empt it as needed. The essential characteristic of tasks is a global attribute, regardless of the team's attention allocation scheme. Therefore, in cases of FIFO, STF, and Priority, any essential task that arrives must be handled first by interrupting current tasks or by waiting directly behind uninterruptible, albeit, nonessential tasks.

Some tasks may not have an expiration time characteristic. If such a task arrives into the system toward the end of a shift's phase and this is not the penultimate phase, then the model is designed to extend the phase time to include the wait time and service time of the final task that arrived within the phase's original threshold. If there is an ending transfer-of-duty period defined, then during the penultimate phase, the model checks if the task is an essential task. Only essential tasks can cause that phase to override the original shift schedule defined by the user (e.g. 8-hour shift) and move the final phase end time, leading to a longer total operator shift time. If the final task that arrived during the penultimate phase is nonessential and it is waiting in the queue at the end of that phase, then it is recorded as an expired task. If that task, though, is currently being serviced, then it is left unfinished as well.

Once the task exits the system, there is a record of how long it waited in the queue and if there was any error on completing the task. For each dispatcher in a shift, the simulation tracks several different statistics. Utilization, the principal measure, is used as a proxy for dispatcher workload. It is defined as the percentage of time a dispatcher spends on task performance out of the total operation time. SHADO records the utilization for each dispatcher in 1-hour intervals and presents the distribution across the number of replications computed. Utilization is an important statistic because decrements in human operator performance are more likely to occur when utilization is below 30% and above 70% (Cummings et al., 2016, 2013; Cummings & Nehme, 2010; Rouse, 1983; Yerkes & Dodson, 1908). The statistics SHADO records are summarized in Table 1.

There are four types of human errors recorded. *Missed* and *incomplete* tasks are artifacts of the simulation based on the timing of events. If tasks have an expiration time or arrive too late to be processed before the end of the shift, then the dispatcher may miss starting them or may be disrupted while processing tasks. The human error probabilities (HEPs) used internally in SHADO ultimately affect the number of *failed* tasks. Whether or not the failed tasks are caught depends on the user-specified error catching chance per team per task. The HEPs were derived from work published by the Rail Safety and Standards Board (Gibson, 2012). Gibson, supported by Network Rail, the Association of Train Operations Companies and London Underground in the United Kingdom, developed a technique for quantifying human error in the railway industry. Although the analysis focused on locomotive engineer tasks, the results have been applied to the context of dispatchers. This is a noted limitation and needs further future research.

The default dispatcher task types identified in Table 5 can be described by generic task types (GTTs) with associated triangular distributions of HEP as listed in Appendix A. GTTs range from skills-, rules-, to knowledge-based tasks (Rasmussen, 1983). According to Gibson (2012), at least 12 and up to 28, usually around 16, in 100 of these tasks are likely to fail and any additional tasks that a dispatcher may have to perform for operations at the very least have a 1 in 10,000 chance of failure. With dispatch decision support systems, the event likelihood could be reduced to 1 in 100,000.

Table 1: Simulation Output Statistics

Output Statistic	Description	Purpose
Utilization	Time on task divided by total (1-hour) time interval	Identify workload of remote operator
Missed Tasks	Tasks that were not started by the remote operator	Identify insufficiency in dispatch operations staffing
Incomplete Tasks	Tasks that were started but not completed by the remote operator	Identify ineffectiveness in remote operator performance
Caught Failed Tasks	Tasks that were completed incorrectly and repeated	Identify inefficiencies in operator performance
Uncaught Failed Tasks	Tasks that were completed incorrectly	Identify riskiness in dispatch operations design

3.3 Usefulness and Limitations of SHADO

An online platform was developed to host open access for stakeholders to use the Simulator of Humans & Automation in Dispatch Operations (SHADO). The landing page is pictured in Figure 12 and additional screenshots are available in Appendix B.

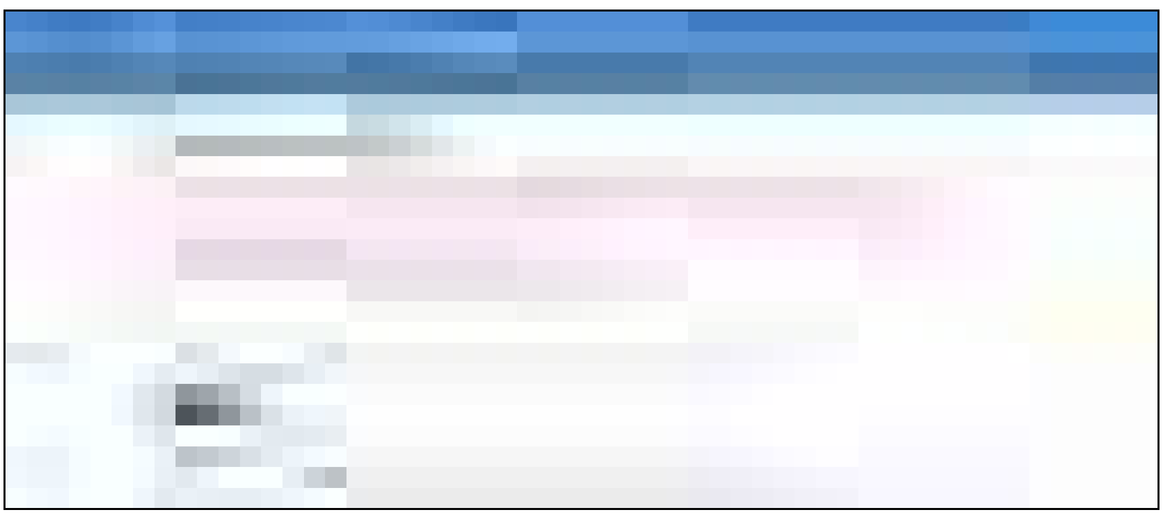


Figure 12: Landing Page of SHADO open online platform

SHADO output is based on information a user defines about the dispatcher teams, the fleets they manage, and tasks associated with their operations. Data on how frequently the tasks are arriving into the dispatchers' queues, how long it may take the dispatchers to process each task, and the dispatchers' chances at completing the tasks successfully form the arrival processes, service processes, and human error probabilities.

Once a user selects their domain (for example, railroad or airline operations), the interface adapts to custom content allows them to adjust input parameters of the shift they intend to simulate. Probability densities like the triangular or uniform distributions are described in layperson terms and tooltips are included to explain how the model works with certain choices a user makes (for example, whether a task is essential) throughout each module of settings. Users have the option to decide whether the teams have artificially intelligent decision aids, and they can force irregular scenarios to simulate how such irregular events may be handled under different system configurations. The final modules of settings allow the user to define which task(s) come from which source(s) and which dispatcher(s) would be responsible.

Although the platform makes it possible for users to make changes to most parameters in the model, there are several assumptions in the design of underlying SHADO model as highlighted in Table 2. These assumptions help simplify the complex dispatch operations center system but result in some limitations. For example, multitasking is modeled by interruption between tasks and not by parallel processing of tasks. All human operators are assumed to experience the same rate of change in cognitive performance due to fatigue as well as the same distribution of time on the same tasks despite any individual differences.

Table 2: Assumptions Designed into Internal Variables of SHADO

Assumption	***Meaning***
Serial Processing	A remote operator can service only one task at a time.
Fatigue	The homeostatic fatigue model from Hursh et al. (2004) was incorporated in that the service time for each task is multiplied by the appropriate fatigue factor depending on the time it arrives into the remote operators' queue for processing such that there is a 1% increase, per hour, in how long it takes each operator to complete tasks.
Common Task Efficiency	A task's service time is drawn from the same distribution regardless of the type of operator completing the task.
Three Types of Artificial Intelligence Decision Aids	1. Agents that assist individual operators reduce the human service time on associated tasks by 30% (partial assistance) or 70% (high assistance); 2. Agents that assist operator teams reduce team coordination time by 50%; 3. Agents that assist fully in operator task performance can handle any task completely at the same, better or worst speed and accuracy as operators on the same team.
Remote Operator Teams	A remote operator may be on a team with other operators and they may have no, partial or a high level of coordination. A partial level of coordination means that they have team communications tasks that arrive on an exponential distribution of once every 10 minutes, lasting an average of 10 seconds. While a high level of coordination means that the team communicates once every 5 minutes. Coordination also leads to greater chances of catching a human error at the desk—30% more likely with partial and 70% more likely with full.
Fleet Autonomy	A fleet with some level of vehicle-to-vehicle communications would demand less from a remote operator. A partial level of fleet autonomy has a decrease in rate of operator task arrivals by 30% while a fleet with high levels of autonomy would reduce the rate of task arrivals by 70%.
Type One Irregular Scenario	Remote operators working during a shift with an irregular scenario of Type One have an irregular and essential task (this concept is described in Section 3.2) that takes anywhere from 20 to 40 minutes, arriving, on average, once every 8 hours of their shift.
Type Two Irregular Scenario	Remote operators working during a shift with an irregular scenario of type two have all tasks that are affected by it arrive 10% more frequently into their operations system.

SHADO was designed to be flexible to model components and entire systems that do not exist today. Users can specify new remote operator roles different from typical dispatchers with new types of tasks, new interarrival and service time distributions, new team structures and fleet compositions. All the while, SHADO facilitates human-centered design of such complex dispatch operations center systems by simulating the impact of fleet-side and office-side interactions on each remote operator's workload. Examples of the outputs from SHADO are pictured below in Figure 13 and Figure 14.

Figure 13: Sample railroad dispatcher workload and error results

Figure 14: Sample airline dispatcher workload and error results

3.4 Chapter Summary

This chapter outlined the structure of SHADO as a conceptual and computational model. The attributes and assumptions of the discrete event simulation model were presented. Attributes were grouped in Figure 9 by those related to the fleet (size, heterogeneity, autonomy), environment, dispatch operations center (AIDA, team coordination, team size, shift schedule, team expertise), and individual dispatcher behavior (attention allocation strategy). Dispatchers, or remote operators, function as serial processors who handle task events arriving from the fleet, environment, and in some cases, their own team in the operations center. With SHADO, users can measure how busy each dispatcher is, how long the tasks wait to be addressed, and how many tasks not completed successfully. In the next chapter, details on how SHADO was validated for use in real cases of railroad and airline dispatch operations are introduced.

4. Model Validation

The Simulator of Humans & Automation in Dispatch Operations (SHADO) was developed to support managerial decisionmakers in planning for staffing and design of dispatch operations for future fleet management. For a simulation to be useful, it must be used. For a simulation to be used, it must be trusted, and hence validated.

The purpose of validation is to increase one's confidence in using a tool. The first goal in validation was to increase internal confidence. Since SHADO was custom-built for dispatch operations modeling, it was important to ensure that the code functions as conceived. Does the model take in input parameters and produce expected results? Does the model respond as expected when the internal parameters are adjusted?

Once internal confidence was gained in how SHADO works, the next goal was to build external confidence. Does the model get results close to what SMEs experience in the real-world? Does the model behave realistically when the initial settings are positively or negatively adjusted? In this chapter, the results of SHADO's multi-stage verification and validation process for rail and air dispatch operations are presented.

4.1 Validation Approach

It is important to validate discrete event simulation models to provide a trusted platform for future research and decision-making (Law & Kelton, 2000). Robinson (1997) describes methods to verify and validate simulation models. These methods are not for proving that a model is infallible but rather to systematically build one's confidence in using the model. There is no way to guarantee that a model perfectly represents the real world, so the goal of this validation process was to ensure that SHADO would sufficiently simulate the performance metric that mattered most in this study: operator workload. SHADO simulates the workload of remote operators managing fleets in various conditions. It takes into consideration the impact of any artificially intelligent systems that support the operators, the operators' staffing structures, and each operator's attention allocation strategy.

The process of validating SHADO included observations of real-world dispatch operations centers in multiple domains of interest, conceptual model validation, computational model verification, computational model validation for each operational domain of interest, and computational model synthesis for generalizability across different domains. Figure 15 visually represents the iterative multi-stage process.

These confidence-building techniques were applied at each stage:

1. **Real-world observation** entailed connecting with subject matter experts from each domain of interest to gather data on factors that affect the workload of remote operators along with relevant inputs and outputs of their operational systems. **Conceptual model validation** involved abstracting the real-world system into a general mathematical structure of adaptable components that interact to result in a metric for workload. Throughout this design stage, reviews

from SMEs were continually incorporated to ensure that the modeler's interpretation closely reflected the real-world system users' experience.

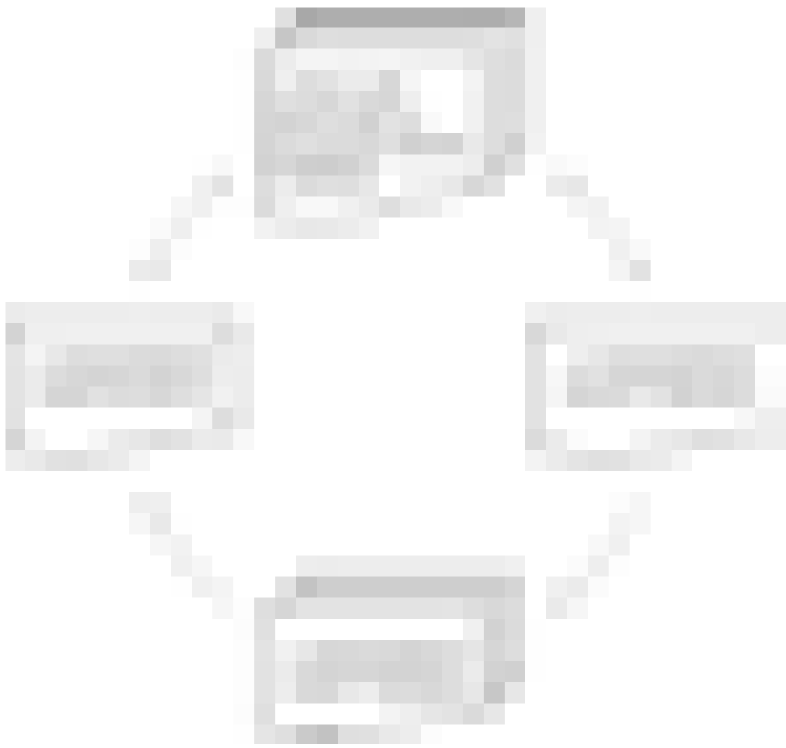


Figure 15: Process of designing and validating SHADO for modeling operator workload in real-world dispatch operations

2. **Computational model verification** occurred early in the process of programming software to represent the conceptual model in Figure 9. This computational model developed was verified by first ensuring that SHADO produced expected analytical outputs under defined inputs. Then, SHADO's results were compared with results from a commercially-available software package that has been considered the industry standard. A sensitivity test was conducted with all 10 key internal parameters (for example, fleet size or shift

schedule) each varied on at least three general levels (for example, 8-hour shift, 10-hour shift, and 12-hour shift).

3. **Computational model validation** is the ongoing process of building confidence in stakeholders, who may be external to the computational model development process, using SHADO for real-world operations research and development. This process was completed for the scope of this thesis with the following three steps:
   - *Data validation* to ensure that the underlying distributions representing the operations accurately represented each domain of interest.
   - *Open-box validation*, a method of testing the internal structures of software, to inspect SHADO part by part with SMEs for each domain of interest.
   - *Black-box validation*, to examine the functionality of software regardless of its internal workings, by using statistical goodness-of-fit tests for holistically comparison of SHADO's results with historical workload data in each domain of interest.
4. **Computational Model Synthesis** was the final step of building general confidence in SHADO. Through sensitivity analyses, researchers determined the input-output relationships within and between the domains of interest in SHADO. This process highlighted key internal parameters which consistently had significant impact on key performance indicators across all the domains of

interest. Results from this work support SHADO's generalizability for future applications.

4.2 Real-world Observation & Conceptual Model Validation

Railroad and airline operations centers served as the basis for validating SHADO with present-day systems. The goal was to build confidence in using SHADO to plan for evolutionary changes in both domains as well as revolutionary concepts of remote operations centers for future systems such as on-demand autonomous air taxi services (Nneji et al., 2018; Nneji, Stimpson, Cummings, & Goodrich, 2017). Therefore, beginning in October 2015, several visits were made to observe real-world and simulated railroad and airline dispatch operations.

4.2.1 Railroad Dispatch Operations

First, researchers traveled to Rio Tinto-Iron Ore of Canada's Quebec North Shore and Labrador Railway in Sept-Iles, Quebec and Amtrak's High-Speed Rail Training Facility in Wilmington, Delaware. These two visits were for building context of the nuances in railroad operations with single- and two-person locomotive crews in freight and passenger services within North America. As a proof of concept to support building SHADO to simulate workload in dispatch operations, a discrete event simulation of locomotive operator workload was developed, validated, and used to analyze the potential human-system performance impact of installing automation technologies on onboard crewmembers (Nneji et al., 2019).

In January 2017, a 3-day visit was made to a dispatch center that manages a Class I railroad network that spans up to 21,000 route miles of track across the US and Canada (Huang et al., 2018). This dispatch center managed both freight and passenger services and coordinated 550 to 750 trains at a time. There were approximately 300 dispatchers, and each controlled a geographical vehicle. During the visit, the researchers participated in seven hours of training on dispatcher work and technologies, observed four dispatchers over 12 hours across two days and two shift schedules, and consulted with the chief dispatcher to verify data gathered and fill in gaps in the researchers' understanding of the railroad dispatch operations. Further details from this visit are reported in Appendix C.

This Class I railroad company provided an exemplar how a railroad dispatch center could grow its network of humans and systems into a vast operation much like many commercial airlines, some of which are presented later in this section. However, for the purposes of gathering high-fidelity data on a representative group of dispatchers in an organization, a smaller railroad dispatch operations center was visited in January 2018.

Rio Grande Pacific Corporation is a holding company for regional freight railroads. RGPC is headquartered near Fort Worth, Texas where their dispatchers are at work dispatching for 12 short-line freight railroads and one local commuter railroad.

The commuter railroad, which operates by centralized traffic control (CTC), covers just 22 miles out of RGPC's over 2400 total miles under management.

RGPC is normally staffed with two dispatchers (pictured in Figure 16 below), one for the freight railroads and the other for the commuter railroad. Documentation (Figure 17) was gathered on the shift schedule of the seven dispatchers that work around the clock on their two desks. Each weekday, two dispatchers work the two desks during the morning (AM, 1st) and afternoon (PM, 2nd) shifts. One dispatcher works both desks during the overnight (ON, 3rd) shifts and for each shift during the weekends.

During the first visit to RGPC in January 2018, the researchers spent two days in Fort Worth and underwent a 3-hour training session on the first day with the chief dispatcher. This was followed by meetings with senior operations managers responsible for railroad technology integration. During the meetings, the managers described present and anticipated challenges with technologies in the dispatch center. This provided useful context for how the conceptual model, discussed in the next section, would be designed. Intermittently during the first day, the chief dispatcher directed the researchers' attention toward the dispatchers at work to note important times of transition in workload. On the second day, the researchers spent nearly 9 hours observing the work of dispatchers at the two freight and commuter desks. The researchers gathered data from direct observation and in-situ interviews. This first visit was concluded with presentations to the senior managers and dispatchers on the conceptual model design.

Following the January visit, the research team received digital copies of RGPC's dispatch operations records. The research team continued communications and returned in March 2018 and May 2018 to verify and validate SHADO's conceptual model design, input parameters and output statistics for present-day operations. And, as will be discussed in Chapter 5, the researchers followed up with the chief dispatcher for several months to analyze results of SHADO used to answer prospective questions for their operations. Over three years of real-world observations established a strong basis for the use of SHADO in the railroad domain and RGPC has provided an excellent testbed for in-depth validation of SHADO.

Figure 16: Rio Grande Pacific Company's two dispatcher desks (left: shortline freight, right: local commuter).

Figure 17: RGPC dispatcher shift schedules.

4.2.2 Airline Dispatch Operations

Like research conducted in the railroad domain, researchers also gathered data from several airline operations. In July 2017, October 2017, and April 2018, three remote operations centers of global airline companies were visited. During each visit, a consistent presentation of the conceptual model design, as presented in Section 4.2.3, was delivered to senior dispatchers and operations managers. The presentations began with two overarching questions: 1) how could and should we strategically distribution functions across teams of [remote operators] and artificially intelligent agents in the [airline] network operations? 2) What are staffing and design considerations to support the work of [remote operators] in the system? The plan for the visit was presented as including interviews of different personnel, observations of how the different personnel coordinate with each other during real operations, and review of databases that the organizations already capture timestamps on different tasks (such as phone calls or use of software) that automatically recorded in computer systems. Details from the visits are reported in Appendix C.

Visiting such global commercial airlines revealed complexities in operations that ultimately led to the selection the 24-7 dispatch operations center of a regional airline operating on a scale like RGPC—Horizon Air—for this dissertation. On any given day, Horizon may dispatch over 300 flights in and out of cities in Canada, and states in the northern and western region of the United States with a fleet of 60 aircraft including Embraer jets and Bombardier turboprops. During the June 2018 visit to Horizon Air, multiple managers and dispatchers were interviewed and observed at work.

There are currently six dispatch desks and five shifts which result in eight staffing configurations at Horizon Air Systems Operations Center. This means that there are hours each day with only one, three, four, or five dispatchers simultaneously on duty. One 8-hour shift is staffed by the chief dispatcher while four 10-hour shift schedules are staffed by dispatchers in direct contact with pilots of the flights they manage. As shown in Figure 18, there are eight rotations in a 24-hour period. After midnight and until 4:05am, there is one overnight (ON) dispatcher on duty at Desk 6. At 4:05am, three dispatchers come on duty for the morning (AM) shift at Desks 1, 2, and 3. They stay on duty past 5:30am when the ON dispatcher leaves. At 7:00am, Desk 4 is filled by a midday (MID) dispatcher and at 9:00am, Desk 5 by a chief dispatcher, both until 5:00pm. At 2:05pm, Desks 1-3 are relieved by three afternoon (PM) dispatchers who work until 12:05am and are joined at 7:30pm by the ON dispatcher.

Figure 18: Horizon dispatchers' desks on duty throughout a day.

Dispatchers working during the morning and afternoon shifts on desks 1, 2, and 3 self-reported their experiences via a novel form that the researchers designed during the visit: Dispatcher's Rough Assessment of Workload-Over Usual Times (DRAW-OUT). DRAW-OUT is a tool designed to allow for the SMEs to recall qualitative aspects of their experienced workload and report it over the course of their shift with the quantitative metric of utilization. Appendix E includes an example of the DRAW-OUT tool used at Horizon.

Researchers gathered the observed task service time and utilization data for every hour of the dispatchers' 10-hour shifts. The senior dispatcher also provided additional information on flights scheduled for multiple days on the three desks and shifts which was used to generate task arrival time distributions. This visit and subsequent calls with a senior dispatcher at Horizon Air provided ample real-world observational data to validate the conceptual model design, input parameters and output results for SHADO.

4.2.3 Conceptual Model Validation

Conceptual model validation is the process of checking how well the real-world variables of a system are captured in the design of the model. As Robinson (1997) and others (Rand & Wilensky, 2006; Sargent, 2005; Yow, Walters, Plott, Laughery, & Persensky, 2005) highlight, an early step of confidence-building typically involves conceptual model validation. For both railroad and airline operations, the conceptual model design was validated before, while, and after real-world observational data on each of the ten parameters were collected.

Ten structured sessions were organized during a 2-year period with over 300 operational leaders from major railroad, airline, and other transportation and logistics organizations across the United States. Additionally, interviews were conducted with over 20 stakeholders representing original equipment manufacturers, market actors, governments and alliances conceptualizing operations for future on-demand air taxi services (Nneji et al., 2017). During each session, the ten internal variables of SHADO were presented and discussed. This gave researchers the opportunity to gather feedback and identify any gaps in understanding and interpretation of the conceptual model of remote operations centers and dispatcher workload. Further details on the conceptual model validation process, including questions that the researchers led the sessions with, are presented in Appendix C.

4.3 Computational Model Verification

The next step in the validation process was to verify that SHADO accurately met expected system specifications by computationally representing the conceptual model that was designed. First, it was important to ensure that SHADO produced target analytical results under defined input parameters. Then, outputs of the key performance indicator, utilization, were compared in sensitivity analyses of seven of the key internal parameters. Simultaneously, parallel tests were conducted during both of these stages using an alternative commercially available software package. This provided an additional layer of verification that SHADO's performance met, and in some cases, exceeded the industry standard.

Two key input parameters of any discrete event simulation are event service time and interarrival time distributions. In the context of SHADO, events are tasks that arrive from the fleet management system and require an operator's attention to respond and provide service to complete the tasks. For this first stage of verification, three generic tasks were defined, each with different service and interarrival time distributions. SHADO was run for 500 replications with settings of ideated operations detailed in Appendix F. As shown in Table 3, SHADO yielded results within thresholds +/-4.5% of expected minutes of service time and +/-8% of expected number of arrival events across the tasks shown. The average total difference between the analytical values calculated mathematically and the simulated values computed using SHADO was -1.26%.

The next stage of verification involved performing a sensitivity analysis of the resulting operator utilization per internal variable in SHADO. With all other variables held constant, each of the 10 internal variables was adjusted. For example, the fleet size factor was tested on three levels with 1 vehicle, 2 vehicles, and 3 vehicles simulated to study the response of operator utilization over 500 replications. As expected, operator utilization was found to increase with fleet size.

Table 4 below summarizes all the results that are detailed in Appendix F.

Results from the sensitivity analyses supported the fact that SHADO was programmed and debugged effectually. SHADO showed the same relationships in internal parameter and utilization output changes one would expect in the real world. However, SHADO is a stand-alone simulation software package that researchers developed using Java. Therefore, Rockwell Automation Technologies' commercial off-the-shelf Arena (Version 15.00.00004) computer program was used to independently verify the architecture of SHADO and build industry confidence in this new tool. A screenshot of the Arena model built is in Appendix F.

SHADO's performance was compared with Arena's across the three input-output variables: (1) task service time, (2) task arrival events, and (3) utilization for the ten internal parameters in sensitivity analyses. These results are also included in Table 3. Comparisons with Arena further verified that the input-output architecture and internal structure of SHADO was computationally accurate. Additionally, SHADO was found to have advanced features that showed it to be more flexible than Arena which had limitations that did not allow for fully testing it simultaneously with SHADO. Therefore, the next steps of computational model validation using data from real railway and airline domains was critical. While verifying the ability of SHADO to behave according to the researcher's conceptual model built internal confidence, the results of SHADO still needed to be externally validated in the context of railway and airline dispatcher workload as described in section 4.5.

Table 3: Comparison of Mean Input Parameter Results from 500 Replications of SHADO versus Arena Simulations

input parameter	Task	Mean Value		
		Target	SHADO	Arena
Service Time (minutes)	**Task 1**	2.5	2.59 (+3.6%)	2.5
	Task 2	1	1.02 (+2%)	1
	Task 3	2	1.91 (-4.5%)	1.67 (-16.5%)
Arrival Events (count)	**Task 1**	42.1	41.98 (-.3%)	40.74 (-3.2%)
	Task 2	17.78	17.71 (-.4%)	17.49 (-1.6%)
	Task 3	2	1.84 (-8%)	1.82 (-9%)

Table 4: Internal Variable-Operator Utilization Relationships

Internal Variable	Relationship with Operator Utilization
Fleet Size	Higher number of vehicles in fleet leads to higher average workload.
Fleet Heterogeneity	Less homogeneity of vehicles in fleet leads to higher average workload.
Fleet Autonomy	Higher levels of fleet autonomy lead to lower average workload.
Exogenous Events	Any of the defined irregularities in operations would increase average workload. There is greater variability in workload in scenarios when a new lengthy task arrives unexpectedly versus in scenarios with poor weather effect of increasing arrival rate of all tasks.
Shift Schedule	Overall, average workload is not affected by shift schedule. However, longer hours reduce the variability in workload.
Team Size	More dispatchers working together reduces average workload and lowers the maximum workload.
Team Expertise	Generalist teams lead to lower average workload whereas some operators in a specialist team may experience considerably higher average workload.
Operator Strategy	Overall, average workload is not affected by strategy, but this ultimately depends on user-defined task parameters and how the tasks are regulated in the operator queue.
Team Coordination	Workload is increased from inter-dispatcher communication tasks, but workload is reduced in future tasks that arrive into the operations center.

Artificial Intelligence Decision Aids	Equal operator reduces workload in the way a second human operator would on a team working together. Task assistant would reduce workload from servicing tasks. Team assistance would reduce workload from time spent on inter-operator communications.

4.4 Computational Model Validation

4.4.1 Data Validation

Data validation is the process of checking sets of information gathered from different sources in the field to be accurate before using them as inputs in the simulation. The goal is to ensure consistency and completeness so the potential for error or loss of data is mitigated. The process of validating data gathered for model simulation with SHADO occurred in multiple stages. The two primary inputs in discrete event simulations like SHADO are task interarrival and service times. Therefore, researchers observed and interviewed SMEs—dispatchers—at Rio Grande Pacific Company (RGPC) and Horizon Air to gather task- and time-related data. With a large dataset of times collected for each task across for different operating conditions, researchers were able to generate interarrival and service time probability density functions. Data validation was an iterative process of presenting the generated distributions for each task back to the dispatchers and better accounting for operational differences as needed to get estimates closer to real-world task times. Before any time-related data was collected, researchers first conducted a task analysis to identify the set of cognitive actions and processes on which dispatchers spent their work time.

4.4.1.1 Railroad Dispatch Operations

A railroad dispatcher's high-level function is to remotely direct and coordinate the safe movement of railroad traffic for his or her territory under management. From that level, there are specific requirements that are consistent across operations in the US. Title 49 Code of Federal Regulations Part 228 Section 17 defines what records dispatchers are required to keep. They total up to 11 types of information related to train movement, including identification, weather, and travel times. Then, there is the General Code of Operator Rules (GCOR). GCOR is over 150 pages of rules and instructions that was formed by committee and with which over 200 railroads agreed to comply. These include the 12 short-line freight railroads and one local commuter railroad that RGPC manages.

Finally, RGPC created an internal handbook, the "Train Dispatcher's Rules/Instructions Manual," for the "safe and efficient movement of trains." This combination of federal regulations, industry standards, and company expectations all lead to dispatchers meeting the two high-level goals of safety and efficiency of the railroad network. The researchers identified ten lower-level tasks for dispatchers working the freight desk and only nine lower-level tasks on the commuter desk. Through subject matter expert interviews and observations with multiple dispatchers at RGPC, the task lists were validated. The task lists for both desks are presented in Table 5 and Table 6 and detailed with the cognitive task analyses reported in Appendix C. These lower-level dispatcher tasks serve to meet the higher-level railroad operational goals.

Table 5: Default Task Timing Input Parameters for Freight Desk per Shift

Dispatcher Task Types	Interarrival Time (minutes)			Service Time (minutes)		
	AM	PM	ON	AM	PM	ON
Actuation (OK)	Expo(16.6)	Logn (43.1,57.6)	Tria(3,5,168)	Unif(2.8,4)		
Actuation (Clear)	Logn (25.9,49.1)	Logn(23.8,43.9)	Tria(2,16,234)	Unif(1.1,2.3)		
Daily Operating Bulletin		Expo(360)			Expo(15)	
Temporary Bulletin Issue	Expo(240)		Expo(1440)	Unif(1.3,4)		
Temporary Bulletin Void & Verify			300			Expo(30)
Other Communications	Expo(17.4)		Expo(60)	Expo(3.4)		
Weather Recording	300		200	Expo(2.5)		
Notetaking	Expo(60)			Expo(1)		
Reporting	Expo(480)			Expo(10)		
Miscellaneous	Expo(60)			Expo(1)		
Transfer-of-Duty	First and Last Task of Shifts			Unif(5,15)		

Table 6: Default Task Timing Input Parameters for Commuter Desk per Shift

Dispatcher Task Types	Interarrival Time (minutes)			Service Time (minutes)		
	AM	PM	ON	AM	PM	ON
Train Movement	Expo(8.1)	Expo(6.8)	Expo(26.7)[1]	Expo(1.7)		
Bulletins	Expo(120)	Expo(240)		Expo(5)		
Temporary Bulletins	Tria(240,480,1440)			Expo(5)		
Bulletin Printing		Expo(450)		Expo(15)		
Other Communications	Expo(60)			Expo(2.8)		
Weather Recording	300		200	Expo(2.5)		
Notetaking	Expo(60)			Expo(1)		
Reporting	Expo(480)			Expo(10)		
Miscellaneous	Expo(60)			Expo(5)		
Transfer-of-Duty	First and Last Task of Shifts			Expo(5)		

Within each day, there are three shifts. The nature of railroad operations is such that crews from the track call in to request track warrants in the morning hours and call back in the evening to clear track warrants with the dispatchers. Because of this difference in busyness and types of incoming calls, the task interarrival time data was separated by the three shifts.

With these three sets of data, Rockwell Automation Technologies, Inc.'s Input Analyzer (Version 15.00.00) software was used to determine the quality of fit of probability density functions to each dataset. This process was verified and repeated with a custom tool that the research team developed in MATLAB (Version R2018a) to

[1] The commuter railroad has low traffic in the first six hours and high traffic in last two hours of the overnight shift.

automatically filter results to the best fit distribution from a group of distributions that used parameters practical for the real-world context of this research. The group was limited to the following common functions: exponential (average interarrival time), uniform (minimum interarrival time, maximum interarrival time), triangular (minimum interarrival time, mode interarrival time, maximum interarrival time), lognormal (average interarrival time, standard interarrival time deviation), and constant interarrival time.

This group of distributions was also used to fit data on task service time. Task service time is an estimate of how long it would take a dispatcher to complete a task. To gather datasets of service time on, for example, *Actuation OK*, *Actuation Clear*, and *Other Communications* tasks, the researchers downloaded records from Media Exchange Interface for End Users (MXIE). MXIE was the desktop client application that RGPC used to keep chronological logs of all the incoming and outgoing calls on the freight desk throughout the morning, afternoon, and overnight shifts. The results of interarrival time and service time distributions for all the dispatcher tasks from both freight and commuter desks are presented in Table 5 and Table 6. As will be discussed in Section 4.4.2, the tables were reviewed with the chief dispatcher and six other dispatchers at RGPC through a process called “open-box validation” to validate that the probability density functions and input parameter estimates, as well as internal parameters and key performance indicator outputs, matched their professional experiences.

4.4.1.2 Airline Dispatch Operations

Like work conducted at RGPC, during the visit to Horizon Air, the research team observed and interviewed multiple managers and dispatchers. Figure 19 shows the desk of one of the dispatchers that the researchers observed at work. Like railroad dispatchers, airline dispatchers have multiple displays to manage their flights. At least one display is maintained for *flight planning* and another for *flight following*, their two primary tasks toward their high-level goal of maintaining safe and efficient flight operations. Desk 5, the chief dispatcher's desk, does not normally have flights allocated for planning. Yet, they maintain overall situational awareness to flexibly support dispatchers during peak periods or emergencies.

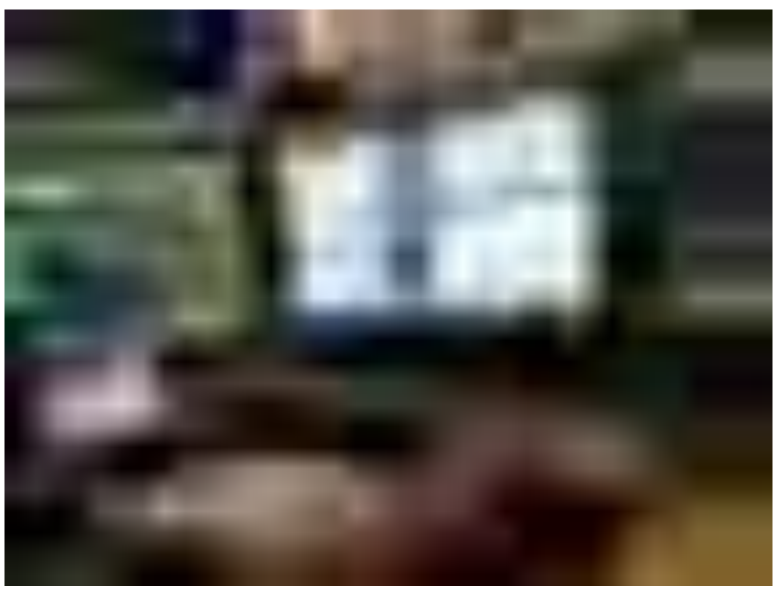


Figure 19: A Horizon dispatcher's desk with multiple displays for flight planning and flight following.

There are three types of flights: short-haul, long haul, and focus. Focus flights require special attention of dispatchers to consider uncommon requirements for aspects like fuel or payload. Short-haul flights have a duration of up to 3 hours while long-haul flights last over 3 hours. There are generally more short haul flights than both long-haul and focus flights.

From flight schedule reports recorded by Horizon Air on five representative days of airline operations in May and June 2018, interarrival time distributions of flight planning and flight following tasks for each desk and shift was generated. Table 7 presents the interarrival time distributions. To generate service time distributions, the researchers designed a quantitative questionnaire to solicit estimates from four dispatchers on how much of their time was spent on flight planning, flight following, and emergency management tasks under different operating conditions. These included representative days when pilots could operate under visual flight rules (VFR), which indicates good weather, and days with poor weather. The completed forms are presented in Appendix E. These task service time distributions are listed in

Table 8.

Table 7: Interarrival Time Parameters for Horizon Air Dispatcher Tasks

Dispatcher Position	*Flight Type*	*Flight Planning Tasks*	*Flight Following Tasks*
AM Desks 1, 2, 3	Short Haul	Expo(10.1)	Expo(8.03)
	Long Haul	Expo(1132.08)	Expo(1125)
	Focus	Expo(451)	Expo(257.14)
PM Desks 1, 2, 3	Short Haul	Expo(16.1)	Expo(9.15)
	Long Haul	No Arrival	Expo(1125)
	Focus	Expo(225)	Expo(180)
Desk 4	Short Haul	Expo(11.2)	Expo(9.64)
	Long Haul	Expo(1200)	Expo(1200)
	Focus	Expo(200)	Expo(160)
Desk 6	Short Haul	Expo(5.1)	Expo(22.05)
	Long Haul	No Arrival	No Arrival
	Focus	Expo(600)	Expo(408.16)

Table 8: Service Time Parameters for Horizon Air Dispatcher Tasks, including Miscellaneous Task Parameters

Flight Type	*Flight Planning Tasks*	*Flight Following Tasks*
Short Haul	Logn(3.3, 1.2)	Unif(1, 3)
Long Haul	Logn(11.5, 20.7)	Expo(10.3)
Focus	Expo(11.5, 21)	Expo(21.7)
Dispatcher Task Type	***Interarrival Time (minutes)***	***Service Time (minutes)***
Miscellaneous	Expo(60)	Expo(5)

As observed, long-haul and focus *flight planning* tasks arrive less frequently than short-haul flights which are planned on average once every 10.1 minutes during the morning shift. *Flight planning* tasks for short-haul flights were found to take dispatchers, on average, less time to complete than for long-haul and focus flights. The researchers discussed the fit distributions with the senior dispatcher to validate that the fit distributions looked reasonable. This data validation process established default input parameters and controlled settings such that researchers could examine the internal workings and output results of SHADO with confidence, as will be presented in the next section.

4.4.2 Open-Box Validation

With SHADO's underlying structure verified and the input parameters for both rail and airline operations validated, the next step was to perform open-box validation. Open-box validation is a method that stems from white-box validation. As Robinson (1997) describes, this method allows testers to investigate whether each component of the computational model sufficiently represents each corresponding real-world element. The difference here is that the testers are the potential users. To elicit feedback from the SMEs in railroad and airline operations, an open platform was developed to be used for demonstrating the Simulator of Humans & Automation in Dispatch Operations on the web (see Figure 12). The Java program of SHADO was run in the backend from the server while the frontend was designed with an accessible graphical user interface. Middleware was also developed to customize the context and default parameters depending on the user's domain of interest.

The researchers scheduled five synchronous online walkthroughs of SHADO with the chief and senior dispatchers at RGPC and Horizon Air. Once the chief and senior dispatchers selected their domain, their web portal would customize an automatic update using the data researchers validated as initial settings in each of their dispatch operations. The dispatchers were asked to each ensure that the content of SHADO was true to their real-world experiences. One submodule-at-a-time, the experts tested cases with components of the model. The results from the first four walkthroughs are shown in Table 2. The final walkthroughs included revisions to the custom content. Following this validation process, these stakeholders accredited the tool to support managerial decision-making and future planning for both railroad and airline operations, as will be discussed in Chapter 5.

Table 9: Subject Matter Expert Face Validation of Settings in Each Submodule of SHADO

Submodule	*Setting #1*	*Face Validity #1*	*Setting #2*	*Face Validity #2*
Shift	Hours	The increments of time (hourly blocks), minimum (1-hour shift), and maximum (12-hour shift), practically represented the dispatchers' realm of possible schedules	Transitions	The periods (beginning and ending) and timing of transitions represent dispatchers' opportunities
Task	Frequency	The minimum and maximum and range for parameters was reasonable for operations	Duration	The unit of time and range of possibilities was realistic and able to simulate different efficiencies
Fleet	Tasks	The options of tasks defined was a good representation and useful to define which fleets shared the same while also representing "other sources" of workload	Traffic	The hourly options were useful to represent trends in operations
Operator	Strategies	The first-in-first-out represented novice response, shortest-task-first represented expert response, and prioritization represented some rule-of-thumbs in companies, useful to move things around	Artificially Intelligent Decision Aids	The three types of agents fairly represented potential technology capabilities. The range in speed and accuracy comparison to humans was logical

4.4.3 Black-Box Validation

The constituent parts of the computational model were found to sufficiently represent the associated real-world elements. These parts included the submodules of shift settings, task settings, fleet settings, and operator settings. The next step was to determine whether the overall model represent the real-world dispatch operations in both railroad and airline with great enough accuracy. To complete this validation process, SHADO was used to replicate the observed empirical results for datasets from the railroad operations and airline operations. The number of replications for each domain was determined using rolling average inspection (Robinson, 2004) and confidence interval method (Law, 2007).

Historical workload data was not accessible in the same form for the desks at RGPC and Horizon. At RGPC, the freight desk could be validated using the company's computer-generated reports of dispatcher utilization. The commuter desk had no such real-world data with which to perform a comparison. Therefore, utilization on the commuter desk was validated by comparing SHADO's results with the expectations and intuition of the chief dispatcher. At Horizon Air, dispatcher-generated reports of their own experiences with utilization using a novel tool (see Appendix E) on three desks there were used to perform black-box validation procedures. Finally, the researchers reviewed the model reports with the SMEs at both RGPC and Horizon Air to perform what's known as a Turing Test of model validity (Schruben, 1980; Turing, 1950).

4.4.3.1 Railroad Dispatch Operations

During the first meeting with the chief dispatcher at RGPC, the researchers identified "talk-and-listen" (T&L) time as the chief's most critical indicator of dispatcher workload. T&L time is the duration of each dispatcher spends on phone calls. On the freight desk, this data has been automatically recorded with the company's computer systems for years. The times are totaled for each hour in the form of a percentage out of total time just as we think of utilization. The chief uses daily, monthly, and annual T&LT reports to quickly decide from his high-level perspective when dispatchers can handle more work or require additional assistance.

The researchers gathered talk-and-listen time-related data from the railroad operations morning (AM), afternoon (PM), and overnight (ON) shifts on the freight dispatcher desk over the course of the same days input data was validated for in March and May 2018. No utilization data could be recorded from January 2018. This spread of data still represents "slow," "average," and "busy" days per the SME's distinction as discussed in Section 4.4.

SHADO was run over 300 replications under the same operational input conditions, including the task interarrival time and service time distributions, for each shift as defined in Table 5. Table 10 summarizes the settings which were limited to validated input data of only the following four T&LT-related tasks: *Actuation OK, Actuation Clear, Temporary Bulletin Issue,* and *Other Communications*.

Utilization results from SHADO were compared to the real-world dispatcher utilization results generated from the company over 6 days x 3 shifts of operations. Tables of the raw T&LT data can be found in Appendix G. Six days of three shifts each were also randomly sampled from SHADO. The two data sets of utilization were recorded in hourly intervals for each of the 18 total shifts. Then a two-sample Kolmogorov-Smirnov (K-S) test was used to ensure the sample of utilization data from SHADO and the sample of utilization data from the real-world come from the same distribution.

Table 10: SHADO Settings for Kolmogorov-Smirnov Test of Dispatcher Utilization per Morning, Afternoon, and Overnight Shift

Settings	**Input Parameter**
Hours	8
Transfer-of-Duty	Beginning of shift Ending of shift
Tasks	Actuation (OK), Actuation (Clear), Temporary Bulletin (Issue), Other Communications
Dispatcher Strategy	First-In First-Out
Dispatcher Error Catching Chance	50% for all tasks
Number of Days	6

MathWorks' MATLAB (version R2018a) *kstest2* function, which returns a test statistic, the asymptotic p-value, and the test decision for the null hypothesis that the data in the two samples are from the same continuous distribution. The K-S Statistic D is defined as the maximum value of the absolute difference between two cumulative distribution functions where each cumulative distribution function is obtained from a list of data points of each sample on which the K-S test is applied. Tables of the resulting

SHADO data can be found in Appendix G. All the *D* values, reported in Table 11, were found to be less than the critical D-value.

The p-value denotes the level of significance with which the null hypothesis may be accepted. Large values of *p*, as we see in Table 11, imply that the cumulative distribution function of the two samples tested are not significantly different. The confidence that both populations do not belong to the same parent distribution is given by (1-*p*) x 100. We find from the K-S test that the distributions of T&LT-related utilization in simulated and real-world operations of RGPC are as if they belong to the same parent distribution, as the test rejects the null hypothesis at $\alpha = 0.05$ level of significance.

Table 11: Results of two sample K-S test for freight dispatcher T&LT-related utilization

Shift	**D**	**p**
AM	0.1250	0.8220
PM	0.1458	0.6521
ON	0.2083	0.2199

Finding that the utilization output from SHADO is not statistically different from the real-world utilization measured from RGPC's freight dispatch operations over a representative sample of days supports the null hypothesis. This increased confidence in modeling with SHADO. Yet, further work was required to validate SHADO for modeling the commuter desk as well.

Because there did not exist similar source of real-world utilization data for the commuter desk like the company's computer-reported measures, the researchers presented a set of results to the chief dispatcher to check how close SHADO simulated his experience and expectations of the average workload and extent of deviation in dispatcher workload between the commuter and freight desks across each shift. The researchers used SHADO to simulate 300 replications of the default parameters in Table 5 and Table 6 using input data that were validated for both desks in Section 4.4. The results are visualized in Figure 20 and detailed in Table 12.

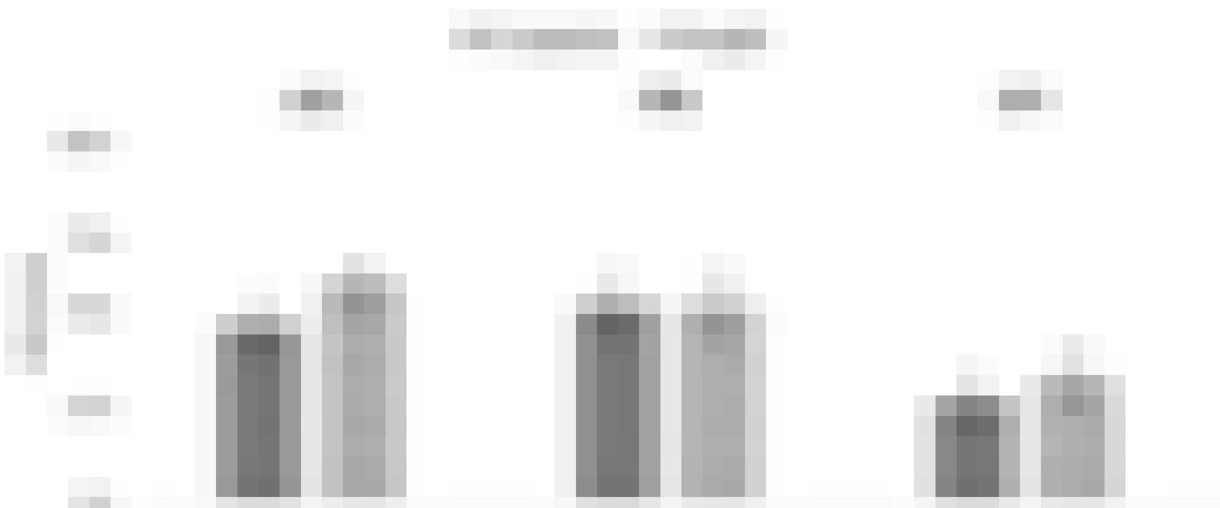


Figure 20: Average (and Standard Deviation) Dispatcher Workload Simulated with SHADO on Commuter and Freight Desk Data per Shift Schedule

Table 12: Average and Standard Deviation, including minimum and maximum averages on Commuter (C) and Freight (F) desks in morning, afternoon, overnight shifts

Shift	AM		PM		ON	
Desk	C	F	C	F	C	F
Average	46.8%	58.5%	52.0%	51.8%	27.2%	32.9%
S.D.	7.0%	8.1%	8.2%	8.7%	6.6%	7.8%
Min Average	29.4%	33.7%	31.0%	25.2%	12.2%	17.3%
Max Average	66.2%	82.1%	74.3%	80.6%	49.1%	60.6%

The research team asked the chief dispatcher how reasonable the results were under the operational conditions and he agreed on multiple findings from SHADO:

1. During an average morning (AM) shift, the freight dispatcher desk has more workload than the commuter dispatcher desk.
2. During an average day, the overnight (ON) shift dispatcher has lower workload from each desk.
3. During an average day, the dispatchers on the freight desk have greater variability in hourly workload than the dispatchers on the commuter desk.
4. During the afternoon (PM) shift, the commuter and freight desk workload averages are comparable although the dispatcher on the freight desk does tend to have spikes with extreme workload.

Finally, the researchers conducted a Turing Test with the chief dispatcher. The chief dispatcher was asked to review two identically structured reports of talk-and-listen time monthly totals for the year 2017. One report was generated by RGPC's own computer logging system. The other report was generated by SHADO following 12 simulation runs of 30 replications each of all three shifts. The settings used, and outputs generated are included in Appendix H. In looking at the two reports, the expert on RGPC railroad operations could not distinguish which results were from the real-world and which results were from simulation. This response was the final feedback that supported SHADO's usefulness in supporting real-world decisionmakers by closely simulating what one could expect from their operations.

4.4.3.2 Airline Dispatch Operations

Unlike RGPC, Horizon Air did not have any formerly computer-generated data logs available to estimate dispatcher utilization. So, to gather utilization data for black-box validation of SHADO's workload simulation in this context of airline operations, researchers designed a novel tool during their visit. Dispatchers working during the morning and afternoon shifts on desks 1, 2, and 3 were asked to self-report their experiences via Dispatcher's Rough Assessment of Workload-Over Usual Times (DRAW-OUT). DRAW-OUT is a tool designed to allow for the subject matter experts to recall qualitative aspects of their experienced workload during an "average" or usual day at work and report it over the course of their shift with the quantitative metric of utilization. Appendix E includes results from the DRAW-OUT tool used at Horizon. The researchers used DRAW-OUT results to systematically gather the utilization data for every hour of the dispatchers' 10-hour shifts. The real-world results were reported for the morning shift by four dispatchers and for the afternoon shift by three dispatchers, for a total of seven datasets.

SHADO was run for 365 replications using the input parameters (see Table 7 and Table 8) and internal parameters (see Table 13) validated for Horizon Air operations. Researchers compared results from SHADO with a representative sample of dispatcher-reported utilization by performing a two-sample K-S test. The goal was to investigate whether the distributions of hourly dispatcher utilization datapoints from the seven datasets was significantly different between simulated and real-world operations at Horizon Air.

Table 13: SHADO Settings for Kolmogorov-Smirnov Test of Air Operations per Morning and Afternoon Shift

Settings	**Input Parameter**
Hours	10
Transfer-of-Duty	Beginning of shift (PM only) Ending of shift (AM only)
Tasks	Short-Haul Flight Planning, Long-Haul Flight Planning, Focus Flight Planning, Short-Haul Flight Following, Long-Haul Flight Following, Focus Flight Following, Miscellaneous
Dispatcher Strategy	First-In First-Out
Dispatcher Error Catching Chance	50% for all tasks
Number of Days	7

The K-S test failed to reject the null hypothesis at the 5% significance level that utilization reported from SHADO versus reported from the DRAW-OUT forms grouped by shift schedule are from the same underlying continuous population. Table 14 shows the results that, for all three desks in the morning and afternoon shift, there was no statistically significant difference in utilization reported by dispatchers versus utilization simulated by SHADO.

Table 14: Results of two sample K-S test for Desks 1-3 dispatcher utilization

Shift	D	p
AM	0.2500	0.1393
PM	0.2333	0.3420

Boundary conditions were also tested to investigate how well SHADO simulated workload in extreme settings. The senior dispatcher who allocates flights at Horizon stated that their threshold for high workload is at over eight flight planning tasks in any hour. They set low workload at less than four flight planning tasks an hour.

SHADO was run with settings presented in Appendix I to simulate these two boundary conditions of input parameters for flight planning tasks during the morning and afternoon shifts. The workload results are shown in Figure 21 and Figure 22.

Figure 21: Average (and S.D.) Dispatcher Utilization over One Year of Morning Shift on Desks 1-3

Figure 22: Average (and S.D.) Dispatcher Utilization over One Year of Afternoon Shift on Desks 1-3

The utilization results of simulations run with settings for six flight planning tasks, and four flight planning tasks, respectively for the morning and afternoon shift are highlighted as green bars in Figure 21 and Figure 22. As presented in Table 15, for the morning shift, the senior dispatcher's goal has been to maintain on average six flight planning tasks per hour and in the afternoon shift, four flight planning tasks per hour, resulting in approximately 60 and 40 flight plans due per desk within each shift respectively. The afternoon shift requires dispatchers to plan and follow three times more *focus* flights which take seven to 20 times longer than *short-haul* flights as shown in Figure 23. So, although, they have only two-thirds number of flights to plan, the additional time is spent on *focus* flights leading to similarly moderate workload.

Earlier work (Cummings & Guerlain, 2007) suggested that operators in similar supervisory control settings experience moderate workload between 30-70% utilization.

Results from SHADO reveal that this finding corresponds with Horizon Air's heuristic for how flights are allocated across desks and shifts. The senior dispatcher also found SHADO's results to follow his expectations. These results increased external confidence for researchers and airline operations center stakeholders using SHADO as a tool to model such operations.

Table 15: Average Number of Tasks during Horizon shifts

Average	Morning	Afternoon
Number of Flight Planning Tasks per Hour	6	4
Number of Focus Flight Following Tasks per Shift	1	3.33
Total Flight Planning Tasks per Shift	60	40

Figure 23: Error bar of relative range of task service times comparing short-haul and focus flight planning

4.5 Model Synthesis

As a final step to extend the model validation process and build general confidence for using SHADO in future operations research, sensitivity analyses were conducted. First, researchers evaluated how errors in input parameters may skew results. Highlighting any variables for which SHADO outputs have significant sensitivity toward ensures that future researchers understand areas with further data validation requirements. Researchers ultimately determined the generalizability of SHADO by identifying the top internal parameters that had the greatest significant impacts on key performance metrics across three different domains. The ten internal parameters of SHADO were each tuned from their default values and results of dispatcher utilization and human error were analyzed for freight railroad, commuter railroad, and airline operations.

With a discrete event simulation model as complex as SHADO, it is important for future users and researchers to understand how sensitive key performance metrics are to changes in input parameters. Sterman (2000) suggests sensitivity analysis as a good test for the model's robustness within reasonable uncertainty in assumptions. Performing a sensitivity analysis includes five major steps:

1. Identifying main input/internal parameters
2. Identifying key outputs
3. Setting parameter levels
4. Designing experiments
5. Analyzing sensitivity of outputs to parameter levels

The basis for these steps is evaluating the sensitivity of the model to errors in parameter values. The purpose is to understand the relationships between inputs and outputs, which can be useful for identifying important variables or errors in model development. Researchers systematically varied parameters of SHADO to determine whether the simulation output is significantly impacted by specific parameters. This information improves model transparency and credibility with stakeholders. The five-step process are detailed for the two analyses conducted with 1) task input parameters and 2) internal design and staffing variables of dispatch operations centers.

4.5.1 Model Sensitivity to Inaccuracies in Task Input Parameters

4.5.1.1 Identifying Main Input Parameters

SHADO is composed of numerous variables. For this analysis, it was important for researchers to prioritize those that have the greatest potential to significantly impact results if estimates are inaccurate. The dataset of task input parameters is composed of interarrival times and mean service times which were thoroughly gathered but with limited access to the industry. So, it is good practice to investigate how this limited access may influence model results for railroad and airline domains. Just as important, these two parameters are the main inputs that SHADO users can change on the platform using data that they themselves have collected on each task. Therefore, it was meaningful to define model confidence within determined margins of inaccuracy for these parameters. If resources are limited for data collection procedures, future users will be aware of which areas to prioritize investments for more precise estimates of the real world.

4.5.1.2 Identifying Key Outputs

For this research, utilization is the primary key performance indicator (KPI) as it is a proxy for workload. Utilization represents an operator's busy time on tasks, so we can expect differences in task service time to correlate with deviations in utilization. The other key indicator of performance is human error, which is described in Table 16 as well. Human error in the form of slips of actions (*missed tasks*) and lapses of memory (*incomplete tasks*) would be expected to increase with decreases in task inter-arrival times and increase with increases in task service times. Mistakes (*failed tasks*) are expected to not change one direction or the other since there is not a direct relationship between human error probability and time in this context. SHADO considers several variables and runs random inter-arrival and service time processes per task, making the model more complex than a theoretical queuing system. Consequently, we cannot expect the different output statistics to respond identically to variances in different default parameters.

4.5.1.3 Setting Parameter Levels

The mean interarrival time and mean service time values were each varied by -75%, -20%, -5%, +5%, +20% and +75%. These six levels allowed researchers to analyze how extremely low, extremely high, and incremental deviations from original inputs may affect key performance indicators. If an extremely low deviation in an input (+/-5%) results in an extremely high deviation in an output, then that is evidence of extreme sensitivity in that relationship. On the other hand, if extremely high deviation in an input (+/-75%) results in relatively little change in output, then the input parameter may not have much significance in the system.

Table 16: Key performance indicators of interest from sensitivity experiments

Output	Computation	Values
Utilization	[illegible]	R = total number of replications; r = index of replications; I = total number of time intervals; i = index of intervals; ρ = utilization; E = number of tasks erred
Human Error	[illegible]	

4.5.1.4 Designing Experiments

Three sets of experiments were conducted across commuter and freight railroad, and airline dispatch operations. The procedures were completed by running twelve (six levels of two factors) experiments for each operational domain as presented in Appendix J. Each experiment was simulated 365 times, involving positive or negative variations to the task inter-arrival and service time parameters, one-factor-at-a-time (see Equation 3). The *train movement*, *other communications*, and *short-haul flight planning* tasks were respectively selected for testing in each domain. When compared with all other tasks in original operations, these three tasks were found to contribute the most to dispatcher workload and error. Table 17 displays the simulation settings that each experiment was based on originally with one dispatcher and one “fleet.”

Table 17: Baseline simulation settings for sensitivity analysis of one dispatcher's performance

Domain	Shift	Shift Turnover	Task Type
Freight Railroad	8-hour	Uniform (minimum = 5, maximum = 10 minutes) at the beginning and end of shift	*Other Communications* with Interarrival Time on Expo (mean = 17.4 mins) and Service Time on Expo (μ = 3.375 mins)
Commuter Railroad		Exponential (μ = 5 minutes) at the beginning and end of shifts	*Train Movement* with Interarrival Time on Expo (μ = 6.8 mins) and Service Time on Expo (μ = 1.7 mins)
Airline	10-hour	Expo (μ = 15 mins) at end of shift	*Short-Haul Flight Planning* with Interarrival Time on Expo (μ = 16.1 mins) and Service Time on Lognormal (μ = 3.3, standard deviation = 1.2 mins)

Equation 3: The transformation of random variables of the arrival or service processes

In Equation 3, the deviation element Δ was varied negatively and positively by 5%, 20%, and 75% from the original 0% change. i_d is the transformed random variable that represents either mean time between the most recently arrived task and the upcoming task or the mean time for a dispatcher to process the task in simulation.

4.5.1.5 Analyzing Sensitivity

The experiments were run for 365 replications and the average of each KPI was recorded. Table 18 shows the results for the default condition in which no parameters were varied.

Table 18: Average key performance metrics from original conditions

Domain	**Dispatcher Utilization**	**Human Error (Missed or Incomplete Tasks)**
Commuter Railroad	26%	1.47 tasks/shift
Freight Railroad	23%	0.43 tasks/shift
Airline	22%	0.14 tasks/shift

Figure 24, Figure 25, and Figure 26 show how sensitive dispatcher utilization and human error are to the interarrival time parameters for each operational domain. Across all the domains, human error was found to be more sensitive than dispatcher utilization to deviations in the input parameters, especially in the cases of a -75% deviation in interarrival times. Figure 27, Figure 28, and Figure 29 show how sensitive the KPIs are to the service time parameters for freight, commuter, and airline dispatch operations. Any deviations in service time distribution were generally found to cause smaller deviations in both KPIs than changes resulting from deviations in interarrival time. The major effect of reduced task interarrival times on increased human error can be explained by the fact that tasks would be arriving so frequently that it would infeasible for a single dispatcher to handle most tasks before they expire either while waiting in the queue or during the service processes following too long delays.

Contrarily, for any increase in interarrival time, utilization responds with an opposite change that is near or smaller in magnitude. For example, a 20% increase in interarrival time respectively leads to 16%, 12%, and 17% average decreases in utilization for commuter, freight, and airline operations.

Figure 24: Freight railroad dispatcher utilization and errors from deviations in interarrival time process of *other communications* tasks

Figure 25: Commuter railroad dispatcher utilization and errors from deviations in interarrival time process of *train movement* tasks

Figure 26: Airline dispatcher utilization and errors from deviations in interarrival time process of *short-haul flight planning* tasks

Figure 27: Freight railroad dispatcher utilization and errors from deviations in service time process of *other communications* tasks

Figure 28: Commuter railroad dispatcher utilization and errors from deviations in service time process of *train movement* tasks

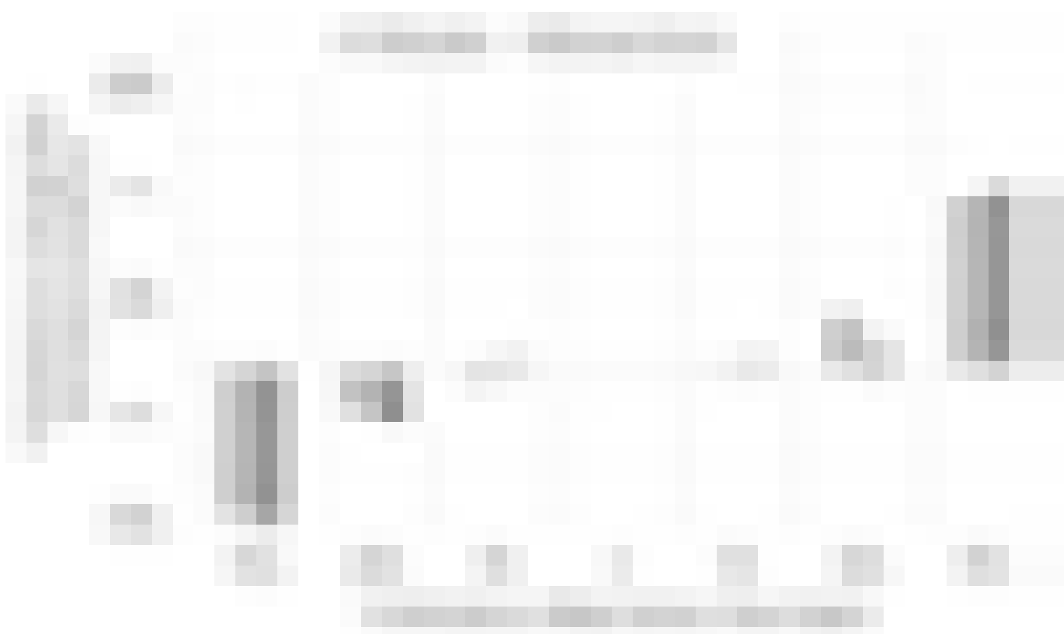


Figure 29: Airline dispatcher utilization and errors from deviations in service time process of *short-haul flight planning* tasks

On average, increasing service time leads to higher workload and more human errors per shift. Results from this sensitivity analysis revealed that human error indicators are most sensitive to changes in both interarrival and service time parameters of tasks. Human error is one of the most important metrics of performance for dispatch centers today as safety is a matter of regulatory compliance and business success. Unfortunately, this information is limited. For the railroad dispatch operations, the only valid source identified was based on results from a human reliability assessment with train crew in the United Kingdom (Gibson, 2012), not dispatchers. Researchers studying human performance of dispatch operators in a similar rail operational setting in Denmark (Thommesen & Andersen, 2012) also referred to this core study.

In aviation, an industry from which the human factors discipline grew, the focus has been on pilot, maintenance or air traffic controller error (U.S. Congress & Office of Technology Assessment, 1988). Considering the central role that dispatchers play in airline network operations and the mandated shared responsibility a dispatcher has with pilots-in-command for the safety of multiple flights at a time (Federal Aviation Administration, 2005), there have been relatively few publications on dispatcher error. Therefore, this modeling limitation of access to human error data presents an opportunity for future research to better validate results of SHADO.

4.5.2 Model Sensitivity to Internal Design & Staffing Variables

4.5.2.1 Identifying Main Internal Variables

SHADO is composed of ten internal variables that users can adjust on the platform depending on the staffing or design of their concept of operations. Therefore, determining the margins of error for these parameters would be meaningful to users. The ten internal variables are fleet size, fleet heterogeneity, fleet autonomy, environment, team coordination, artificial intelligence decision aids, shift schedule, team size, team expertise, and operator attention allocation strategy.

The first four variables listed contribute to majority of tasks that dispatchers are responsible for handling in railroad and airline operations. Therefore, it is expected that variations in these internal parameters would likely impact dispatcher utilization. The next two variables are likely to impact dispatcher error because of the human and automation assistance provided to dispatchers. The next three variables, largely based on staffing decisions by operational managers, may have differing effects.

For example, the longer a dispatcher is at work with more hours in their shift, the greater fatigue they are expected to experience. Fatigue would in turn lead to more time on tasks and perhaps increase in workload and human errors due to delays. Another variable in staffing is team size. The more dispatchers working together on the same fleets and tasks, the less each dispatcher is expected to be utilized. Then, the team expertise has a range of possibilities in that on specialist teams it can be expected that one dispatcher may end up with higher workload than others while on generalist teams, the workload would be balanced.

Finally, the dispatcher's attention allocation strategy's influence on workload or error is not clear but it is something that many experts debate about in the industry. During visits to the dispatch operations centers, experienced dispatchers often highlighted their preference toward shortest-task-first while novice dispatchers were more likely to take the first-in-first-out approach. Studying how changes to this final variable as well as the earlier nine will provide a more holistic understanding of SHADO and the real-world systems of freight, commuter, and airline dispatch operations.

4.5.2.2 Identifying Key Outputs

The two key performance indicators selected for the sensitivity analysis with task input parameter deviations in Section 4.5.1 were again identified here. Utilization represents dispatcher workload. Human errors involve missed as well as unfinished tasks. These results are averaged and used to study SHADO's sensitivity with changes to internal design and staffing parameters.

4.5.2.3 Setting Parameter Levels

The ten internal variables were systematically varied within the realm of normal or reasonable parameters of dispatch operations centers to study the resulting change in the key outputs. Table 19 and Table 20 detail the levels used in sensitivity analysis in the two railroad operations and the airline operations. Each factor was varied with one of the options that a user may choose in simulation. There were two levels for operator strategy; three levels of fleet size, fleet heterogeneity, fleet autonomy, team size, and team expertise; four levels of exogenous events and artificial intelligence, and five levels of team coordination which also included levels of artificial intelligence support. The basic levels were similar across the different operational domains. However, the number of tasks and fleets was incongruent with airline operations, so the organization of the final specialist level was across of a team of three as opposed to a team of two dispatchers. Additionally, the original shift schedule in airline operations is two hours longer than railroad operations so the level of variance differed in that effect.

Table 19: Levels of internal parameter variance for freight and commuter railroad dispatch operations factors

Railroad Factors	*Original (1)*	*Levels of Variance*				
		(2)	*(3)*	*(4)*	*(5)*	*(6)*
Fleet Size	1 railroad set	2	3			
Fleet Heterogeneity	Homogeneous	2 v 1	1 v 1 v 1			
Fleet Autonomy	None	Some	Full			
Exogeneous Events	None	Both	Derailment	Weather		
Artificial Intelligence Decision Aids	None	Equal Operator	Some Task Assistance	Full Task Assistance		
Team Coordination	None	Some	Full	Full w/ Some AI	Full w/ Full AI	
Team Size	1 dispatcher	2	3			
Team Expertise	Generalists	50:50 Specialists	25:75 Specialists			
Shift Schedule	8 hours	2	4	6	10	12
Operator Strategy	FIFO	STF				

Table 20: Levels of internal parameter variance for airline dispatch operations factors

Airline Factors	*Original (1)*	*Levels of Variance*				
		(2)	*(3)*	*(4)*	*(5)*	*(6)*
Fleet Size	1 flight type	2	3			
Fleet Heterogeneity	Homogeneous	2 v 1	1 v 1 v 1			
Fleet Autonomy	None	Some	Full			
Exogeneous Events	None	Both	Medical	Weather		
Artificial Intelligence Decision Aids	None	Equal Operator	Some Task Assistance	Full Task Assistance		
Team Coordination	None	Some	Full	Full w/ Some AI	Full w/ Full AI	
Team Size	1 dispatcher	2	3			
Team Expertise	Generalists	50:50 Specialists	33:33:33 Specialists			
Shift Schedule	10 hours	2	4	6	8	12
Operator Strategy	FIFO	STF				

4.5.2.4 Designing Experiments

Another three sets of experiments were conducted across freight and commuter railroad, and airline dispatch operations. The procedures were completed by running 26 (one to six levels of ten factors) experiments for each operational domain as presented in Appendix K. Each experiment was simulated 365 times, involving variations to the internal design and staffing parameters, one or two factors at a time. Each experiment was based on original parameters from Section 4.4 with all earlier defined tasks in each simulation of the afternoon shifts.

4.5.2.5 Analyzing Sensitivity

The experiments were run for 365 replications and the average KPIs were recorded. Table 21 shows the results for the default condition in which no parameters were varied.

Table 21: Average key performance metrics from original conditions

Domain	**Dispatcher Utilization**	**Human Error (Missed or Incomplete Tasks)**
Freight Railroad	51%	1.5 tasks/shift
Commuter Railroad	48%	2.5 tasks/shift
Airline	59%	12 tasks/shift

Figure 30 – Figure 39 show how sensitive dispatcher utilization and human error are to the internal design and staffing parameters for freight railroad operational domain. Remaining charts for the commuter railroad and airline dispatch operations are included in Appendix L. Results across the domains showed that the number one factor for dispatcher performance was fleet size (see Figure 30). In the real-world, this means how many railroad operations or how many flight operations a dispatcher is managing during their shift is the greatest determinant of workload and error. Other factors such as operator strategy and exogenous events had relatively little influence on how a dispatcher performed overall.

Figure 30: Freight Dispatcher Workload and Errors with Changes to Fleet Size

Figure 31: Freight Dispatcher Workload and Errors with Changes to Fleet Heterogeneity

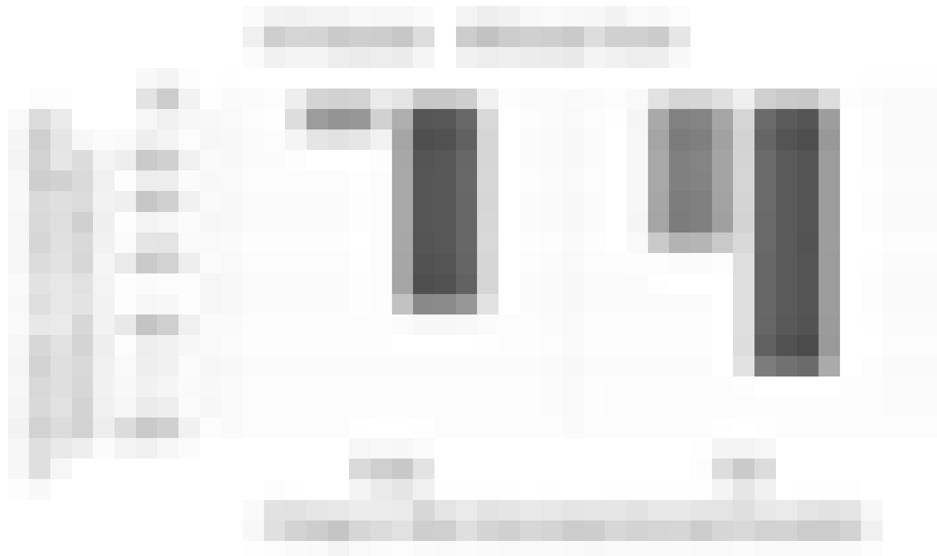


Figure 32: Freight Dispatcher Workload and Errors with Changes to Fleet Autonomy

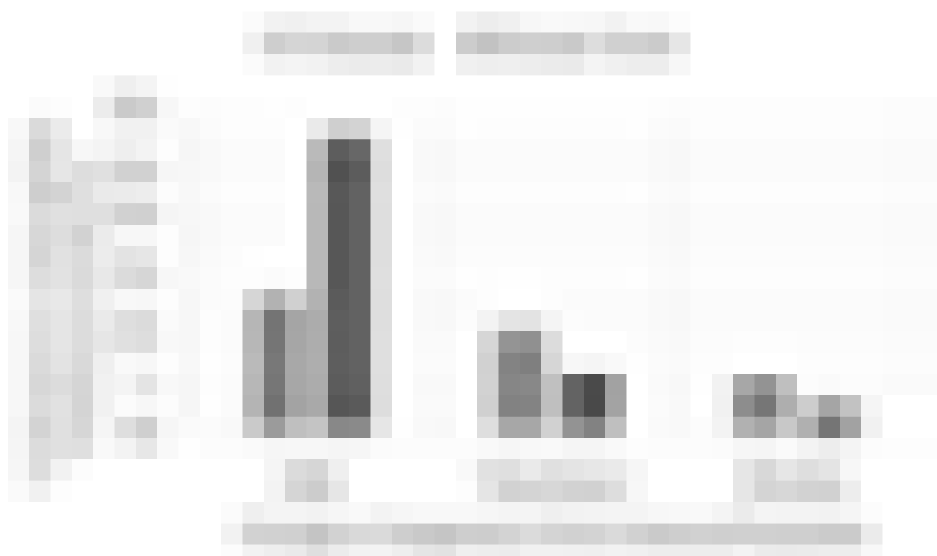


Figure 33: Freight Dispatcher Workload and Errors with Changes to Exogenous Events

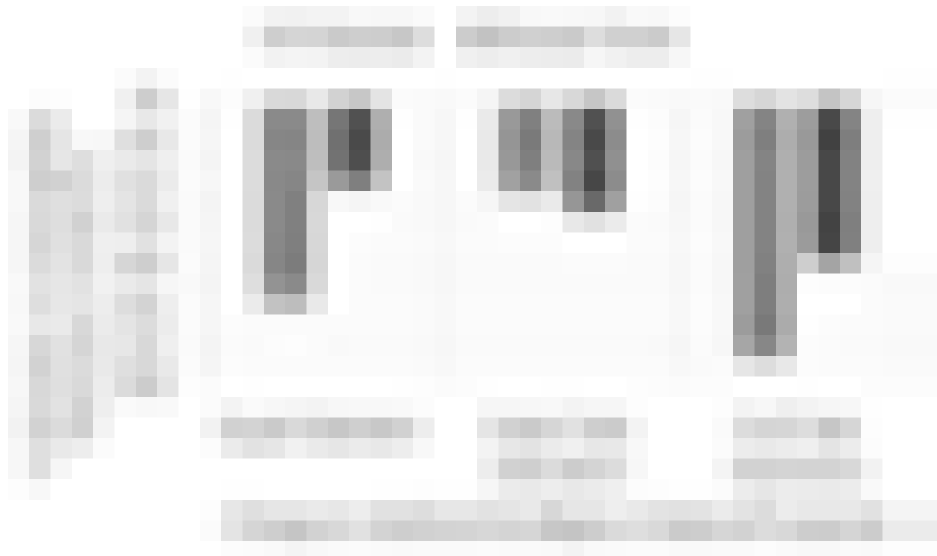


Figure 34: Freight Dispatcher Workload and Errors with Changes to AIDA

Figure 35: Freight Dispatcher Workload and Errors with Changes to Team Coordination

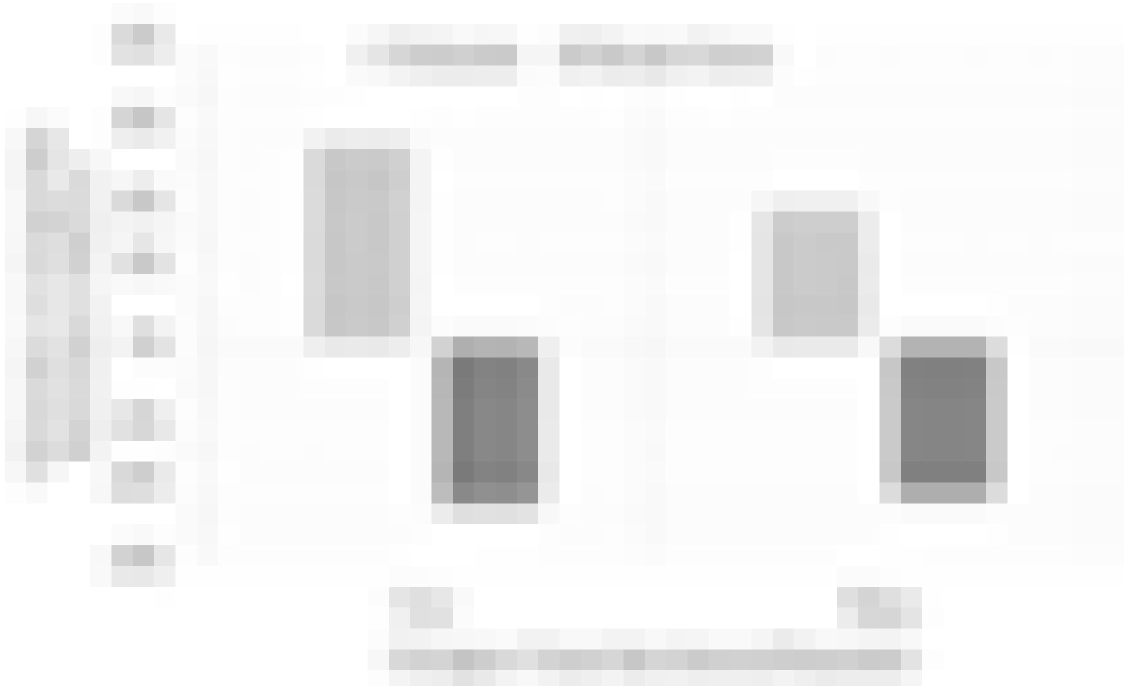


Figure 36: Freight Dispatcher Workload and Error Deviation with Changes to Team Size

Figure 37: Freight Dispatcher Workload and Errors with Changes to Team Expertise

Figure 38: Freight Dispatcher Workload and Errors with Changes to Shift Schedule

Figure 39: Freight Dispatcher Workload and Errors with Changes to Strategy

As shown in Figure 35, Figure 31, and Figure 32, team coordination, fleet heterogeneity, and fleet autonomy were freight railroad operational factors that followed behind the significance of fleet size. In commuter railroad and airline dispatch operations, as reported in Appendix L, team expertise, fleet heterogeneity, and team size instead follow fleet size in how influential they are on performance. In the next section, we will discuss how conducting these 78 experiments across three operational domains revealed some insightful results about what matters when designing and staffing dispatch operations centers.

4.5.3 Model Generalizability for Different Dispatch Operations

When we reflect on the in-depth sensitivity analyses of 114 experiments conducted across railroad and airline domains, freight and commuter operations, 8-hour and 10-hour shifts, and the multitude of diverse characteristics, we see how generalizable SHADO for modeling different dispatch operations. Thirty-six of the 114 experiments were run by deviating task input parameters from the norm in three operational domains. Results showed that allowing tasks to arrive up to 75% more frequently than normal would be detrimental to operational safety due as dispatchers' workload would approach 100% and dispatcher error from missing starting on and completing tasks would spike uncontrollably.

Changes to average task service times led to expected results in dispatcher performance. Reducing or increasing task times generally led to workload and error also reducing or increasing in tandem. However, these changes were generally smaller in magnitude. So, across the different dispatch operations when prioritizing resources for data collection or task redesign, resources should go toward the interarrival time factor first. Additionally, better recording human error and effects of different kinds of error on operational performance would support future efforts to test changes in dispatch centers.

Along with these general findings, there were some contrasting results. Freight railroad dispatch operations was found to be more robust to deviations in task frequency and duration in that dispatcher error did not change as much as it did in commuter railroad or airline settings. In fact, when service time was increased or decreased by the most extreme deviations, human error had a smaller response than workload.

Contrasts and similarities can also be found by studying how deviations in task input variables compare to changes in how dispatch operations are designed or staffed as we see with the remaining 78 of 114 experiments conducted. Fleet size had approximately twice the impact on human error in freight railroad domain more than in either of the other two domains which also had fleet size as their most significant factor of performance. The freight dispatch desk does entail 10 different tasks whereas commuter entails nine and airline entails eight tasks. So, the explanation may be that the impact of these combinations of tasks doubling or tripling would indirectly decrease time between tasks and more time overall busy with tasks such that it outweighs the impact of tuning the most influential task to arrive extremely frequently.

4.6 Chapter Summary

These sensitivity analyses were the final steps in the process of validation through verification, gathering and refining data, and testing the simulation as an open-box and as a black-box with SME review throughout, enhanced both internal and external confidence in using SHADO. Conceptual models were affirmed by SMEs. A computational model was developed from this conceptual model and verified in multiple methods including with software accepted as the industry standard.

Different sources of data were available at Rio Grande Pacific Company and Horizon Air to validate input parameters for simulating dispatcher workload in the two companies. At RGPC, many historical documents were used to categorize and estimate tasks and times. At Horizon Air, a questionnaire was created to solicit similar information directly from several dispatchers.

The online SHADO platform gave the chief and senior dispatchers at RGPC and Horizon Air the ability to participate in open-box testing. This was a medium to gather additional feedback to refine the underlying SHADO model and conduct black-box testing. Black-box testing involved statistical goodness-of-fit testing, Turing testing, and boundary condition testing to ensure that the workload predictions in SHADO are not significantly different from what one would anticipate in the real-world based on the wide range of data sources.

The final steps of sensitivity analyses built general confidence in using SHADO across different operational domains. Multiple input and internal variables were tested to investigate the impact of task design, fleet characteristics, operations center dispatch support systems and staffing and individual approaches on dispatcher workload and error. Dispatcher performance was found to be most sensitive to changes in task interarrival times and in fleet size in freight railroad, commuter railroad, and airline operations.

Through these quantitative and qualitative approaches, SHADO has been shown to reliably model real-world dispatcher workload in railroad and airline dispatch operations at Rio Grande Pacific Company and Horizon Air. In the next chapter, SHADO will also be shown to be useful in modeling for the future. With the confidence gained in modeling present-day operations, stakeholders including managers in both domains will use SHADO to support decision-making on concepts of operations that are otherwise too expensive to test in the real world.

5. Predictive Simulation

SHADO can be used predictively to test what-if scenarios to support operations managers in understanding *when* and *why* their dispatchers may be over- or under-utilized during a given shift, which can be especially useful to determine what the impact will be of staffing or design changes. Such models would be useful in investigating *how* a dispatcher's workload may change with changes in the fleet and/or operations center. SHADO can be used to explore how a dispatcher's work, as well as overall system efficiency and safety, could be improved. Having developed and validated the discrete event simulation model used in SHADO, the next step is to illustrate how SHADO can be used to conduct a prospective study of human workload in different system configurations of railroad and airline operations.

5.1 Railroad Dispatch Operations

Companies like Rio Grande Pacific Company are often interested in how much more work their current workforce can reasonably manage. The chief dispatcher has been considering options for expanding their shortline freight business. He is interested in knowing what the potential staffing implications would be if RGPC increases the number of railroad operations.

How might the workload of the single dispatcher currently at the desk be affected by such changes and when may it be too much work for them to safely and efficiently handle? It remains unclear how much RGPC could scale up their operational size from the current set of 12 shortline freight railroads with their current staffing of just one dispatcher at that desk. Hiring a new dispatcher is a long process of recruiting, interviewing, assessing, training, and managing to reduce the likelihood of turnover.

On the commuter desk, RGPC plans to install positive train control over the next year in keeping with the congressional mandate (110th US Congress, 2008). So, the chief dispatcher used SHADO to explore another question, which is how might the integration of automation into the local commuter dispatch operations affect dispatcher workload at that desk? Many railroad companies have expressed concern about the impact of PTC technology, a form of automation, on their railroad operations. PTC is a system of three main components: trackside devices, onboard locomotive computers, and the dispatch operations office. Trackside devices communicate with the onboard locomotive computers to provide switch positions and other vital information about the railroad track conditions. The onboard locomotive computer is always on to compute the current train speed and brake time requirements.

The dispatch operations component of PTC functions as a centralized controller of the network of trains. From the network operations control center, dispatchers can monitor locations and velocities of trains and issue movement authority. Along managing train movement, dispatcher also maintain safety and efficiency of other operations on the track include maintenance-of-way crews who may be doing railroad construction work. Testing such operational changes in the real world would take more time and money than RGPC can afford, so SHADO was the ideal tool to investigate the potential impacts of expanding operations as well as installing automation on dispatcher workload.

5.1.1 What if RGPC increases the size of short-line freight rail operations?

To examine the impact of increasing the size of short-line freight rail operations, SHADO was tuned to multiply, by a factor of size, the average number of task arrivals from the present-day operations validated for their current 12 railroads. SHADO was run for 300 days using the default parameters presented in Section 4.4 for each shift of freight operations for each of the blocks of 12 additional railroads. All other variables except for railroad operational size were held constant. Figure 40 presents results from simulating AM, PM, and ON shifts on the freight dispatcher desk.

Figure 40: Average Dispatcher Utilization Results with Increase in Size of Railroad Operations in AM, PM, and ON Shifts on Freight Desk

In each consecutive run, the railroad operational size was multiplied from the present-day 12 railroads to 24 railroads, 36 railroads, 48 railroads, and so on up to 96 railroads on the 8th trial. The freight dispatcher during the morning shift reached 100% utilization by the third trial at 36 railroads, during the afternoon shift by the fourth trial at 48 railroads, and during the overnight shift by the eight trial at 96 railroads. From analysis in Chapter 4, it was found that time of day determines the rate of arrival for certain tasks. The overnight shift was shown to be slower which resulted in present-day dispatchers averaging 32% utilization compared to the 50% or 60% measured for afternoon and morning dispatchers. So, as was found in this prospective analysis, the overnight desk can manage two to four times as many railroads as their coworkers who work during the other times of the day.

These results should be taken with caution. Working dispatchers to 100% may be possible but may not be favorable for human performance. This is important to note too as an average value less than 100% may also result in many periods at or near 100% which would be unacceptable for dispatchers working in such safety critical operations. So, the question for the short-line freight rail dispatch desk may be asked in two questions. First, if RGPC increases the size of freight railroad operations under management, at what point may the sole dispatcher begin to underperform? And second, at what point might we need to increase dispatcher staffing to handle the increase in operational size at the freight desk?

During the morning and afternoon shifts, the results suggest that, on average, dispatchers could handle up to 24 railroads and during the overnight shift, the dispatcher may be able to oversee almost 36 railroads. By the 36th railroad, the morning dispatcher would be maxed out. This would be the case during the afternoon shift by the 48th railroad. On the contrary, the overnight shift results show that on average, the dispatcher capacity would not reach maximum until 96 railroads. However, it is not recommended to work dispatchers at these maximum conditions for long periods of time as human error and other delays may become unreasonable for safe and efficient operations.

Maximum utilization is 100%, at which point there are no additional mental resources available for personnel to accomplish additional tasks. The upper bound of workload for optimal sustained operator performance has previously been found at the 70% utilization threshold (Cummings & Nehme, 2010; Rouse, 1983). Levels of utilization below 30% have also been associated with poor performance as operators are prone to boredom and distraction (Cummings et al., 2016, 2013; Yerkes & Dodson, 1908).

These results suggest that adding operations for this specific setting will increase a dispatcher's workload and that additional dispatchers will likely be needed for the morning and afternoon shifts somewhere around 18 railroads. However, a single overnight dispatcher can theoretically handle significantly more, upwards of 48 railroads.

Before deciding the expand the number of railroads under management, human-in-the-loop experiments may be useful to determine the likely distribution of operator utilization over the course of each shift at smaller increments of increasing by one railroad at a time. The data available at this time is limited in that one must assume that each additional block of 12 railroads behave identically which may not be the case with new railroad operations across the nation.

The chief dispatcher agreed with the trend found per shift of railroad operations expansion. According to the chief, today's morning dispatcher is nearing their threshold for high workload which is around 5.5 hours of talk-and-listen time during the 8-hour shift. Results from SHADO show utilization at the desk today to be 60% and this concurs with the fact that 5.5 out of 8 is 68.75%. This also supports the long-held theory that 70% utilization is a real threshold of workload across domains of dispatch operations. The chief dispatcher found the results of this prospective analysis to be useful in long-term planning about when to hire additional staff and decision-making about which times of day to accept additional business opportunities. In this way, RGPC can maximize profits in dispatch operations by managing more railroads while maintaining staffing at levels that give dispatchers just enough to do without overburdening any teammate with work.

5.1.2 What if RGPC installs automation in the local commuter rail operations?

On the current RGPC commuter desk, the dispatchers digitally control tracks and signals for one railroad. Positive train control (PTC), a system that automates emergency braking for each locomotive in a rail network, could increase the workload of these dispatchers. Current designs of PTC would require additional displays that dispatchers would need to interface with for *train movement* tasks. Dispatchers may become more involved in the control loop than ever before to remotely manage multiple train systems in the case of crew error or computer malfunction. Although there are expected safety benefits for crew working on the railroad, it is not clear how PTC may ultimately impact dispatchers.

Interestingly, a previous study of European dispatch operations centers (Sharples, Millen, Golightly, & Balfe, 2010) found that automation in operations led to dispatchers spending less time on interaction and paperwork tasks. On the other hand, dispatchers were found to spend more time on planning, communicating via phone calls or in the transfer-of-duty periods, and doing miscellaneous tasks.

In SHADO, dispatcher interaction with PTC was reflected through longer dispatcher time on *bulletins*, *temporary bulletins*, *other communications*, *miscellaneous* and *transfer-of-duty* tasks. Less time was spent on *train movement*, *bulletin printing*, *weather recording*, *notetaking*, and *reporting*. The before-and-after using data from the European study to adjust input parameters are detailed in Appendix M.

SHADO ran 300 simulated days per four scenarios per the three shifts, resulting in 12 experiments. The results of present-day operations (using commuter desk settings from Table 6) were compared with results of the most likely, best-, and worst-case scenarios during the morning (AM), afternoon (PM), and overnight (ON) shifts with the PTC augmentation. The average predicted dispatcher utilization results are plotted in Figure 41 and presented in detail in Appendix M.

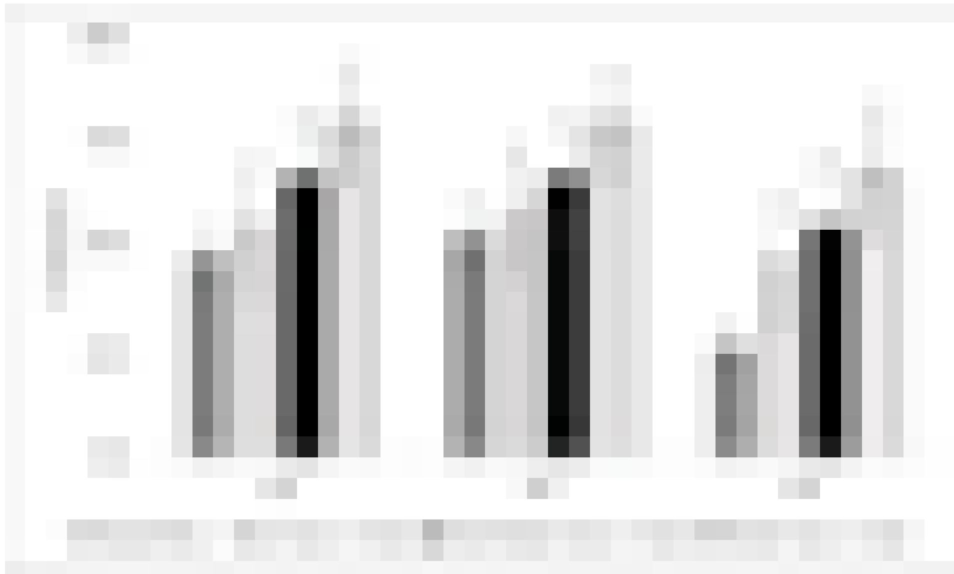


Figure 41: Average and S.D. Utilization on Commuter Desk per Shift per Automation Scenario

Workload of the present-day commuter dispatcher desk were compared with the potentially best-case scenario with automated operations for each shift. The results of the statistical analyses are presented in Table 22. The difference between utilization of the commuter dispatcher during present-day operations versus during the best case of automated operations in AM, PM, and ON shifts are statistically significant.

Table 22: Comparison of Means from Present-Day to Best-Case with Automation Commuter Dispatcher Utilization

SHIFT	AM	PM	ON
AVERAGE DIFFERENCE	8.4%	5.6%	19.6%
STANDARD ERROR	0.009	0.01	0.009
T-STATISTIC	8.963	5.522	21.35
DF	598	598	598
SIGNIFICANCE LEVEL	P < 0.0001	P < 0.0001	P < 0.0001

Although statistically significant results were found in this second prospective analysis, in the context of dispatcher workload, the impact of best or most likely cases of automation on performance during the AM and PM shift may not be practically significant. On average, the changes in workload for the commuter desk during the morning and afternoon shift in the best-case of automation may not result in noticeable differences in dispatcher performance as utilization is maintained between 30% - 70%. It is only in the worst-case scenario with automation in which average workload results near or beyond 70% utilization where dispatcher performance may begin to decrement.

The greatest difference in utilization with automation versus in present-day operation was found during the overnight shift, which resulted in approximately 20% greater workload in the best-case scenario. It appears that for the overnight shift, additional time on tasks from the new work paradigm may benefit the dispatcher's average workload, increasing it from low to moderate levels above 30% utilization at 47%. However, this analysis does not consider how unmanageable this role may be if the one overnight dispatcher manages both commuter and freight rail operations as RGPC currently staffs today. Overall, these results suggest that increased locomotive automation could increase workload for a dispatcher in this specific setting, but within manageable levels.

The data and analyses used here are based on work from operations in Europe. Therefore, decisions based on the results presented here alone have limited confidence. Further data and analyses in the context of U.S. operations as companies like RGPC implement PTC would provide practical support on the impact of automation on dispatcher workload. In discussing results from both case studies with the chief dispatcher at RGPC, there was validation based on the SME's experience with railroad dispatch operations.

The chief dispatcher agreed with the method used to project the impact of future changes in the railroads on personnel. The key metric of performance in their operations center is talk-and-listen time. This corresponds with the estimates gathered from the European study to simulate the impact of scenarios with automation on dispatcher busy time. The chief dispatcher SHADO to be advantageous for him to use as he could systematically adjust parameters without much ado about the interactions of different system components. Additionally, the ability to simulate across multiple days of data was useful to explore the potential range of results.

Before this study, the chief dispatcher notionally believed that while there are many benefits to safety along the railroad, the addition of PTC may have an adverse effect on the commuter dispatcher, by significantly raising his or her usual workload. So, seeing the results from SHADO confirmed but also tempered some of his concern. As shown earlier, workload is projected to increase but within manageable levels for the dispatcher staff.

5.2 Airline Dispatch Operations

Airline companies are investigating transformations needed in their operations to maintain or improve levels of performance. The introduction of on-demand mobility, which is a form of personal air transportation that meant to shorten the travel time within and between metropolitan regions by leveraging advanced vertical takeoff and landing (VTOL) designs and other technologies (e.g. vehicle autonomy and distributed electric propulsion) (Nneji et al., 2017), would change an airline’s fleet size with more, albeit smaller, aircraft and flight schedule with less predictable über-short-haul routes.

The autonomy in their fleet would also change with higher levels of vehicle-to-vehicle communication. And if there is a diversity of flight management requirements within this new category, fleet heterogeneity would also change. Given all these changes that are revolutionary as opposed to evolutionary, a dispatcher workload model would be very beneficial in the planning of network operations.

Other business changes in airlines prompt further considerations into how operations may be transformed. In January 2019, Horizon Air moved their dispatch operations center from Portland, Oregon to the headquarters of their parent company, Alaska Air Group, in Seattle, Washington. This move was not expected to change the overall fleet and environmental parameters that Horizon dispatchers are used to managing. However, the move did present some challenges and opportunities in deciding how to staff and train dispatchers. With this move, Horizon had an opportunity to improve dispatcher training and SHADO was employed to inform such planning as early as June 2018.

Another present consideration of Horizon Air today is how dispatchers are assigned their roster of flights. The process indiscriminately assigns flights based primarily on one metric: the average number of flights per desk per hour. Alaska Airlines, on the other hand, distributes flights based primarily on geography. Each dispatcher at Alaska is responsible for flights heading to specific destinations. This is believed to help dispatchers build on their expertise in understanding the unique trends of their area. As with the rail case studies, SHADO can also help diagnose possible present-day workload inefficiencies in these airline operations. By using SHADO to explore these two different philosophies, stakeholders may better understand how to best merge the dispatchers in Seattle.

5.2.1 What if new dispatchers are trained to operate differently?

There are three levels of dispatch operations decision-making that can be varied in SHADO. The first two are on the organizational and team levels dealing respectively with strategic design factors (e.g. artificial intelligence decision aids) and tactical staffing factors (e.g. shift schedule) in the operations center. The final level is that of the individual dispatcher. On this level, the internal variable that represents decision-making is operator attention allocation strategy. Different dispatchers may employ different strategies in how they allocate their attention.

In SHADO, different strategies are represented by how tasks are queued for dispatchers. These strategies include first-in-first-out, shortest-task-first, and priority and they are initially set by users per each operator type or team. The priority scheme is defined by user preferences in which case the user would rank tasks and set any tasks that are of essential priority and any that may be interrupted by higher priority tasks.

This operator variable is one that can be changed in the real-world through training. Dispatchers may be taught to strategize how they approach their task load with certain sequencing or prioritizations rules. SHADO can be used to explore how changes in task sequencing strategies that reflect different training paradigms may ultimately affect dispatcher workload and performance during normal and irregular conditions.

A comparative analysis of dispatcher workload for eight simulated scenarios was conducted, adjusting the operator strategy or operational condition in each scenario.

Four levels of strategy and two types of conditions were defined (Table 23). Normally, novice dispatchers begin working with a first-in-first-out (FIFO) approach to manage emergent tasks. This means that the operator responds to tasks in the chronological order in which they arrive into the system. More experienced dispatchers have reported taking a shortest-task-first (STF) approach. To represent alternate training regimes, SHADO was run with three additional queuing schemes:

- **STF approach**. The operator brings to the front of the line tasks they know will take less time to complete.
- **Priority #1 (Pr1) approach**. The operator handles 1) flight following tasks then 2) flight planning tasks then 3) miscellaneous tasks.
- **Priority #2 (Pr2) approach**. The operator handles 1) *focus flight*-related flight following and planning tasks along with *long-haul* flight following tasks, then 2) *long-haul* flight planning, *short-haul* flight following and planning, then 3) miscellaneous tasks.

Irregular operating conditions in this study include both a medical emergency and poor weather during one shift. The fleet presents the medical emergency as an uninterruptible task that would disrupt whatever the dispatcher was attending to once it arrived into the system and require 20 to 40 minutes of dispatcher attention. The environment presents the factor of poor weather with a heightened frequency of *flight following* tasks for all three types of flights. Normal operations do not involve any such exogenous events. Exploring both irregular and normal conditions helps to identify if any of the four strategies are robust enough to allow dispatchers to maintain similar levels of workload on days when things do not go as planned as they would during more typical days.

Table 23: Design of Experiments

Experiment	*Level of Strategy*	*Type of Condition*
1	FIFO	Normal
2	STF	Normal
3	Priority #1	Normal
4	Priority #2	Normal
5	FIFO	Irregular
6	STF	Irregular
7	Priority #1	Irregular
8	Priority #2	Irregular

SHADO was simulated for 365 replications of the initial settings at each level of strategy for either condition. A 4 (strategy) x 2 (condition) two-way factorial analysis of variance (ANOVA) was calculated on dispatcher utilization. Assumptions of normality and homogeneity of error variance were met in this analysis.

The main effect of strategy was significant, $F(3, 2912) = 165.21$, $p < .001$. The main effect of condition was significant, $F(1, 2912) = 56.23$, $p < .001$. The interaction effect of strategy x condition was not significant, $F(3, 2912) = 0.42$, $p = 0.7360$. Dispatchers using the FIFO, STF, Priority #1, and Priority #2 strategies do not experience statistically significantly different workload (as measured by utilization) based on operational condition. The p-values for the four types of operator strategies, two types of operational conditions, and interaction between condition and strategy indicate that operator strategy and operational condition each affect dispatcher utilization but there is no evidence of an interaction effect between the two.

Post-hoc Tukey's honest significant difference tests showed that dispatcher workload during normal operations (M = 68.02%, SD = 7.07%) differed significantly (p < .001) from dispatcher workload during irregular operations (M = 69.82%, SD = 6.99%). The multiple comparison test also showed that there was no significant difference (p = 0.9253) in workload of dispatchers taking a FIFO versus STF strategies. However, there were significant differences (p < .001) in workload when comparing FIFO and Priority #1, FIFO and Priority #2, STF and Priority #1, STF and Priority #2, and Priority #1 and Priority #2 strategies. A dispatcher behaving with the Priority #1 strategy experienced the greatest workload based on utilization (M = 73.33%, SD = 7.28%), lower workload with the Priority #2 (M = 68.81%, SD = 6.81%), and the least workload with FIFO (M = 66.67%, SD = 5.87%) and STF (M = 66.88%, SD = 6.20%) strategies. The utilization means and comparison intervals of the two groups of conditions and four groups of strategies are depicted in Figure 42.

Figure 42: Mean utilization with comparison intervals (alpha = .001) for each level of operator strategy in normal (N) and irregular (O) operations

In simulated operations, the Priority #2 strategy allowed dispatchers maintain workload levels during irregular conditions just as manageable as during normal conditions. Therefore, for the airline manager considering strategies to train dispatchers on, the first-in-first-out and shortest-task-first approaches yield indistinguishable results in workload. However, dispatchers do experience a statistically noticeable difference in workload during irregular conditions, leading them to experience workload not much different than a dispatcher taking the Priority #2 approach during normal and irregular conditions. So, if the airline managers expect frequent or severe irregular operations, according to results from SHADO, it may be wise to train with the strategy of Priority #2. The Priority #1 strategy is not advisable in either condition as it appeared to always lead to significantly higher workload than any of the other tested strategies.

The senior dispatcher at Horizon Air found the results from SHADO to be surprising in that the prioritization that is often touted may lead to lower performance due to workload being above 70%. A common heuristic is that the senior dispatcher reminds new dispatchers is that they should prioritize the aircraft in the sky, then the aircraft preparing to take off, then everything else. Within the simulation, this was modeled by treating the flight following tasks for short-haul, long-haul, and focus flights as highest priority. Then, the flight planning tasks for all types of flights were at the next level of priority. And finally, miscellaneous tasks were at the lowest level of priority. The results showed that a dispatcher operating with this approach may experience high workload with an average of 72% utilization, compared to the 66% experienced with the

FIFO approach. Under contingency conditions, utilization with Priority #1 strategy would rise to 74% compared to just under 68% with FIFO.

As the senior dispatcher and researchers reviewed the results further, the senior dispatcher highlighted that the result is not so concerning. In fact, dispatchers newly on the job do take the FIFO approach and experienced dispatchers take the STF approach, however it is in concert with their prioritization heuristic as a matter of safety. If they are ever faced with an aircraft in the sky calling with an issue that is pressing but may not have arrived into their queues first or may take significantly longer than other tasks, they quickly respond and reallocate resources as needed. And, in the cases of their workload spiking beyond the optimal range, they have the chief dispatcher present to support by offloading lower priority tasks. The reality is that people dynamically switch from one strategy to the other. The results found with SHADO are to be considered in the context of this reality and these tests would be well-suited with additional data on the timing of such switches.

5.2.2 What if Horizon takes Alaska's approach to staffing flights?

Horizon Air currently operates with a generalist team of dispatchers, in which all dispatchers can plan and follow any type of flight on their desk. *Flight planning* involves dispatchers referencing up-to-date information to submit details for each flight required for pilot approval prior to takeoff. It is typically completed in one-go and takes longer for *focus* and *long-haul* flights than *short-haul* flights, whereas *flight following* tasks arise intermittently requiring a dispatcher's brief attention to monito flight conditions. Horizon Air may consider changing their approach to staffing dispatchers by specialty, in a way like how Alaska Airlines operates. At Alaska's operations center, certain flights are only allocated to certain dispatchers. To test the effect of such a change, SHADO was used to simulate workload of dispatchers for a shift at work on Horizon's current flight schedule. The three generalist desks became three specialist desks, assigned flights based on geography.

The airline's goal has been to maintain approximately 60 flights per shift per dispatcher. So, for the morning shift, the airline's usual dispatch distribution process resulted in the dispatchers working at Desks 1, 2, and 3 being randomly assigned 59, 63, and 58 flights respectively. As shown in Table 24, Desk 1 was assigned only short-haul flights to plan, Desk 2’s roster included four focus flights, and Desk 3’s included one long-haul flight.

Table 24: Number of flights to plan and follow per original desks

Dispatcher Desk	*1*	*2*	*3*
Short-Haul Planning	59	59	57
Long-Haul Planning	0	0	1
Focus Planning	0	4	0
Short-Haul Following	45	45	46
Long-Haul Following	0	0	1
Focus Following	0	2	0

Currently, dispatchers’ flights are divided up without consideration of geography. A new configuration of the three desks was designed for this experimental simulation. The researchers identified Horizon's two major landing locations as Seattle-Tacoma International Airport and airports in the state of Oregon. Therefore, the new configuration includes two desks that would be predominantly responsible for handling flights specifically landing at these two geographical hubs. The third desk *includes all other flights which land anywhere in* California and farther out to Missouri, where the single long-haul flight of the morning shift lands at Kansas City International Airport. This multiregional desk is assigned 53 flights, including the four focus flights and one long-haul flight. The Seattle desk is assigned 63 flights while the Oregon desk is assigned 62 flights. All the flights on both the Seattle and Oregon desk are short-haul.

With these new assignments, interarrival time distributions were redrawn from the modified flight schedule based on when dispatchers would be due to submit flight plans to the pilots-in-command before aircraft takeoff times. The flight schedule was modified from the documents the senior dispatcher originally shared (see Chapter 4) from five days of airline operations. The resulting flight schedule per desk is reported in Appendix N. The majority of interarrival time distributions are exponential. The focus flight planning tasks for the multiregional *desk* arrive on a lognormal distribution with mean of 165 +/- 104.5 minutes as there are few and far between.

SHADO was simulated for 365 replications of the initial settings at each desk for either the original-generalist team or geographic-specialist team. A 3 (desk) x 2 (team expertise) two-way factorial analysis of variance (ANOVA) was calculated on dispatcher utilization. Assumptions of normality and homogeneity of error variance were met in this analysis. The main research question is: did the different staffing allocations result in different utilization results?

The interaction effect of desk x team expertise was significant, $F(2, 2184) = 38.3148$, $p < .001$. So, a one-way ANOVA was run on the simple effects and no significant difference (p = .2431) was found when collectively comparing the three desks before and after the reassignments. However, a closer look in Figure 43 shows a significant difference between the workload distribution on Desk 2 and all other desks. Another one-way analysis of variance (ANOVA) was used to determine whether dispatcher utilization differed based on expertise within or between two teams of the three potential geographical specialist desks and three present-day generalist desks.

Figure 43: Multiple comparison of means (alpha = .001) of utilization per dispatcher geographic (Multiregional, Seattle, Oregon) or original desk (1, 2, 3).

The analysis showed significant differences among six groups of 365 data points each, $F(5,2184) = 40.09$, $p < .001$. A dispatcher working on the original generalist Desk #2 experienced the greatest workload based on utilization (M = 52.12%, SD = 7.22%), significantly less workload on the new Multiregional (M = 48.63%, SD = 7.88%) and Oregon (M = 48.75%, SD = 5.96%) specialist desks, lower workload on the new Seattle specialist desk (M = 46.89%, SD 5.79%) and original generalist Desk #3, and the least workload on the original generalist Desk #1 (M = 46.01%, SD = 5.72%).

Post-hoc Tukey's honest significant difference tests showed that dispatcher workload on the original Desk #2 differed significantly ($p < .001$) from each of the other desks. Workload on the new Multiregional specialist desk differed significantly ($p < .001$) from workload on Desk #1 but was not significantly different from the new Oregon ($p = .9999$) and Seattle ($p = .0045$) specialist desks or Desk #3 ($p = .0277$).Workload on the Oregon desk also differed significantly from workload on the original generalist Desk #1 ($p < .001$) but was not significantly different from the Seattle ($p = .0017$) specialist desk or generalist Desk #3 ($p = .0122$). On the Seattle desk, dispatcher workload was not significantly different from what it was on the original generalist Desks #1 ($p = .4569$) and #3 ($p = .9946$). Dispatchers on the original generalist Desk #1 experienced workload not significantly different than those on Desk #3 ($p = .1739$). Figure 44 reports the absolute differences between pairs of desks.

Figure 44: Absolute Differences of Dispatcher Utilization between Pairs of Desks

The absolute differences between estimated mean utilizations on geographical desks with specialist dispatchers was 1.73% for Multiregional and Seattle ($p = .0011$), .12% for Multiregional and Oregon ($p = .9656$), and 1.86% for Seattle and Oregon ($p < .001$). On the other hand, 6.11%, 1.14%, 4.97% were the absolute differences between estimated mean utilizations on original desks with generalist dispatchers for Desk 1 and Desk 2 ($p < .001$), Desk 1 and Desk 3 ($p = .0455$), and Desk 2 and Desk 3 ($p < .001$), respectively. So, in simulated operations, dispatchers in the original desk placement had a wider difference in workload experienced within their team than in the new geographically placed desks. Dispatchers working on the present-day Desk #2 experienced significantly higher workload than dispatchers working on any of the other two generalist desks as well as any of the three potential specialist desks.

Therefore, there may benefits found in transitioning from the current generalist desk approach to specialist desks based on geography of flights. As found using SHADO, the inter-dispatcher utilization on the specialist team does not vary as widely as it does on the generalist team. In prior work (Mekdeci & Cummings, 2009), specialist (or “mechanistic”) teams were shown to perform better than generalist (or “organic”) teams with quicker objective completion times but there is a concern that such team structures would perform worst in cases where the task load is unequal. If any new flights are added to the network in the future, it will be important to re-evaluate the task load distribution as well as functional training such that in the event of extreme irregular operations, the team working can quickly redistribute tasks to manage any spikes in dispatcher workload and reduce the opportunities for error.

The senior dispatcher found this prospective analysis to be revealing on two fronts. First, in present-day operations, differences in workload for dispatchers working similar desks during the same shift were found that were previously unrecognized. The current flight assignment process makes it challenging for the senior dispatcher to get a comparative view of how work may differ between desks. Rather, one desk-at-a-time is allocated a set of flight schedules that seems reasonable.

Second, the senior dispatcher concurred with the finding that there is no significant difference in average workload with the original generalist team versus geographic specialist team structure. Since Alaska Airlines and other companies have used the latter for many years while Horizon and yet other companies continued to use the former, signified to the senior dispatcher that there was not any "ground-breaking" support for one structure over the other. Until now. Analyzing results from SHADO, there was a significant difference found in the workload experienced by one dispatcher from present-day operations to the rest of the group of dispatchers. So, there exists an imbalance in workload wherein one dispatcher has high workload while others overall experience moderate workload. Although the current Horizon method strives to balance total number of flights around 60, the new method proposed here would result in a balance of around 4 hours and 48 minutes total busy time out of their 10-hour shift per desk. The new method leads to less disparity in the distribution of workload across the three dispatchers on duty. More importantly, this would allow all three dispatchers to maintain moderate levels of workload between 30-70% utilization.

SHADO provides the airline with a tool to simulate how busy their dispatchers may end up becoming due to possible time spent on flight planning and flight following, with thoughtful attention to the complexities of managing long-haul and focus flights embedded in the design of the simulator. However, the experience that dispatchers gain over time from working on the same regions may lead to better long-term performance which SHADO has not been designed to fully capture. And it should be emphasized that these interpretations are strictly just for these sets of dispatch operations and the model would need to be calibrated for each specific application.

5.3 Chapter Summary

Here, SHADO was applied to answer four questions across two domains. First, SHADO was used to assist a chief dispatcher of both commuter and freight railroad dispatch operations in investigating: 1) at what point in railroad operational expansion may the company need to hire additional dispatchers? 2) How dispatcher workload might be affected by the impact of automation? In airline dispatch operations, the airline company could explore: 3) how training methods may affect dispatcher workload? 4) What may be the effect of changing the way flights are assigned to dispatchers on duty?

Due to the nature of Rio Grande Pacific Company's railroad dispatch operations, prospective analyses needed to be conducted on two different sets of data. One dataset came from the freight desk and the other from the commuter desk. The two operate independently and therefore require their own input and internal parameter settings. More institutional records were accessible for data on the freight railroad operations while researchers needed to investigate beyond the United States to discover data on how automation may impact the commuter railroad dispatcher workload. So, answers to the second question are limited based on a hypothesis that technology that would ultimately be implemented at the company would uphold similar results. The railroad commuter work flow is more comparable to the airline dispatcher work flow at Horizon Air. Horizon dispatchers also have a predictable schedule that they begin each shift with and they must respond in time to operational responsibilities to keep the vehicles safely and efficiently in route. The most significant task, in terms of time, for the two different dispatchers was related to this goal: *train movement* for the commuter and *short-haul flight planning* for the passenger operations. In all three operational domains, questions about team coordination or artificially intelligent decision aids were found to be beyond the realms of possibilities and therefore not meaningful for either company to explore. As shown in this work, numerous pressing questions that do matter to the stakeholders could be investigated instead. With the open platform released from this work, RGPC, Horizon, and other transportation companies can use SHADO to reduce on upfront costs in resources of experimenting with their actual human operators in the loop and in the longer-term risks of making decisions without thoughtful consideration of human factors issues such as operator workload.

6. Conclusions

Given that advanced automation will be increasingly used in both fleets and network operations, there is a need to better understand how the insertion of such technologies will likely impact human operators from both an efficiency and a safety perspective. While simulation models have been routinely developed to explore such questions in aviation, there has been little work in extending such objective and quantitative approaches to dispatch operations in any domain. Even so, most of the operations research discounts human factors. To address this gap, this thesis outlined the development of such a model that allows stakeholders to consider the human factors in planning the future dispatch workforce and innovative transportation system designs.

Railroad and airline companies have multiple stakeholders with diverse interests who demand public safety, logistic efficiency, job security, profitability, and technological innovation. With these competing demands, there is a need for a tool to prospectively explore how changes to some variables may influence other important variables. With increasing forms of technology in transportation systems across the US, an objective method for better understanding the potential implications on the human performance is urgently needed. The model simulation was developed to support various decision makers in designing and staffing of dispatch operations in both railroad and airline settings. Building from the literature review, the core elements of the underlying model of the Simulator of Humans & Automation in Dispatch Operations were outlined in Chapter 3. In Chapter 4, the results of a multi-stage simulation verification and validation were presented. Chapter 5 presented prospective case studies using the model in simulation. This chapter concludes by highlighting the novelty, usefulness, limitations, and potential of SHADO for future work.

6.1 Novelty

This research began with three questions:

1. How can we develop a workload model of dispatchers managing fleets in railroad and airline operations centers so that stakeholders can explore future concepts of operations?
2. What parameters in the model are most influential to dispatcher workload in these dispatch operations centers?
3. What are the limitations of the model and how generalizable is it for fleets in railroad, airline, and future transportation systems?

The first question was answered by studying the literature on dispatcher workload and dispatch operations modeling in the fields of human factors engineering, operations research, aerospace, and robotics. Key elements of dispatch operations centers that may influence dispatcher workload were identified. Utilization, the percent of busy time over total time allotted, was defined as the metric for workload. Through real-world observations and interviews with subject matter experts (SMEs) in dispatch operations centers, timing of tasks—when tasks occur, how long it takes dispatchers to complete tasks, and the potential risks that arise when tasks are not completed on time—was found to be critical to human-systems performance. Therefore, discrete event simulation, a queuing-based modeling method, was employed to model dispatcher workload in numerous operational conditions.

The discrete event simulation (DES) modeling approach was supported by studies on human supervisory control of robots (Gao, Cummings, & Solovey, 2014) and unmanned vehicle systems (Nehme, 2009). Yet, it had not been reported before with a focus on dispatch operations centers of civil transportation systems where dispatchers supervise networked fleets. Using DES here was advantageous in that teams of human supervisory controllers (in this case, dispatchers) could be represented coordinating and managing fleets of vehicles (which can also share information between vehicles) at the task-level, which was the focus of this effort. Human operators were modeled as serial processors of tasks and the DES maintained a record of key performance metrics like dispatcher utilization and error over a number of replications which represented different days of the same shift settings.

The 10 key elements of dispatch operations centers stemmed from task load factors of the fleets and environment, strategic system design variables for dispatcher decision support, tactical staffing variables determining task allocation, and operator behavior. The task load factors included (1) fleet size, (2) fleet heterogeneity, (3) fleet autonomy, and (4) exogenous events (e.g. train derailment or emergency landing) which could affect the timing requirements of tasks dispatchers need to address. System design variables are those decided in long-term planning by strategic leaders regarding operations center infrastructure for (5) artificial intelligence decision aids and (6) team coordination. Tactical staffing variables are those decided in mid-term planning by tactical leaders regarding (7) shift schedules, (8) team expertise, and (9) team size. And, (10) attention allocation strategy is the final level of decision-making per dispatcher which would ultimately affect the order that he or she handles tasks in their queue.

Each type of task is represented by event interarrival and service time parameters. Additionally, the human error probability—chance of dispatcher making a mistake—can be defined for each task. These characteristics describe how a task is designed and the input parameters can be varied if tasks are redesigned. As discussed in Chapter 4, these input and internal parameters of how dispatch operations are designed and staffed can have different effects on human workload and performance.

To answer the second research question, model sensitivity to changes to input and internal parameters was analyzed in 114 experiments across railroad and airline domains, freight and commuter operations, 8-hour and 10-hour shifts, and a multitude of other operational conditions. The human error performance indicator was found to be more sensitive than dispatcher utilization when input parameters were deviated from original values, particularly when there was a -75% deviation in task interarrival times. This result was found across all the domains of operations. And, any deviations in task service times were generally found to cause smaller responses in dispatcher workload and performance than changes resulting from deviations in interarrival time.

When internal parameters were changed, the sensitivity analysis showed that the fleet size factor resulted in the most significant difference in dispatcher workload and performance. This can be visualized in the following Pareto charts (Figure 45 - Figure 50) of the effect of internal design and staffing parameters on dispatch operations across multiple domains. When fleet size was doubled, dispatcher utilization increased by 73% on the freight railroads' desk, increased by 65% on the commuter railroad desk, and increased by 48% on the airline desk. Moreover, the number of dispatcher errors, in terms of missing tasks or not completing tasks in time, approximately tripled across all three operational domains when fleet size was doubled. When fleet size was tripled, human error was found to increase 20-, 10-, and 6-fold on each respective desk.

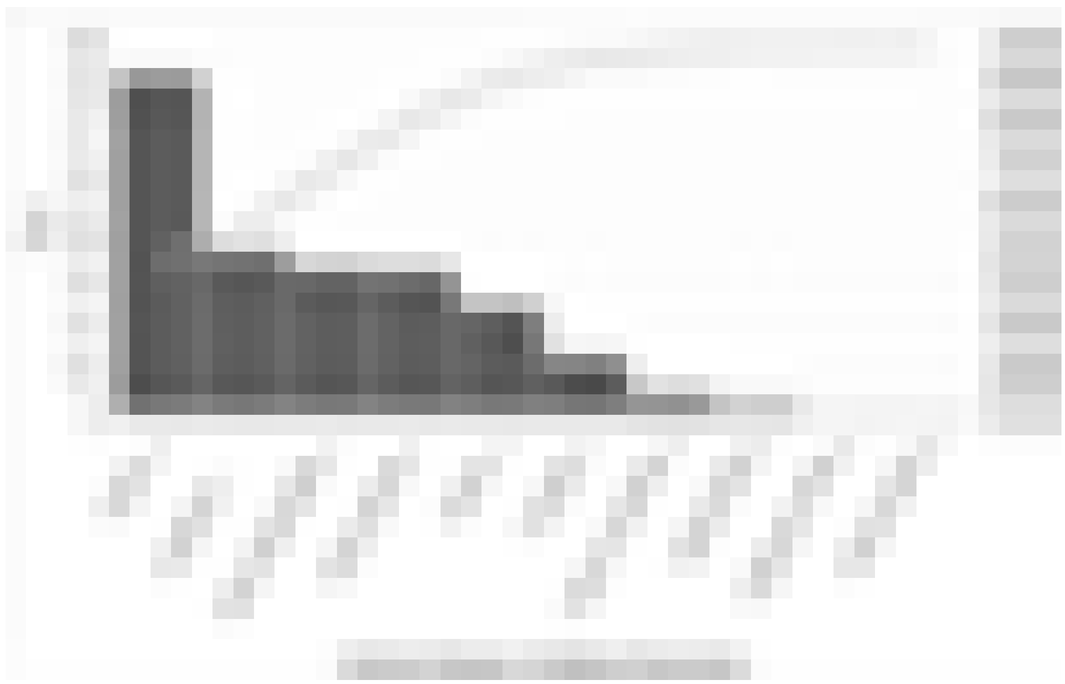


Figure 45: Pareto chart of effect of internal parameters on freight railroad dispatcher workload

Figure 46: Pareto chart of effect of internal parameters on freight railroad dispatcher error

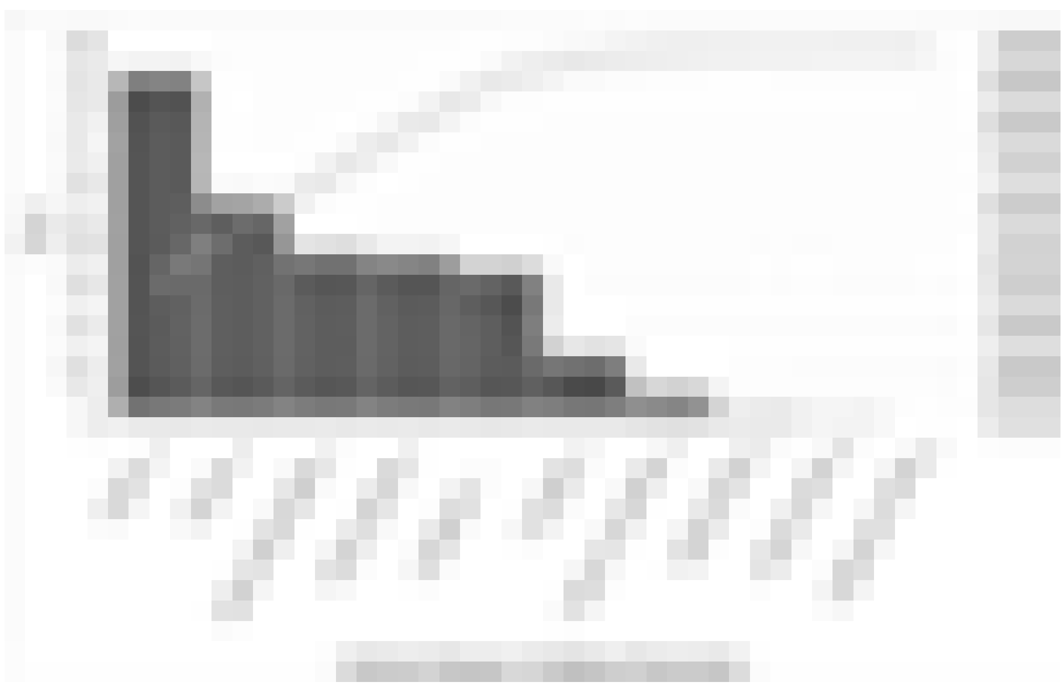


Figure 47: Pareto chart of effect of internal parameters on commuter railroad dispatcher workload

Figure 48: Pareto chart of effect of internal parameters on commuter railroad dispatcher error

Figure 49: Pareto chart of effect of internal parameters on airline dispatcher workload

Figure 50: Pareto chart of effect of internal parameters on airline dispatcher error

So, the top contributor to dispatcher workload and error was found to be the number of railroad operations or number of flight operations a dispatcher has at their desk during their shift. The model was found to be generalizable in other areas too. The next leading contributors to dispatcher performance across domains consistently included fleet heterogeneity, fleet autonomy, team size, team coordination, and AIDA. Team expertise made a significant difference in commuter railroad and airline dispatch operations but not in freight railroad dispatch operations. On the other hand, operator strategy and exogenous factors as modeled were not found to have any significant impact on dispatcher workload.

In the process of answering the first two questions, an answer to the third question arose. The resulting model was validated internally with industry software, externally with real-world operational data from railroad and airline dispatch operations, and with the sensitivity analyses, was found to be generalizable across the multiple operational domains. Consulting with partners in the transportation industry, a platform was designed for running the model as a discrete event simulation online. As will be discussed in the next section, this platform, the Simulator of Humans & Automation in Dispatch Operations (SHADO), was found useful by stakeholders to answer what-if questions about their network operations management.

6.2 Usefulness

SHADO is a tool that allows stakeholders to rapidly prototype numerous scenarios. This tool gives users immense control to design their operations centers to meet their specifications with more than 10^18 possible combinations of input parameters. SHADO can simulate historical, present-day and future concepts of operations. The ability to model human performance with results reported on dispatcher workload and error over up to 10,000 days with realistically random distributions each day is another novel contribution to the railroad and airline industries.

Because of the model's novelty, it was important to make sure that the code functioned as conceived. Did the model take in input parameters and produce expected outputs? Did the model respond as expected to adjustments to the internal parameters? Once internal confidence was built in how SHADO works, the next goal was to build external confidence, i.e., increasing the confidence in railroad and airline stakeholders using the tool to model real dispatch operations. Did the model get results close to experts' experience in the real-world under the same initial settings? Did the model behave realistically when the initial settings were adjusted positively or negatively?

At two real-world companies, Rio Grande Pacific Company and Horizon Air, SHADO was used to test what-if scenarios and support various stakeholders in understanding *when* and *why* dispatchers may be over- or under-utilized during specific shifts. SHADO was found useful in investigating *how* a dispatcher's workload may change with changes in dispatch operations. In the first company, SHADO was used to answer the question: what if RGPC increases the size of short-line freight rail operations? How might this affect dispatcher workload and additional staffing requirements? Results from this prospective analysis with SHADO showed the chief dispatcher at RGPC that the dispatcher working the overnight desk could theoretically manage two to four times as many railroads as their coworkers who work during the other times of the day. With the suggestion that adding operations for this specific setting would increase the dispatchers' workload and that additional dispatchers would likely be needed for the morning and afternoon shifts when the operational size grew to 18 railroads, the chief dispatcher could better plan for future business growth with human factors in mind.

Other questions surrounding the potential impact of automation on the local commuter railroad dispatcher desk, and at Horizon Air, regarding the potential impacts of new training or staffing regimes, could all be explored by tuning the input and internal parameters of the model in SHADO. Performing prospective analyses across the multiple domains built further credibility in how generalizable the model could be to different dispatch operations. The work flow of commuter railroad dispatchers and airline dispatchers was found to be even more comparable than commuter and freight dispatcher work flows in the same organization. Yet, SHADO could be used to study current and future operations for the diverse set of possibilities.

Throughout the model development process, SMEs were asked to review the inputs, the internal model and the outputs. These SMEs included the chief dispatchers, senior dispatchers, and other dispatchers at work. Researchers walked through the steps of setting shift conditions, task characteristics, fleets, and dispatcher roles in SHADO, and then explored the results that SHADO produces. The open-box validation process resulted in an open software platform online http://apps.hal.pratt.duke.edu/shado-webdev.

Any stakeholder can access the platform, which has undergone multiple iterations of usability testing. The platform allows users to input custom settings, run multiple scenarios, interact with dynamic data visualizations, save decisions, and download human-system performance results. Stakeholders from RGPC and Horizon Air approved the user-friendly design, the usefulness of the underlying computational tool and the ability it gives them to test out ideas that would otherwise be too expensive and time-intensive to try in the real-world.

6.3 Limitations

As with any model, there are several limitations of SHADO. First, use of the model requires that the user have some representation of the underlying distributions of the task inter-arrival times and operator service times. While the researchers were able to obtain relatively accurate numbers for validation tasks, these would need to be updated for every new application of SHADO.

Second, as described in Section 3.2, the model sourced human error probability parameters from locomotive crew estimates and not from dispatcher estimates. Unfortunately, little data was found on dispatcher error rates so there needs to be significantly more research in this area.

Third, the model does not account for some characteristics of the human operators that could impact performance such as the hours of sleep prior to the shift. As presented in Appendix L, a more complex model that considers other factors like recent operator work history and shift time of day, along with shift duration, may provide better predictions for operator workload during extremely long shift schedules. However, it is not clear whether the inclusion of such variables would improve model fidelity, so this represents another area of future research.

The results of the prospective analyses are limited in that our input data assumes that the relative effect of automation technology found in Europe are applicable to the United States and that RGPC's new railroads will replicate the influence of their current railroads. The prospective analyses can be improved with additional resources to gather more precise data on the nature of positive train control in dispatching and on the nature of calls, bulletins, and other sources of dispatcher workload different types of railroads.

Despite model limitations, as has been demonstrated here, SHADO can approximate workload levels for dispatchers with different operational responsibilities and schedules. The internal, external, and face validation with industry standard software, empirically collected data and subject matter expert interviews provide confidence in SHADO's representation of the real-world dispatch operations system. The simulation results were found to be consistent with dispatcher workload trends as experienced by those who have worked directly in rail dispatch operations.

6.4 Potential

Along with developing a generalizable dispatcher workload model and identifying parameters most influential to dispatcher workload, other important contributions were made here which can be used in future work in the field. While validating the model in the airline domain, a tool was developed to gather dispatcher-generated workload data, the *Dispatcher's Rough Assessment of Workload-Over Usual Times* (DRAW-OUT). The airline company did not already have a database of workload recordings that could be shared, and the nature of the airline dispatcher's work organization was such that researchers decided to avoid intrusive methods. And, for validating a discrete event simulation, researchers found more qualitative methods that existed to be too far removed from the level of contextual data required. DRAW-OUT allows researchers to capture work-specific information and workload experienced in the metric of utilization. The presentation was also found to be designed highly usable amongst dispatchers. Dispatchers may complete the form over the course of their shifts or all at once. The format also makes it possible for researchers to collect data without being physically present during the process. Finally, the DRAW-OUT tool was found useful to managers to quickly recognize discrepancies in how dispatchers under the same conditions experienced similar or different workload over time.

Another important contribution from this research is the database of task and time information gathered from multiple operational domains. This database was processed to generate probability density functions of distributions of interarrival and service times of tasks. The database includes approximations of likelihood of mistakes (human error probability) which was developed by connecting relevant data from a prior human reliability assessment in the railroad industry with original cognitive task analyses conducted during this thesis research across the freight and commuter railroad, and airline dispatch operations.

The Simulator of Humans & Automation in Dispatch Operations (SHADO) provides stakeholders with a tool to rapidly explore concepts, the first step in any systems engineering development process. This process can be followed by not just academic or government researchers but also analysts within companies to similarly model their unique operations using our customizable SHADO platform. Altogether, a foundational set of parameters that are generalizable to define the work flow structure of railroad and airline dispatch operations today is provided, along with potential future scenarios including advanced technologies integrated into operations.

One of the goals for building SHADO was to provide a predictive platform to help planners investigate how changes in operations may affect human-system performance. Thus, SHADO was used to explore four future scenarios with the short-line freight railroad and local commuter railroad and airline dispatcher desks, and so results reported here are limited to these applications. Simulating real-world data with the underlying model revealed that out of the ten potential variables that could affect dispatch operations, the variable that matters most to dispatcher workload and performance is fleet size. More work is required to identify differences in input parameters for new transportation systems and how these findings could be generalized beyond railroad and airline operations. But the potential of SHADO is great and much of this work has proven methods of data collection, model design, and system testing that introduce innovative opportunities for research that have not been explored for these operational domains in the United States. Moreover, SHADO can be applied to other modes of transportation that similarly rely on dispatchers as railroads and airlines have for over a century.

Appendix A: Human Error Probabilities

This appendix presents the human error probabilities associated with each dispatcher task.

Table 25: Human Error Probabilities per Freight Dispatcher Task Type

Freight Dispatcher Task Types	Generic Task Type Description	Triangular Probability Distribution		
		Minimum	Mode	Maximum
Actuation (OK) **Actuation (Clear)**	Completely familiar, well designed, highly practiced task which is routine	.008%	.04%	.7%
Daily Operating Bulletin	Restore or shift a system to original or new state, following procedures with some checking	.08%	.3%	.7%
Temporary Bulletin Issue	Identification of situation requiring interpretation of alarm/indication patterns;	2%	7%	17%
Temporary Bulletin Void & Verify	Restore or shift a system to original or new state, following procedures with some checking	.08%	.3%	.7%
Other Communications	Simple response to a dedicated alarm and execution of actions covered in procedures	.008%	.04%	.7%
Weather Recording	Fairly simple task performed rapidly or given insufficient or inadequate attention	6%	9%	13%
Notetaking	Skill-based tasks when there is some opportunity for confusion	.2%	.3%	.4%
Reporting	Fairly simple task performed rapidly or given insufficient or inadequate attention	6%	9%	13%
Miscellaneous				
Transfer-of-Duty	Skill-based tasks when there is some opportunity for confusion	.2%	.3%	.4%

Table 26: Human Error Probabilities per Commuter Dispatcher Task Type

Commuter Dispatcher Task Types	Generic Task Type Description	Triangular Probability Distribution		
		Minimum	Mode	Maximum
Train Movement	Completely familiar, well designed, highly practiced task which is routine	.008%	.04%	.7%
Bulletins	Restore or shift a system to original or new state, following procedures with some checking	.08%	.3%	.7%
Temporary Bulletin Issue	Identification of situation requiring interpretation of alarm/indication patterns;	2%	7%	17%
Other Communications	Simple response to a dedicated alarm and execution of actions covered in procedures	.008%	.04%	.7%
Weather Recording	Fairly simple task performed rapidly or given insufficient or inadequate attention	6%	9%	13%
Reporting	Skill-based tasks when there is some opportunity for confusion	.2%	.3%	.4%
Bulletin Printing	Fairly simple task performed rapidly or given insufficient or inadequate attention	6%	9%	13%
Miscellaneous				
Transfer-of-Duty	Skill-based tasks when there is some opportunity for confusion	.2%	.3%	.4%

Table 27: Human Error Probabilities per Airline Dispatcher Task Type

Airline Dispatcher Task Types	Generic Task Type Description	Triangular Probability Distribution		
		Minimum	Mode	Maximum
Short-Haul Flight Planning	Completely familiar, well designed, highly practiced task which is routine.	.008%	.04%	.7%
Long-Haul Flight Planning				
Focus Flight Planning				
Short-Haul Flight Following				
Long-Haul Flight Following				
Focus Flight Following				
Miscellaneous				
Transfer-of-Duty	Skill-based tasks when there is some opportunity for confusion	.2%	.3%	.4%

Appendix B: Simulator of Humans & Automation in Dispatch Operations

This appendix presents screenshots of the SHADO online platform.

Figure 51: Example of shift settings page

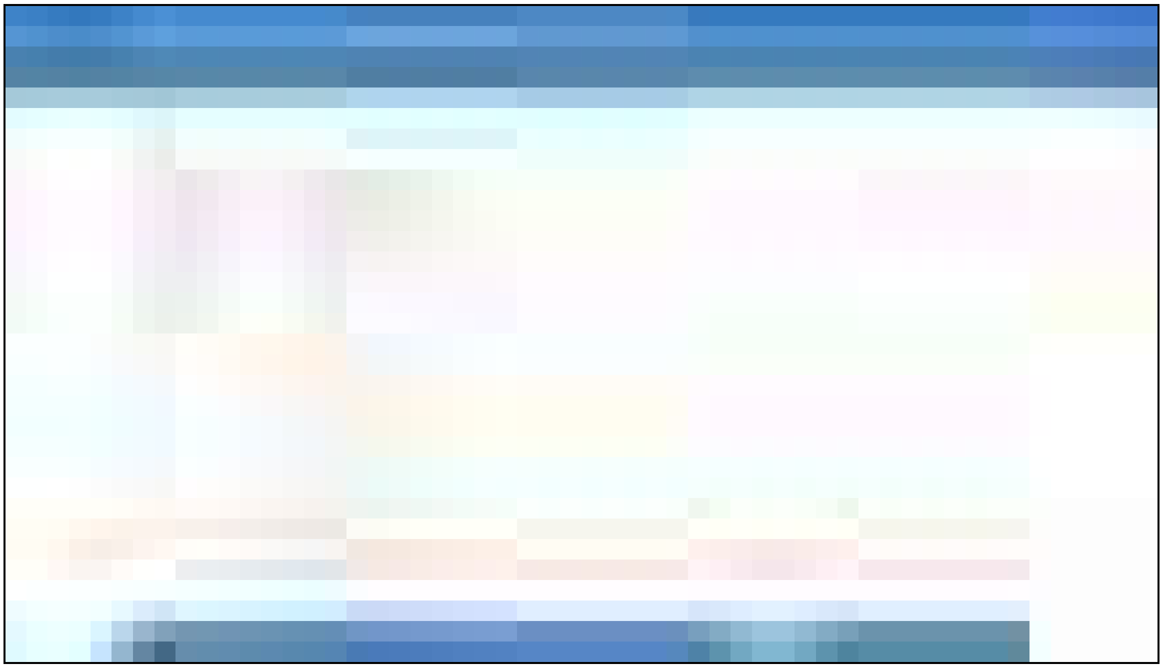


Figure 52: Example of first section of Tasks' settings page for a task

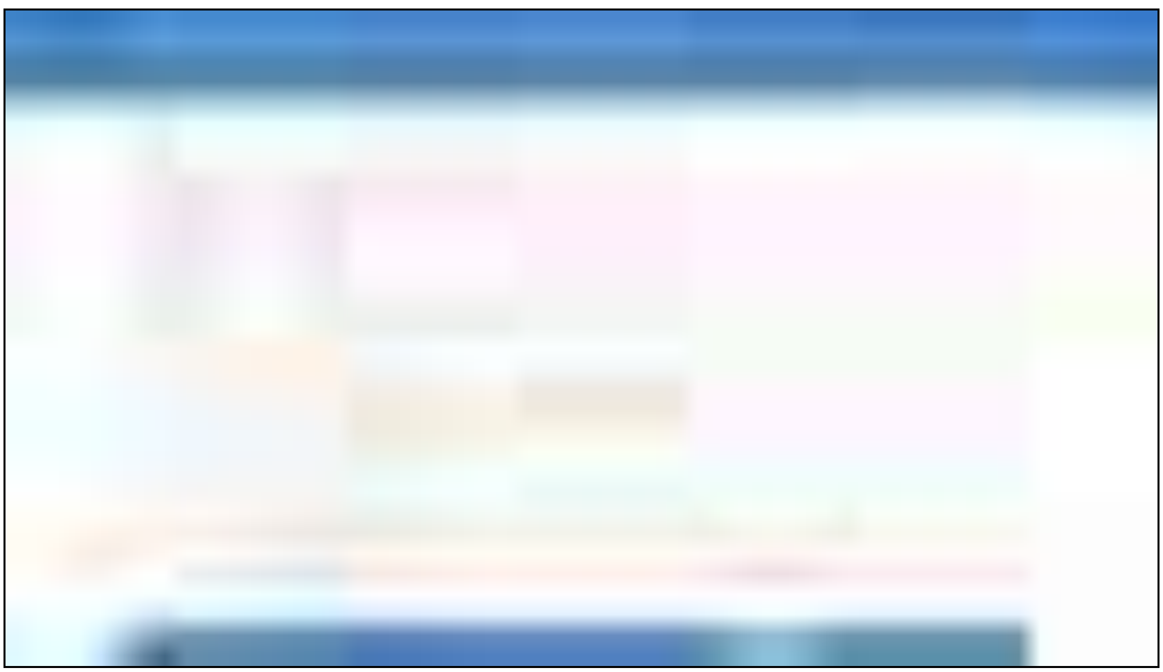


Figure 53: Example of final section of Tasks settings page of a task

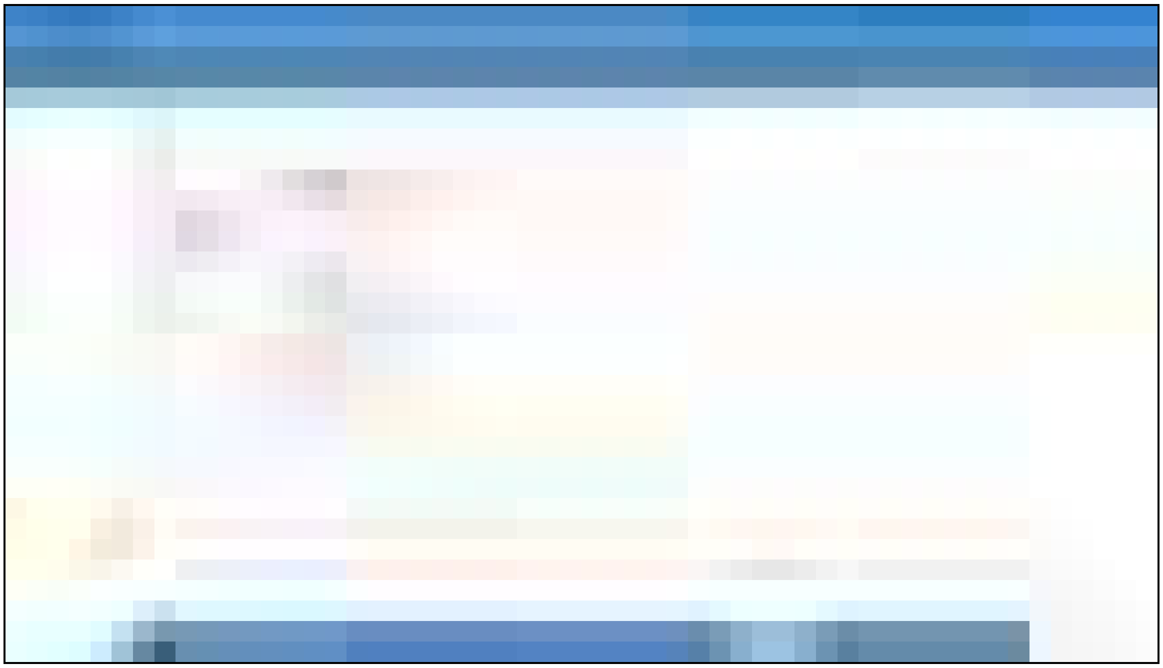


Figure 54: Example of railroads' settings page for tasks from Other Sources

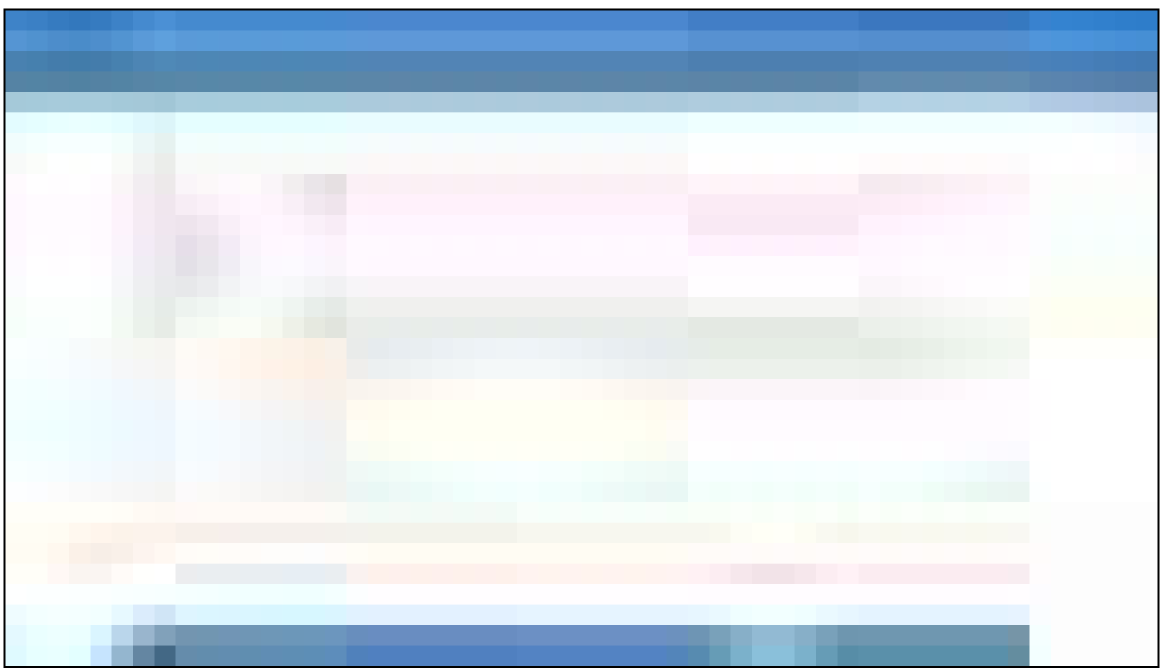


Figure 55: Example of dispatchers' settings page

Figure 56: Example of high-level results page of workload and time per task type

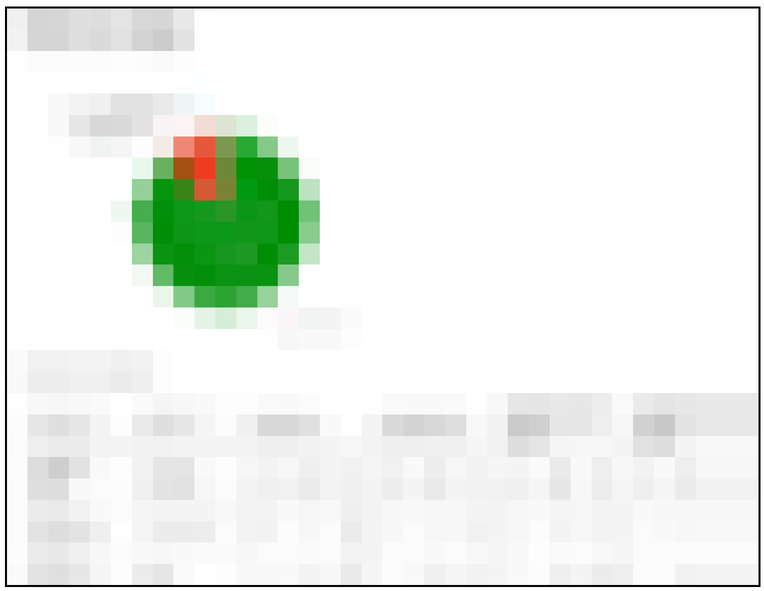


Figure 57: Example of high-level results page of failed tasks

Appendix C: Additional Data Gathered for Cognitive Task Analyses in Dispatch Operations

On Day 1 of the January 2017 visit to the Class I railroad, the chief dispatcher condensed what is normally 10 weeks of classes and 18 weeks of on-the-job experience into six hours of training. Day 2 began with one hour of training on the company's dispatch scheduling interface and symbol system. Then, two first-shift dispatchers were each observed for three hours in the morning and early afternoon. Another three hours of observation was repeated on Day 3 with a final set of two dispatchers. Following these four observations, the researchers debriefed with the chief dispatcher to review recordings, verify data gathered.

During the visit to the global commercial airline company in April 2018, two researchers conducted field observations at the airline for over three days to determine airline dispatcher workflow and tasks. On the first day, the researchers interviewed people in the following roles to gain a thorough understanding of the airline operations: dispatch instructor, information technology (IT) systems manager, IT dispatch software developer, and fleet manager. The dispatch instructor was a career dispatcher and gave the researchers a "crash course" into airline dispatcher functional requirements. The IT systems manager gave the researchers a tour of the operations center and introduced the departments that surround and interface with the dispatchers' work. The researchers completed the day by observing two dispatchers of domestic flights for two hours each during their evening shift.

On the second day, the researchers observed the daily morning briefing call which is hosted in the airline operations command room and includes operational leaders calling in remotely from the entire airline network. There, managers and representatives from each department and station update the attendees on the status of the airline and any anticipated network conditions that may affect the safety and efficiency of flights for the day. A special assignment supervisor for domestic flight control explained the role of each person speaking and later shared how he assigns flights and distributes workload for each dispatcher desk. Two researchers respectively observed two dispatchers managing international flights for two hours during their morning shift. Again, during their morning shift, on the third day, one researcher observed two additional dispatchers at work on domestic flights for two hours each.

To proceed with further development of SHADO, the real-world observations discussed in Section 4.2 were instrumental for validation of the conceptual model presented in Chapter 3. In Chapter 2, previously published models from domains relevant to remote operations centers were demonstrated. This prior work informed the design and interaction of ten key internal parameters of SHADO: fleet size, fleet heterogeneity, fleet autonomy, environment, team coordination, artificially intelligent decision aids, shift schedule, team expertise, team size, and operator attention allocation.

In November 2016, researchers convened meeting at Duke University with two senior leaders from a Class I railroad. The leaders were responsible for automation (namely, positive train control) technology, safety and human-systems performance in dispatch operations of the large railroad. This was the same railroad that researchers later visited, as reported in Section 4.2. The purpose of this initial meeting was for researchers to deliver a presentation and gather feedback from the leaders on the preliminary conceptual model design of railroad dispatch operations' human-automation factors that affect dispatcher workload. Similarly, structured sessions were organized several times over the course of two years for operational leaders at RGPC and again at an Airline Dispatcher's Federation (ADF) workshop hosted by NASA Ames Research Center where nearly every commercial airline in the United States was represented, including regional airlines such as Horizon Air.

As shown in Table 28, there was a purpose to each question. For example, researchers asked about how the leaders' operational centers were organized. This question led to answers that validated the concept of teams in SHADO. Researchers identified a commonality in this case across the railroads and the airlines: dispatchers do not normally work in teams which require interpersonal communication. Rather, each dispatcher operates independently on their tasks at hand. Another question, about what work people do, revealed that dispatcher-to-dispatcher communication does happen but typically during the shift turnover period, wherein, the dispatcher that is going off duty briefs the dispatcher that is coming on duty at the same desk.

This set of questions was foundational to validate each submodule of SHADO's conceptual model. The leaders even provided further details that supported later steps in the validation process of SHADO. These will be discussed in the next sections.

Table 28: Questions Posed to Subject Matter Experts at Railroads and Airlines

Question	Purpose
How is the ops center organized (for example, do the people work in teams? How many people per shift?)?	To validate concept of teams
What work do people do (for example, answering calls or completing paperwork)?	To validate concept of tasks
Where does the work come from and go to in the network (for example, a pilot calls in for help and the dispatcher calls another to share information)?	To validate concept of fleets
How often do those tasks need to be addressed in a shift?	To validate concept of task arrival process
How long does it take to complete each task?	To validate concept of operator service process
How do off-nominal situations (for example, unexpected severe weather) affect the workload of the ops and delays in the network?	To validate concept of exogenous events
How do you measure system performance in your operations?	To validate concept of utilization

The researchers found that dispatchers sit in pods with other dispatchers, but each dispatcher is solely responsible for their fleet. When a dispatcher needs a break or is overloaded, the senior dispatcher helps to redistribute their work for a limited time. Railroad dispatchers manage 1-12 railroads per 8-hour shift. Airline dispatchers manage 30-60 flights per 10-hour shift.

Dispatchers are responsible for preparing and tracking each trip from the fleet they manage. Dispatchers begin their shifts with debrief communications from the dispatchers that are going off duty. As they begin to situate themselves for the shift ahead, the fleets and others in the network may demand their attention. Railroad dispatchers set and release bulletins for who is using what track when. Airline dispatchers plan flights and follow each flight. Dispatchers work to ensure safe and efficient fleet operations.

From the first visit to RGPC, dispatcher tasks were identified and initial estimates of input parameters for each task in the dispatch center were gathered. The task types (Table 5 and Table 6) of the two dispatcher desks were validated. Dispatchers on the commuter desk were found to perform functions differently. For example, *actuation OK* and *clear* tasks performed via phone communications and paperwork on the freight desk can be summarized as *train movement* tasks which are performed by clicking on a computer in anticipation of a more predictable schedule.

The arrival times of *actuation*, *train movement*, and *other communication* tasks were identified from the records and time distributions were generated. These are reported in the tables. Appendix D includes a copy of a track warrant form where the arrival times of *actuation OK* and *clear* tasks on the dispatcher desk were gathered. Distributions generated for other tasks found from observations and data mining were validated over the course of several follow up calls and visits with RGPC.

Miscellaneous tasks were identified as another source of task load. From conversations, we realized the importance of including time spent going to the restroom or to get fresh air as people in this work environment do not have established break times. Including these led to more realistic simulations of how dispatchers use their time.

Service times for *miscellaneous* and other tasks were estimated from multiple days of observations in January, March, and May 2018 and interviews with dispatchers from each shift to get a range of possibilities from the spectrum of experiences at each desk. The chief dispatcher was then interviewed to validate the final distributions. The *daily* *operating bulletin* task was found to only be performed during the PM shift on the freight desk whereas there were at least two *bulletin* tasks per shift on the commuter desk. The transfer-of-duty was estimated to take anywhere from 5 to 15 minutes on the freight desk and last for an average of 5 minutes on the commuter desk, during the beginning and ending of each shift. Dispatchers on the commuter desk were estimated to spend more time on *miscellaneous* tasks and this can be expected from the nature of their work being more predictable with consistently scheduled train movements than the freight desk.

At Horizon Air, the dispatchers also have a phone dock to receive and make calls between airport station crew, pilots, maintenance, and other personnel within and beyond the airline. When *planning* for a single flight, a dispatcher must check several conditions before releasing the plan to pilots. The dispatcher has access to information about weather, aircraft equipment, airports, passenger and cargo payload, fuel, and flight paths available through air traffic control (ATC). Dispatchers needs to consider compliance with regulations, safety of the flight, and efficiency of the airline. If any of these are disregarded, the company is at risk of Federal Aviation Administration (FAA) fines, the pilot and passengers are at risk of flying in unsafe conditions, and the dispatcher themselves at risk at penalties on the job for wasting fuel, causing poor customer service due to turbulence or delays, or leading to the formerly listed risks.

Flight plans are due two hours prior to scheduled flight departure times for *short-* and *long-haul* flights while *focus* flights require flight plans at least three hours prior to departure. *Flight following* is difficult to capture in observations as it largely involves monitoring and is often interrupted by other tasks with hard deadlines, like *flight planning*. The purpose of *flight following* to maintain situational awareness for all flights on their desk. In fact, *flight following* is more of a priority than flight planning since it requires rapid response from dispatchers with relevant network information to any aircraft flying at any time.

Finally, *emergency management* tasks are not as common as *flight planning* and *flight following* tasks but are a part of the job. Dispatchers must remotely support flight crews in dealing with emergencies. For example, if there is a sick passenger, the crew makes a call to the dispatcher who then makes calls to their chief dispatcher, the medical crews, and personnel at alternative landing airports.

Appendix D: Blank Track Warrant Form

Figure 58: Copy of form dispatcher uses to record times of actuation OK and clear tasks

Appendix E: Data Gathered from Horizon Air

Figure 59: Dispatcher's Rough Assessment of Workload, Over Usual Times (DRAW-OUT) blank tool

Figure 60: Copies of DRAW-OUT forms from dispatchers on usual workload experienced during various shifts/desks

Figure 61: Copies of Dispatcher-Generated Service Time Estimates for Tasks in various conditions

Appendix F: Verification with Arena

We used the following input settings in both SHADO and Arena. Our goal was to ensure that SHADO mathematically computes results not significantly different from Arena across eight different variables with three to four different levels each. For these verification tests, we changed one-factor-at-a-time to ensure that we controlled for complex interactions and we used the settings in Table 29 for our base case.

Table 29: Initial Input Parameter Settings for Verification Tests

Parameter	**Initial Value**		
Shift	8 hours		
Traffic	High		
Days	500		
Desks	1		
Fleets	1		
Error Catching Chance	50%		
Strategy	FIFO		
Expertise	All tasks		
Tasks Generated	All		
Tasks	Task 1	Task 2	Task 3
Arrival Distribution (minutes)	EXPO(60/240)	EXPO(60/11.4)	EXPO(60/27)
Service Distribution (minutes)	UNIF(1,3)	UNIF(2,3)	UNIF(.75, 1.25)
Human Error Probability	0.04%		

We tested the desk size parameter by adjusting the number of dispatchers from 1 to 2 to 3. We tested the fleet size parameter by adjusting the number of vehicles from 1 to 2 to 3. We tested the shift schedule parameter by simulation 2-hour, 4-hour, 6-hour, and 8-hour shifts.

We tested the desk expertise by running two simulations each with 2 dispatchers. The first simulation with a homogeneous desk where both dispatchers handled tasks from the same queue and the second with a heterogeneous desk where the two dispatchers each had separate responsibilities of tasks lined up in their queues.

We tested fleet heterogeneity by running four simulations of 6 vehicles each. The first simulation had 6 homogenous vehicles, all transmitting the same types of tasks to dispatchers. The second simulations had two fleets, one with three vehicles of Task 1 and the other fleet with three vehicles generating Task 2. Our third simulation also had two fleets, the first with four vehicles generating Task 1 and the second with two vehicles generating Task 2. Our last simulation her modeled three fleets, each with two vehicles generating Task 1, Task 2, and Task 3, respectively.

We tested fleet autonomy by running three simulations. The first simulation was with the control variables (no level of vehicle-to-vehicle communications). The second was with partial communications and the third with high level of communications. Partial signifies that 30% less tasks are generated from this fleet while high signifies that 70% less tasks are generated.

We tested the dispatcher strategy parameter by testing a FIFO strategy, priority strategy, and shortest task first strategy in three different simulations. We tested the extreme conditions parameter by changing the parameter from none to Type 1 (e.g. train derailment) to Type 2 (e.g. poor weather) to both. Type 1 generated a new task that lasted longer than other tasks. Type 2 extended the times on all tasks.

Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, and Table 37 present results of the verification tests. The average utilization is computed over 500 replications. The max average represents the results of the one replication with the highest average utilization across all the time intervals within that shift. The max represents the maximum utilization value for any one interval across all replications.

Table 30: Verification of Dispatcher Type (desk) Size Parameter

NUMBER OF DISPATCHERS	UTILIZATION STATISTICS					
1	Software	average	max average	min average	max	min
	SHADO	26.42%	36.69%	14.41%	100.00%	0.00%
	Arena	24.68%	40.26%	14.60%	100.00%	0.00%
2	Software	average	max average	min average	max	min
	SHADO	12.44%	19.25%	7.66%	80.85%	0.00%
	Arena	13.22%	18.56%	7.27%	81.64%	0.00%
3	Software	average	max average	min average	max	min
	SHADO	8.33%	12.64%	4.43%	58.44%	0.00%
	Arena	8.81%	12.75%	4.84%	63.24%	0.00%

Table 31: Verification of Fleet Size Parameter

NUMBER OF VEHICLES	UTILIZATION STATISTICS					
1	Software	average	max average	min average	max	min
	SHADO	24.51%	35.27%	14.73%	100%	0.00%
	Arena	26.42%	40.26%	14.60%	100%	0.00%
2	Software	average	max average	min average	max	min
	SHADO	47.31%	61.33%	37.02%	100.00%	0.00%
	Arena	53.28%	70.79%	39.48%	100.00%	0.00%
3	Software	average	max average	min average	max	min
	SHADO	67.10%	93.60%	56.04%	100.00%	0.00%
	Arena	78.73%	96.83%	59.05%	100.00%	0.00%

Table 32: Verification of Shift Schedule Parameter

HOURS ON DUTY	UTILIZATION STATISTICS					
2	Software	average	max average	min average	max	min
	SHADO	25.90%	49.50%	5.74%	100%	0.00%
	Arena	25.87%	49.54%	9.69%	100%	0.00%
4	Software	average	max average	min average	max	min
	SHADO	25.84%	44.24%	7.47%	100%	0.00%
	Arena	26.25%	46.53%	14.38%	100%	0%
6	Software	average	max average	min average	max	min
	SHADO	26.50%	39.03%	16.48%	100%	0.00%
	Arena	26.35%	39.11%	15.76%	100%	0.00%
8	Software	average	max average	min average	max	min
	SHADO	26.45%	35.96%	17.40%	100.00%	0.00%
	Arena	26.42%	40.26%	14.60%	100.00%	0.00%

Table 33: Verification of Dispatcher Type Parameter

TYPES OF DISPATCHERS	UTILIZATION STATISTICS					
SIMULATION 1, TYPE 1 OF 1 DISPATCHER #1 OF 2	Software	average	max average	min average	max	min
	SHADO	12.50%	18.53%	7.55%	86.80%	0.00%
	Arena	13.30%	19.04%	8.23%	84.21%	0.00%
SIMULATION 1, TYPE 1 OF 1 DISPATCHER #2 OF 2	Software	average	max average	min average	max	min
	SHADO	12.50%	18.53%	7.55%	86.80%	0.00%
	Arena	13.30%	19.04%	8.23%	84.21%	0.00%
SIMULATION 2, TYPE 1 OF 2 DISPATCHER #1 OF 1	Software	average	max average	min average	max	min
	SHADO	23.13%	33.27%	12.99%	100.00%	0.00%
	Arena	22.94%	34.52%	11.03%	100.00%	0.00%
SIMULATION 2, TYPE 2 OF 2 DISPATCHER #1 OF 1	Software	average	max average	min average	max	min
	SHADO	3.81%	6.93%	1.53%	45.03%	0.00%
	Arena	3.73%	6.07%	1.54%	45.35%	0.00%

Table 34: Verification of Fleet Heterogeneity Parameter

FLEET HETEROGENEITY	UTILIZATION STATISTICS					
FLEET 1 OF 1: 6 VEHICLES WITH TASK 1	Software	average	max average	min average	max	min
	SHADO	2.51%	5.40%	0.07%	42.42%	0.00%
	Arena	2.47%	5.14%	0.05%	50%	0.00%
FLEET 1 OF 2: 3 VEHICLES WITH TASK 1, FLEET 2 OF 2: 3 VEHICLES WITH TASK 2	Software	average	max average	min average	max	min
	SHADO	6.89%	10.52%	4.59%	53.15%	0.00%
	Arena	6.80%	10.28%	3.89%	45.46%	0.00%
FLEET 1 OF 2: 4 VEHICLES WITH TASK 1, FLEET 2 OF 2: 2 VEHICLES WITH TASK 2	Software	average	max average	min average	max	min
	SHADO	5.47%	8.43%	3.16%	49.55%	0.00%
	Arena	5.36%	8.49%	2.78%	45.38%	0.00%
FLEET 1 OF 3: 2 VEHICLES WITH TASK 1, FLEET 2 OF 3: 2 VEHICLES WITH TASK 2, FLEET 3 OF 3: 2 VEHICLES WITH TASK 3	Software	average	max average	min average	max	min
	SHADO	27.11%	35.21%	18.01%	99.93%	0.00%
	Arena	26.60%	35.38%	18.48%	100%	0.00%

Table 35: Verification of Fleet Autonomy Parameter

LEVEL OF VEHICLE-TO-VEHICLE COMMUNICATIONS	UTILIZATION STATISTICS					
NONE	Software	average	max average	min average	max	min
	SHADO	27.04%	41.49%	18.53%	100%	0.00%
	Arena	26.02%	40.26%	14.60%	100%	0.00%
PARTIAL VEHICLE-TO-VEHICLE COMMUNICATIONS	Software	average	max average	min average	max	min
	SHADO	19.06%	29.34%	9.82%	100%	0.00%
	Arena	18.54%	26.79%	10.87%	100%	0.00%
HIGH VEHICLE-TO-VEHICLE COMMUNICATIONS	Software	average	max average	min average	max	min
	SHADO	8.10%	15.23%	3.16%	99.77%	0.00%
	Arena	7.89%	15.11%	2.76%	100.00%	0.00%

Table 36: Verification of Dispatcher Strategy Parameter

DISPATCHER ATTENTION ALLOCATION STRATEGY	UTILIZATION STATISTICS					
FIRST-IN, FIRST-OUT	Software	average	max average	min average	max	min
	SHADO	26.74%	36.93%	17.97%	100%	0.00%
	Arena	26.42%	40.26%	14.60%	100%	0.00%
PRIORITY	Software	average	max average	min average	max	min
	SHADO	26.41%	38.92%	16.93%	100%	0.00%
	Arena	26.47%	37.59%	17.02%	100%	0.00%
SHORTEST TASK FIRST	Software	average	max average	min average	max	min
	SHADO	26.05%	36.05%	16.10%	100%	0.00%
	Arena	26.47%	36.70%	17.02%	100%	0.00%

Table 37: Verification of Extreme Conditions Parameter

EXTREME CONDITIONS	UTILIZATION STATISTICS					
NONE	Software	average	max average	min average	max	min
	SHADO	26.74%	40.26%	15.87%	100%	0.00%
	Arena	26.42%	40.09%	14.60%	100%	0.00%
TYPE 1	Software	average	max average	min average	max	min
	SHADO	29.00%	43.56%	19.39%	100%	0.00%
	Arena	32.20%	60.50%	15.94%	100%	0.00%
TYPE 2	Software	average	max average	min average	max	min
	SHADO	29.20%	39.74%	18.51%	100%	0.00%
	Arena	29.12%	40.16%	16.08%	100%	0.00%
TYPE 1 AND TYPE 2	Software	average	max average	min average	max	min
	SHADO	30.85	52.32%	20.29%	100%	0.00%
	Arena	35.46%	65.96%	17.67%	100%	0.00%

Table 38: Verification of Team Coordination Parameter

TEAM COORDINATION	UTILIZATION STATISTICS					
NONE	Software	average	max average	min average	max	min
	SHADO	13.62%	21.56%	7.20%	95%	0.00%
	Arena	13.22%	18.56%	7.27%	81.64%	0.00%
33% COORDINATION	Software	average	max average	min average	max	min
	SHADO	14.53%	22.56%	8.11%	98%	0.00%
	Arena	10.11%	14.12%	7.18%	54.59%	0.00%
66% COORDINATION	Software	average	max average	min average	max	min
	SHADO	15.39%	23.13%	9.18%	100%	0.00%
	Arena	5.66%	7.49%	4.17%	29.07%	0.00%

We reported average utilization computed over 500 days. The max average represents the results of the one day with the highest average utilization across all the 10-minute intervals within that shift; the max represents the maximum utilization value for any one 10-minute interval across all days. SHADO and Arena results generally agreed with the maximum percentage of disagreement occurring from the Extreme Conditions internal parameter of both types of exogenous events (Type 1 could be a train derailment that introduces new task, Type 2 could be poor weather that increases times on all related tasks). In that case, SHADO had ~3% higher minimum average utilization, ~13% lower maximum average utilization, and ~5% lower overall average utilization compared to Arena. Yet, SHADO and Arena reported the same overall minimum and maximum utilizations at 0% and 100%, respectively.

Figure 62: Screenshot of Arena model in simulation software

Appendix G: Railroad Talk-and-Listen Time Data

Table 39 - Table 44 are copies of utilization data gathered directly from RGPC's recording system. Results from SHADO are included in further below.

Table 39: Talk-and-Listen Time Data per Shift on March 14, 2018

March14S0	March14S1	March14S2	March14S3
Grand Total 0:52:02	Grand Total 3:22:10	Grand Total 1:38:19	Grand Total 0:08:47
12:00 AM - 1:00 AM 0:14:29	6:00 AM - 7:00 AM 0:12:27	2:00 PM - 3:00 PM 0:25:20	10:00 PM - 11:00 PM 0:02:48
1:00 AM - 2:00 AM 0:11:03	7:00 AM - 8:00 AM 0:36:43	3:00 PM - 4:00 PM 0:18:33	11:00 PM - 12:00 AM 0:05:59
2:00 AM - 3:00 AM 0:05:52	8:00 AM - 9:00 AM 0:41:34	4:00 PM - 5:00 PM 0:06:29	
3:00 AM - 4:00 AM 0:02:13	9:00 AM - 10:00 AM 0:33:05	5:00 PM - 6:00 PM 0:11:57	
4:00 AM - 5:00 AM 0:04:55	10:00 AM - 11:00 AM 0:12:39	6:00 PM - 7:00 PM 0:09:44	
5:00 AM - 6:00 AM 0:13:30	11:00 AM - 12:00 PM 0:38:39	7:00 PM - 8:00 PM 0:11:06	
	12:00 PM - 1:00 PM 0:16:10	8:00 PM - 9:00 PM 0:04:34	
	1:00 PM - 2:00 PM 0:10:53	9:00 PM - 10:00 PM 0:10:36	

Table 40: Talk-and-Listen Time Data per Shift on March 16, 2018

March16S0	March16S1	March16S2	March16S3
Grand Total 0:19:57	Grand Total 3:34:06	Grand Total 1:28:53	Grand Total 0:02:17
12:00 AM - 1:00 AM 0:02:35	6:00 AM - 7:00 AM 0:08:50	2:00 PM - 3:00 PM 0:24:24	10:00 PM - 11:00 PM 0:02:04
1:00 AM - 2:00 AM 0:02:39	7:00 AM - 8:00 AM 0:32:56	3:00 PM - 4:00 PM 0:20:09	11:00 PM - 12:00 AM 0:00:13
2:00 AM - 3:00 AM 0:03:07	8:00 AM - 9:00 AM 0:30:54	4:00 PM - 5:00 PM 0:08:46	
3:00 AM - 4:00 AM 0:00:00	9:00 AM - 10:00 AM 0:59:13	5:00 PM - 6:00 PM 0:09:54	
4:00 AM - 5:00 AM 0:01:49	10:00 AM - 11:00 AM 0:15:26	6:00 PM - 7:00 PM 0:06:27	
5:00 AM - 6:00 AM 0:09:47	11:00 AM - 12:00 PM 0:24:04	7:00 PM - 8:00 PM 0:15:15	
	12:00 PM - 1:00 PM 0:18:02	8:00 PM - 9:00 PM 0:03:58	
	1:00 PM - 2:00 PM 0:24:41	9:00 PM - 10:00 PM 0:00:00	

Table 41: Talk-and-Listen Time Data per Shift on March 17, 2018

March17S0	March17S1	March17S2	March17S3
Grand Total 0:16:20	Grand Total 1:15:13	Grand Total 1:16:37	Grand Total 0:03:16
12:00 AM - 1:00 AM 0:00:20	6:00 AM - 7:00 AM 0:23:18	2:00 PM - 3:00 PM 0:10:08	10:00 PM - 11:00 PM 0:03:16
1:00 AM - 2:00 AM 0:09:35	7:00 AM - 8:00 AM 0:10:36	3:00 PM - 4:00 PM 0:14:44	11:00 PM - 12:00 AM 0:00:00
2:00 AM - 3:00 AM 0:00:00	8:00 AM - 9:00 AM 0:14:09	4:00 PM - 5:00 PM 0:11:44	
3:00 AM - 4:00 AM 0:03:29	9:00 AM - 10:00 AM 0:09:23	5:00 PM - 6:00 PM 0:03:50	
4:00 AM - 5:00 AM 0:00:46	10:00 AM - 11:00 AM 0:05:27	6:00 PM - 7:00 PM 0:01:21	
5:00 AM - 6:00 AM 0:02:10	11:00 AM - 12:00 PM 0:00:48	7:00 PM - 8:00 PM 0:10:14	
	12:00 PM - 1:00 PM 0:08:34	8:00 PM - 9:00 PM 0:14:41	
	1:00 PM - 2:00 PM 0:02:58	9:00 PM - 10:00 PM 0:09:55	

Table 42: Talk-and-Listen Time Data per Shift on May 16, 2018

May16S0	May16S1	May16S2	May16S3
Grand Total 0:48:00	Grand Total 3:36:50	Grand Total 2:44:27	Grand Total 0:10:27
12:00 AM - 1:00 AM 0:01:27	6:00 AM - 7:00 AM 0:20:06	2:00 PM - 3:00 PM 0:17:45	10:00 PM - 11:00 PM 0:04:44
1:00 AM - 2:00 AM 0:04:51	7:00 AM - 8:00 AM 0:41:31	3:00 PM - 4:00 PM 0:31:40	11:00 PM - 12:00 AM 0:05:43
2:00 AM - 3:00 AM 0:04:44	8:00 AM - 9:00 AM 0:41:11	4:00 PM - 5:00 PM 0:27:49	
3:00 AM - 4:00 AM 0:04:03	9:00 AM - 10:00 AM 0:38:22	5:00 PM - 6:00 PM 0:13:31	
4:00 AM - 5:00 AM 0:11:06	10:00 AM - 11:00 AM 0:21:26	6:00 PM - 7:00 PM 0:25:07	
5:00 AM - 6:00 AM 0:21:49	11:00 AM - 12:00 PM 0:21:07	7:00 PM - 8:00 PM 0:25:36	
	12:00 PM - 1:00 PM 0:21:59	8:00 PM - 9:00 PM 0:11:16	
	1:00 PM - 2:00 PM 0:11:08	9:00 PM - 10:00 PM 0:11:43	

Table 43: Talk-and-Listen Time Data per Shift on May 18, 2018

May18S0	May18S1	May18S2	May18S3
Grand Total 0:36:22	Grand Total 3:15:12	Grand Total 2:25:25	Grand Total 0:19:26
12:00 AM - 1:00 AM 0:00:00	6:00 AM - 7:00 AM 0:18:49	2:00 PM - 3:00 PM 0:27:53	10:00 PM - 11:00 PM 0:09:58
1:00 AM - 2:00 AM 0:05:40	7:00 AM - 8:00 AM 0:33:49	3:00 PM - 4:00 PM 0:27:39	11:00 PM - 12:00 AM 0:09:28
2:00 AM - 3:00 AM 0:09:38	8:00 AM - 9:00 AM 0:23:59	4:00 PM - 5:00 PM 0:18:19	
3:00 AM - 4:00 AM 0:01:51	9:00 AM - 10:00 AM 0:33:17	5:00 PM - 6:00 PM 0:19:22	
4:00 AM - 5:00 AM 0:00:30	10:00 AM - 11:00 AM 0:09:59	6:00 PM - 7:00 PM 0:14:25	
5:00 AM - 6:00 AM 0:18:43	11:00 AM - 12:00 PM 0:20:35	7:00 PM - 8:00 PM 0:18:30	
	12:00 PM - 1:00 PM 0:31:55	8:00 PM - 9:00 PM 0:14:11	
	1:00 PM - 2:00 PM 0:22:49	9:00 PM - 10:00 PM 0:05:06	

Table 44: Talk-and-Listen Time Data per Shift on May 20, 2018

May20S0	**May20S1**	**May20S2**	**May20S3**
Grand Total 0:28:41	Grand Total 1:02:36	Grand Total 0:47:04	Grand Total 0:07:01
12:00 AM - 1:00 AM 0:00:00	6:00 AM - 7:00 AM 0:05:11	2:00 PM - 3:00 PM 0:05:30	10:00 PM - 11:00 PM 0:00:00
1:00 AM - 2:00 AM 0:05:07	7:00 AM - 8:00 AM 0:02:43	3:00 PM - 4:00 PM 0:07:05	11:00 PM - 12:00 AM 0:07:01
2:00 AM - 3:00 AM 0:00:00	8:00 AM - 9:00 AM 0:19:26	4:00 PM - 5:00 PM 0:01:37	
3:00 AM - 4:00 AM 0:06:31	9:00 AM - 10:00 AM 0:06:40	5:00 PM - 6:00 PM 0:06:48	
4:00 AM - 5:00 AM 0:07:29	10:00 AM - 11:00 AM 0:10:05	6:00 PM - 7:00 PM 0:07:08	
5:00 AM - 6:00 AM 0:09:34	11:00 AM - 12:00 PM 0:09:33	7:00 PM - 8:00 PM 0:04:52	
	12:00 PM - 1:00 PM 0:07:07	8:00 PM - 9:00 PM 0:08:39	
	1:00 PM - 2:00 PM 0:01:51	9:00 PM - 10:00 PM 0:05:25	

Table 45: Hourly Freight Dispatcher Utilization per Shift of Default Talk-and-Listen Tasks

Table 46 are results from simulation.

Table 46: SHADO Results of Hourly Freight Dispatcher Utilization per Shift of Default Talk-and-Listen Tasks

	AM	PM	ON
Hour 1	63.41%	56.27%	24.69%
Hour 2	16.96%	21.63%	0.00%
Hour 3	6.25%	2.08%	0.00%
Hour 4	42.93%	23.75%	3.54%
Hour 5	61.02%	30.57%	3.82%
Hour 6	15.01%	38.01%	23.98%
Hour 7	14.85%	6.36%	0.00%
Hour 8	82.69%	36.79%	23.66%
Hour 1	53.28%	39.38%	21.77%
Hour 2	43.68%	26.61%	4.80%
Hour 3	34.88%	46.56%	0.00%
Hour 4	8.84%	32.07%	0.00%
Hour 5	22.97%	2.51%	0.54%
Hour 6	16.46%	4.11%	24.32%
Hour 7	21.73%	45.66%	9.41%
Hour 8	51.03%	24.98%	18.20%
Hour 1	70.06%	70.65%	25.21%
Hour 2	12.89%	38.72%	2.06%
Hour 3	10.48%	4.52%	0.00%
Hour 4	29.35%	20.52%	1.92%
Hour 5	5.93%	25.27%	0.00%
Hour 6	23.38%	30.18%	1.90%
Hour 7	20.91%	2.40%	0.00%
Hour 8	27.34%	22.83%	22.22%

	AM	PM	ON
Hour 1	58.01%	54.32%	38.91%
Hour 2	4.93%	4.79%	15.37%
Hour 3	9.69%	16.14%	4.00%
Hour 4	0.00%	11.44%	0.03%
Hour 5	17.41%	3.44%	5.81%
Hour 6	13.21%	28.93%	6.24%
Hour 7	43.55%	5.31%	9.48%
Hour 8	35.23%	23.39%	27.93%
Hour 1	98.29%	47.42%	21.77%
Hour 2	13.07%	4.84%	1.96%
Hour 3	7.46%	13.20%	0.00%
Hour 4	18.63%	6.68%	6.59%
Hour 5	39.21%	12.07%	2.95%
Hour 6	19.87%	12.31%	2.74%
Hour 7	34.76%	18.88%	24.35%
Hour 8	37.82%	31.96%	42.02%
Hour 1	68.00%	42.61%	25.18%
Hour 2	35.71%	32.28%	2.11%
Hour 3	6.68%	12.06%	8.30%
Hour 4	22.90%	3.99%	0.31%
Hour 5	14.04%	19.58%	9.40%
Hour 6	39.33%	32.57%	0.39%
Hour 7	27.60%	9.01%	14.65%
Hour 8	78.23%	35.89%	29.41%

Appendix H: Turing Test Data for Railroad Operations

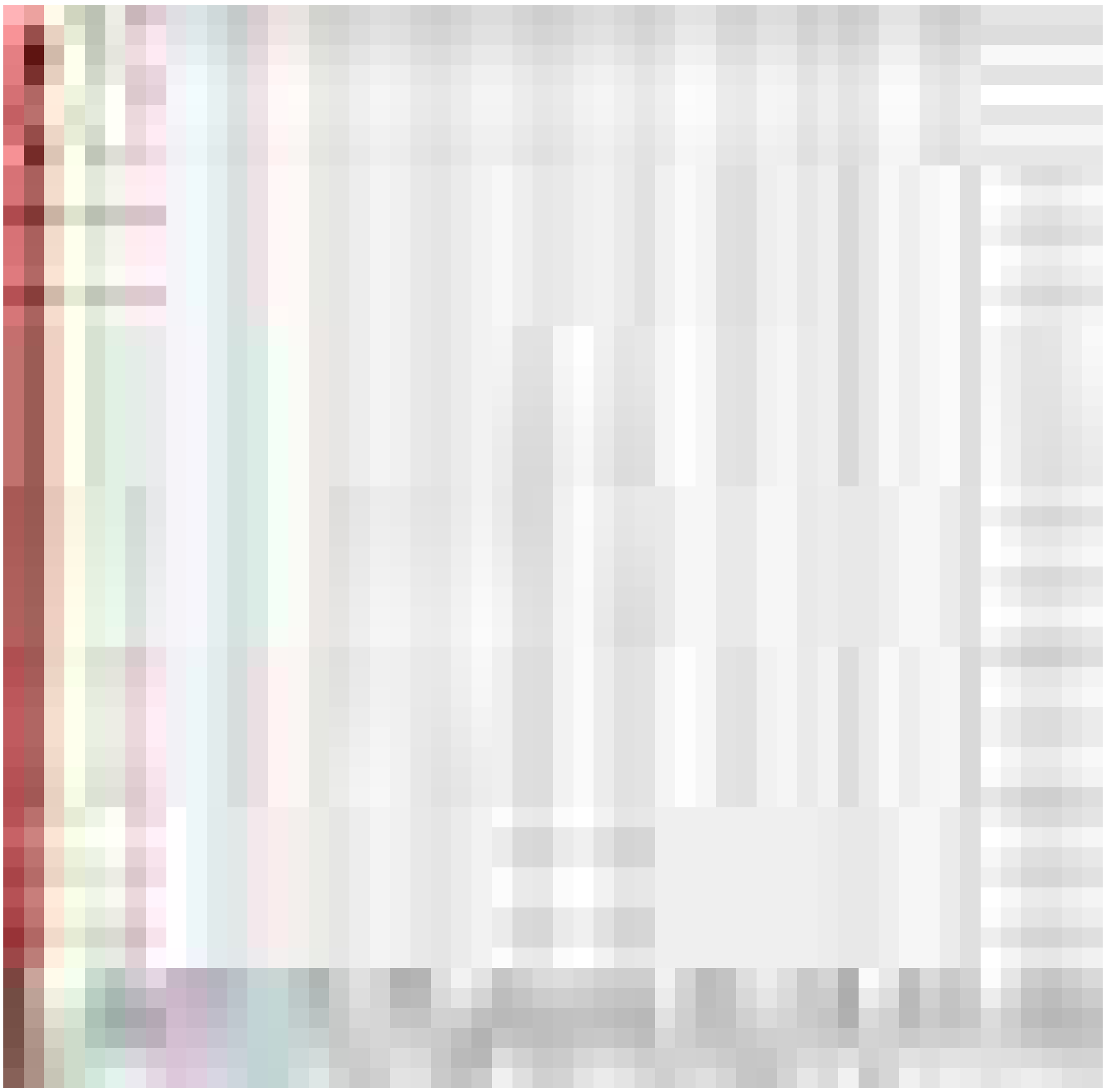


Figure 63: Copy of RGPC's overview records of talk-and-listen time

Table 47: Results from SHADO simulating same settings as RGPC per month over one year

MONTH	NUMBER OF REPLICATIONS	TOTAL TLT (HOURS)
JANUARY 2017	31 days	161.5
FEBRUARY 2017	28 days	114.5
MARCH 2017	31 days	145.5
APRIL 2017	30 days	138.75
MAY 2017	31 days	156.3
JUNE 2017	30 days	154.5
JULY 2017	31 days	155.2
AUGUST 2017	31 days	154.3
SEPTEMBER 2017	30 days	141.2
OCTOBER 2017	31 days	155.25
NOVEMBER 2017	30 days	145.9
DECEMBER 2017	31 days	171.17
AVERAGE MONTH 2017		149.5058333
MEDIAN MONTH 2017		154.4

Appendix I: Modified Input Parameters for Horizon Hourly Heuristic Test

Table 48: Interarrival Time Parameter Changes for Short-Haul (SH), Long-Haul (LH) and Focus (F) Flight Planning (FP) tasks

	Low Workload	*Moderate Workload*	*High Workload*
Expo Interarrival SH FP Parameter (minutes)	30	10.1	6
Avg Number of SH FP	2	5.940594059	10
Expo Interarrival LH FP Parameter (minutes)	-	as is	
Avg Number of LH FP	0	as is	<1
Expo Interarrival F FP Parameter (minutes)	-	as is	
Avg Number of F FP	0	as is	<1
Avg Utilization	0.2824667936	0.6661291833	0.8871650438
SD Utilization (from 365 reps)	0.04654113861	0.06286448546	0.03958303462

Appendix J: Design of Experiments for Sensitivity Analysis with Deviations in Task Input Parameters

Table 49: Design of Experiments for Sensitivity of Analysis with Task Inputs for Freight and Commuter Railroad and Airline Dispatch Operations in SHADO

Experiment #	*Deviation in Inter-arrival time*	*Deviation in Service time*
1	-75%	0%
2	-20%	0%
3	-5%	0%
4	5%	0%
5	20%	0%
6	75%	0%
7	0%	-75%
8	0%	-20%
9	0%	-5%
10	0%	5%
11	0%	20%
12	0%	75%

Appendix K: Design of Experiments for Sensitivity Analysis with Deviations in Internal Variables

Table 50: Design of Experiments for Sensitivity of Analysis with Internal Design and Staffing **Variables for Freight and Commuter Railroad and Airline in SHADO**

Experiment #	*Fleet Size*	*Fleet Heter.*	*Fleet Autonomy*	*Exogeneous*	*AIDA*	*Team Coordination*	*Team Size*	*Team Expertise*	*Shift Schedule*	*Operator Strategy*
1	2	1	1	1	1	1	1	1	1	1
2	3	1	1	1	1	1	1	1	1	1
3	3	2	1	1	1	1	1	1	1	1
4	3	3	1	1	1	1	1	1	1	1
5	3	1	2	1	1	1	1	1	1	1
6	3	1	3	1	1	1	1	1	1	1
7	1	1	1	2	1	1	1	1	1	1
8	1	1	1	3	1	1	1	1	1	1
9	1	1	1	4	1	1	1	1	1	1
10	1	1	1	1	2	1	1	1	1	1
11	1	1	1	1	3	1	1	1	1	1
12	1	1	1	1	4	1	1	1	1	1
13	1	1	1	1	1	1	2	1	1	1
14	1	1	1	1	1	2	2	1	1	1
15	1	1	1	1	1	3	2	1	1	1
16	1	1	1	1	1	4	2	1	1	1
17	1	1	1	1	1	5	2	1	1	1
18	1	1	1	1	1	1	3	1	1	1
19	1	1	1	1	1	1	2	2	1	1
20	1	1	1	1	1	1	2	3	1	1
21	1	1	1	1	1	1	1	1	2	1
22	1	1	1	1	1	1	1	1	3	1
23	1	1	1	1	1	1	1	1	4	1
24	1	1	1	1	1	1	1	1	5	1
25	1	1	1	1	1	1	1	1	6	1
26	1	1	1	1	1	1	1	1	1	2

Appendix L: Results from Sensitivity Analysis with Deviations in Internal Variables

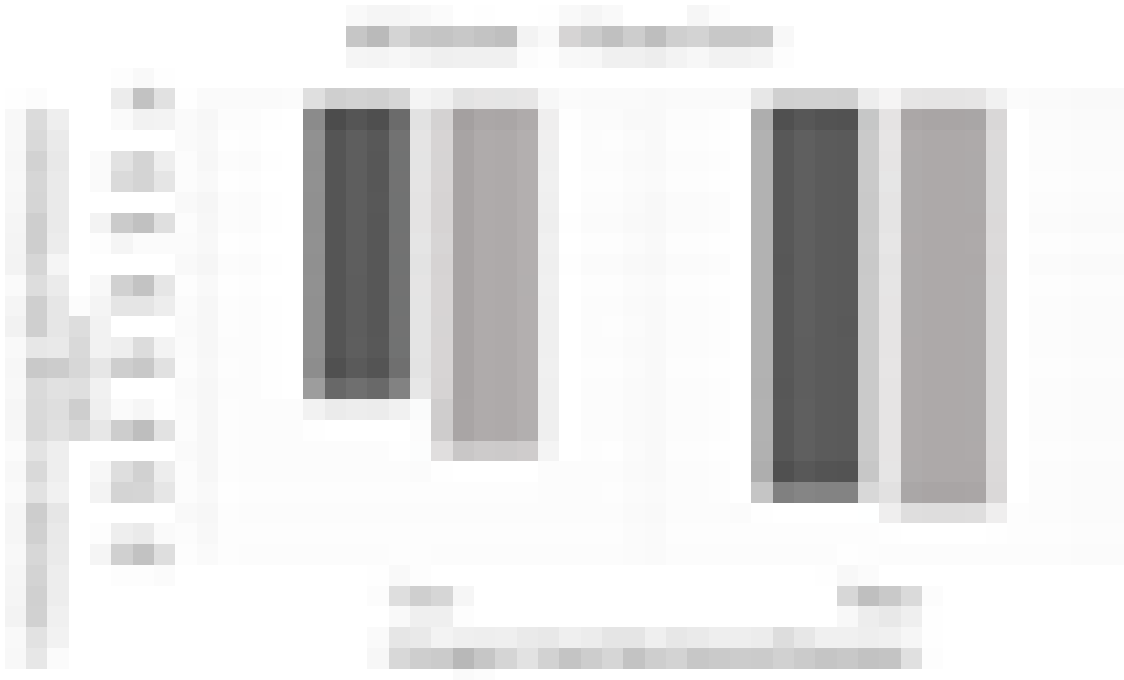


Figure 64: Commuter Dispatcher Workload and Error Deviation with Changes to Team Size

Figure 65: Commuter Dispatcher Workload and Error Deviation with Changes to Team Expertise

Figure 66: Commuter Dispatcher Workload and Error Deviation with Changes to Team Coordination

Figure 67: Commuter Dispatcher Workload and Error Deviation with Changes to Shift Schedule

Figure 68: Commuter Dispatcher Workload and Error Deviation with Changes to Operator Strategy

Figure 69: Commuter Dispatcher Workload and Error Deviation with Changes to Fleet Size

Figure 70: Commuter Dispatcher Workload and Error Deviation with Changes to Fleet Heterogeneity

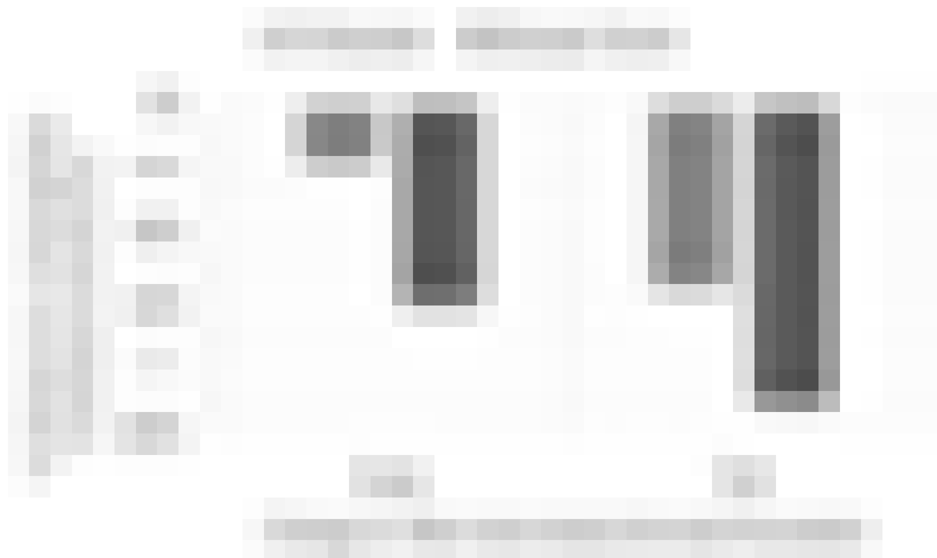


Figure 71: Commuter Dispatcher Workload and Error Deviation with Changes to Fleet Autonomy

Figure 72: Commuter Dispatcher Workload and Error Deviation with Changes to Exogeneous Event

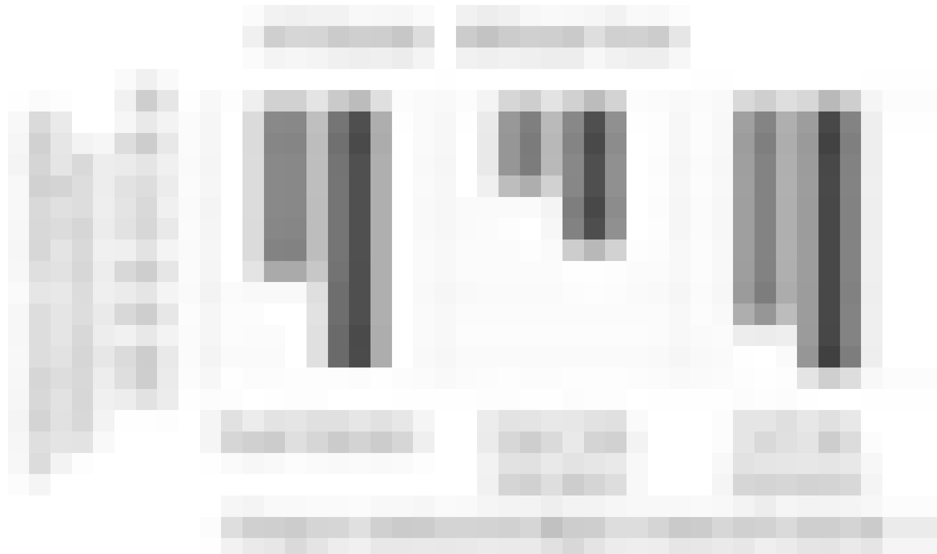


Figure 73: Commuter Dispatcher Workload and Error Deviation with Changes to AIDA

Figure 74 - Figure 83 show results from Horizon Air data used in SHADO.

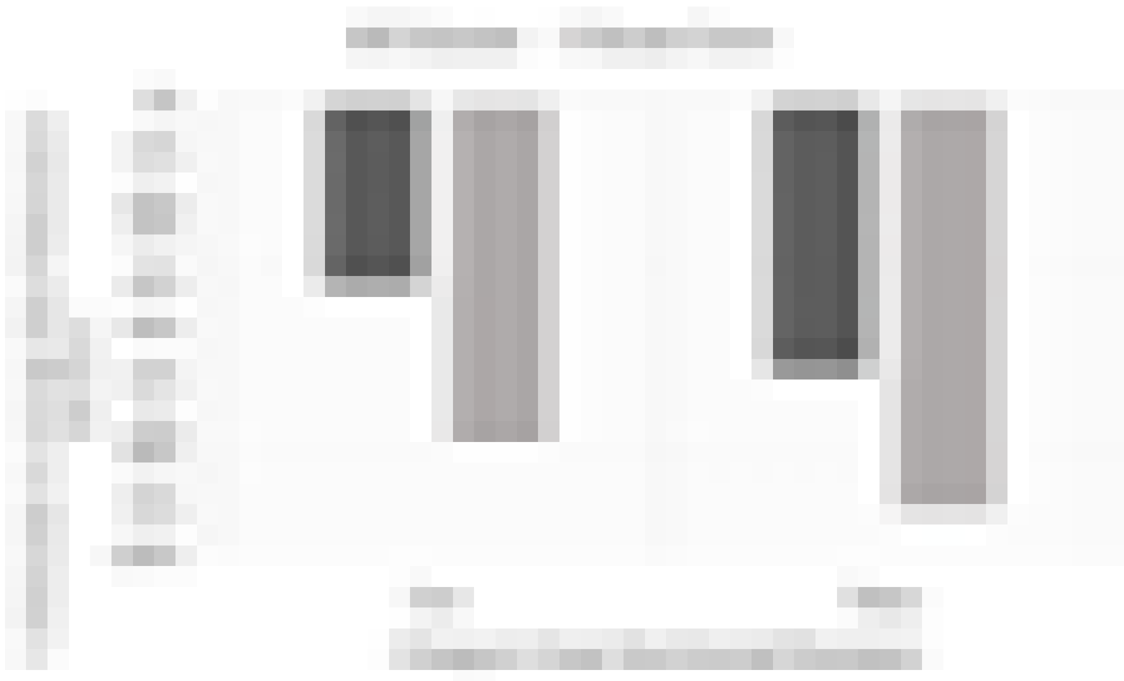


Figure 74: Airline Dispatcher Workload and Error Deviation with Changes to Team Size

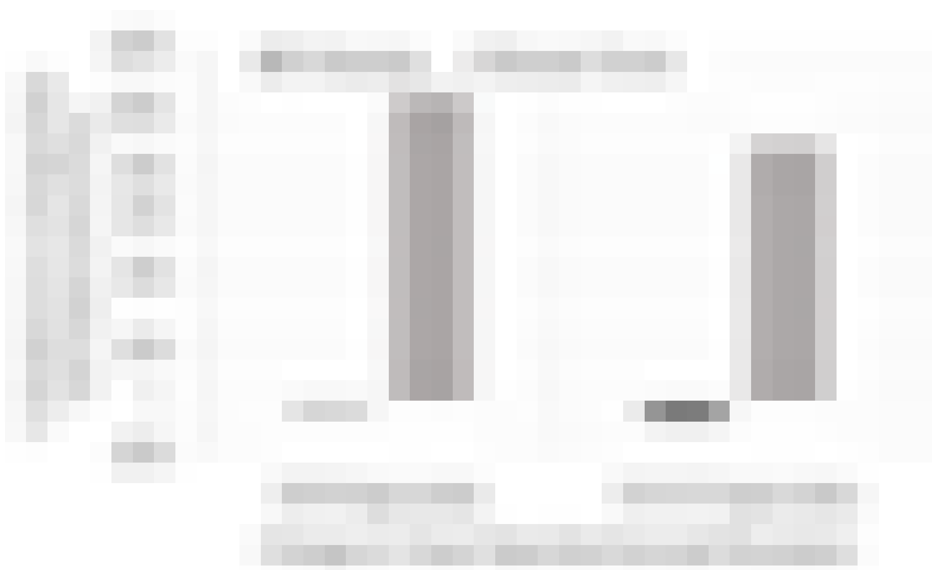


Figure 75: Airline Dispatcher Workload and Error Deviation with Changes to Team Expertise

Figure 76: Airline Dispatcher Workload and Error Deviation with Changes to Team Coordination

Figure 77: Airline Dispatcher Workload and Error Deviation with Changes to Shift Schedule

Figure 78: Airline Dispatcher Workload and Error Deviation with Changes to Operator Strategy

Figure 79: Airline Dispatcher Workload and Error Deviation with Changes to Fleet Size

Figure 80: Airline Dispatcher Workload and Error Deviation with Changes to Fleet Heterogeneity

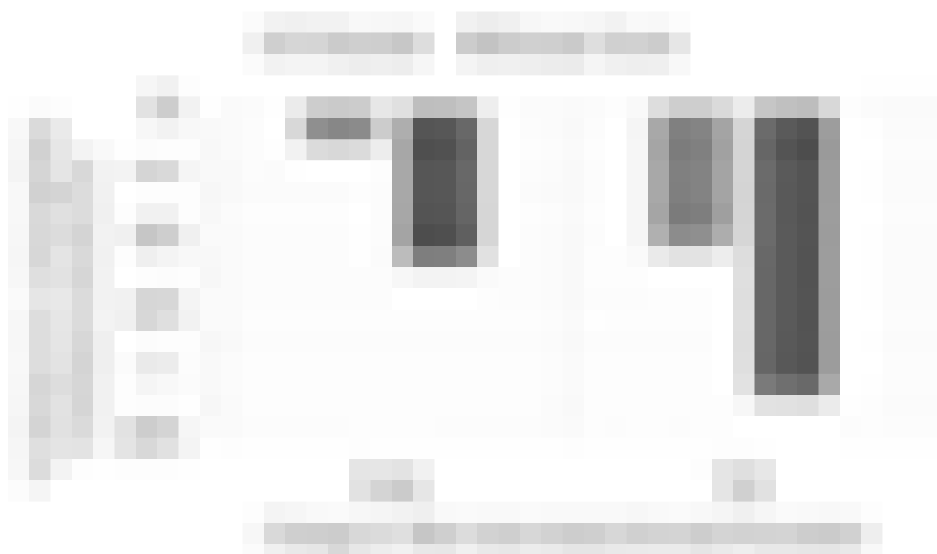


Figure 81: Airline Dispatcher Workload and Error Deviation with Changes to Fleet Autonomy

Figure 82: Airline Dispatcher Workload and Error Deviation with Changes to Exogenous Events

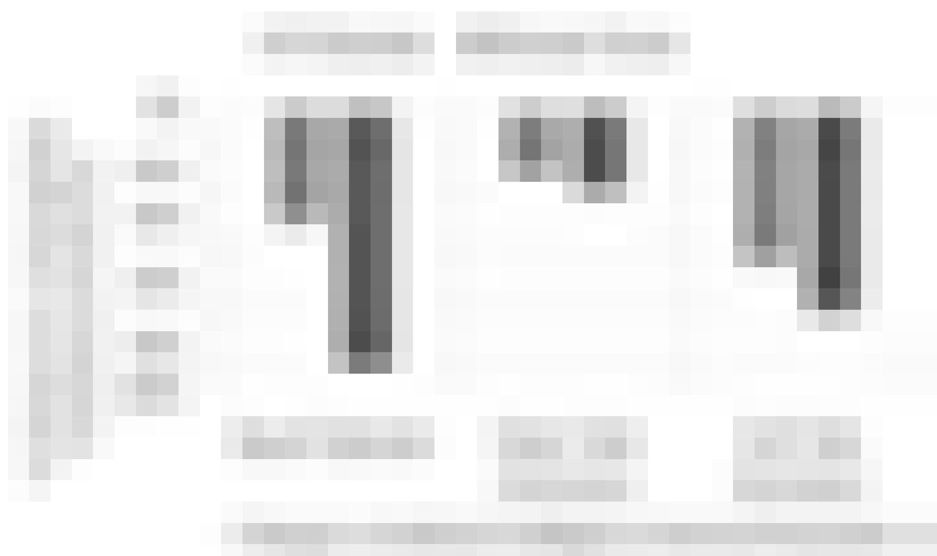


Figure 83: Airline Dispatcher Workload and Error Deviation with Changes to AIDA

In addition to studying the 10 key internal variables, two underlying models of fatigue (Hursh et al., 2004) and human error (Gibson, 2012) were investigated. Five different cases of human error probability parameters—each changed by a magnitude of

10. In two charts below, the impact of variances in human error probability (HEP) estimates derived by Gibson are presented. Figure 84 shows that there is an overall effect on dispatcher workload and error. Lower HEP results in smaller changes in workload whereas increasing HEP with equal magnitude results in much larger increases in average workload.

Figure 84: Impact of changes in human error probability on changes in dispatcher workload

In this next graph, Figure 85, we see that reducing the human error probability reduces total number of erred tasks. This total number includes the failed tasks due to HEP but also slips and lapses due to time constraints at work. There are significantly more overall erred tasks when HEP is increased by same magnitudes.

Figure 85: Impact of changes in human error probability on changes in overall number of erred tasks

Finally, Figure 86 presents the impact of fatigue on dispatcher workload for incremental as well as extreme time at work. The underlying homeostatic linear model of fatigue does not have significant impact on average workload.

Figure 86: The impact of duration of time at work on workload

Appendix M: Data from Simulating Automation at RGPC

Sharples et al. (2010) shared a summary of distribution of observed behaviors of dispatchers (better known as "signalers" in the UK) for non-automatic and automatic VDU systems with mean and S.D. percentage for each related task. We used the percent change (Table 51) from conventional to automatic to adjust the mean service time parameters (Table 52).

Table 51: Service Times Validated and from Observational Study (Sharples et al., 2010)

Dispatcher Task Types	**Default Mean (minutes)**	*Observed Conventional Relative Mean*	*Observed Automatic Relative Mean*	*Observed Conventional Relative Max*	*Observed Automatic Relative Min*	*Observed Conventional Relative Min*	*Observed Automatic Relative Max*
Train Movement	1.7	11.57%	7.50%	16.67%	3.28%	6.47%	11.72%
Bulletins	5	8.33%	11.14%	14.47%	5.00%	2.19%	17.28%
Temporary Bulletins	5	8.33%	11.14%	14.47%	5.00%	2.19%	17.28%
Bulletin Printing	15	1.57%	0.58%	3.99%	0.00%	0.00%	1.68%
Other Communications	2.8	5.34%	7.78%	9.72%	4.09%	0.96%	11.47%
Weather Recording	2.5	1.57%	0.58%	3.99%	0.00%	0.00%	1.68%
Notetaking	1	11.78%	9.04%	16.91%	4.57%	6.65%	13.51%
Reporting	10	1.57%	0.58%	3.99%	0.00%	0.00%	1.68%
Miscellaneous	5	4.38%	22.38%	9.71%	4.38%	0.00%	40.38%
Transfer-of-Duty	5	1.46%	2.08%	2.95%	0.00%	0.00%	4.47%

Table 52: Adjusted Service Time Parameters with Three Cases of Automation

Dispatcher Task Types	Service Time (minutes)		
	Best	**Most Likely**	**Worst**
Train Movement	Expo (0.86)	Expo (1.1)	Expo (1.2)
Bulletins	Expo (5.97)	Expo (6.7)	Expo (11.42)
Temporary Bulletins	Expo (5.97)	Expo (6.7)	Expo (11.42)
Bulletin Printing	Expo (5.54)	Expo (5.54)	Expo (6.32)
Other Communications	Expo (3.3)	Expo (4.08)	Expo (11.93)
Weather Recording	Expo (0.92)	Expo (0.92)	Expo (1.05)
Notetaking	Expo (0.69)	Expo (0.77)	Expo (0.8)
Reporting	Expo (3.69)	Expo (3.69)	Expo (4.21)
Miscellaneous	Expo (20.79)	Expo (20.79)	Expo (25.55)
Transfer-of-Duty	Expo (7.12)	Expo (7.12)	Expo (7.58)

Table 53: Default versus Mode Automated Commuter Dispatcher Utilization

	Commuter AM	**Commuter PM**	**Commuter ON**	**Mode Auto Commuter AM**	**Mode Auto Commuter PM**	**Mode Auto Commuter ON**
Average	46.82%	51.95%	27.20%	66.24%	66.15%	53.95%
Standard Deviation	6.98%	8.16%	6.56%	15.82%	15.73%	17.49%
Min Average	29.44%	30.99%	12.20%	29.10%	31.13%	14.25%
Max Average	66.24%	74.25%	49.09%	100.00%	100.00%	99.13%

Table 54: Worst versus Best Automated Commuter Dispatcher Utilization

	Worst Auto AM	Worst Auto PM	Worst Auto ON	Best Auto AM	Best Auto PM	Best Auto ON
Average	79.04%	78.93%	67.00%	55.26%	57.56%	46.84%
Standard Deviation	14.17%	13.15%	16.76%	14.74%	15.59%	14.52%
Min Average	38.90%	38.17%	27.86%	17.55%	24.92%	80.86%
Max Average	100.00%	100.00%	99.51%	98.15%	98.38%	15.83%

Figure 87 shows the proportion of time spent on four high-level task categories on the railroad commuter dispatcher desk: miscellaneous tasks, communication tasks, paperwork tasks, and control tasks. Communication tasks include both phone calls and the transfer-of-duty conversation. Paperwork includes bulletins, bulletin printing, weather recording, notetaking, and reporting. The one control task is train movement. Across the three shifts, the relative time dispatchers spend talking and doing non-functional tasks is predicted to increase with any of the three cases of automation.

Figure 87: Relative Time on Tasks w/ Different Cases of Automation

Appendix N: Modified Horizon Air Schedule for Geographically Reallocated Flights

Table 55: Proposed Oregon Desk Schedule during AM and PM shifts with Estimated Times of Departure (ETD) and Arrival (ETA) and Computed Interarrival Times (IATs)

ORIGIN	*ETDPDT*	*ETD IATs*	*DESTINATION*	*ETAPDT*	*HAUL*
Portland, Oregon, USA	600		Medford, Oregon, USA	656	SHORT
Bellingham, Washington, USA	600	0	Portland, Oregon, USA	700	SHORT
Spokane, Washington, USA	615	15	Boise, Idaho, USA	722	SHORT
Seattle, Washington, USA	625	10	Portland, Oregon, USA	713	SHORT
Portland, Oregon, USA	625	0	Spokane, Washington, USA	728	SHORT
Portland, Oregon, USA	635	10	Boise, Idaho, USA	748	SHORT
Pasco, Washington, USA	645	10	Portland, Oregon, USA	730	SHORT
Seattle, Washington, USA	645	0	Boise, Idaho, USA	812	SHORT
Seattle, Washington, USA	700	55	Eugene, Oregon, USA	806	SHORT
Seattle, Washington, USA	705	5	Medford, Oregon, USA	831	SHORT
Seattle, Washington, USA	710	5	Spokane, Washington, USA	806	SHORT
Boise, Idaho, USA	710	0	Spokane, Washington, USA	817	SHORT
Redmond, Washington, USA	720	10	Portland, Oregon, USA	757	SHORT
Seattle, Washington, USA	720	0	Pasco, Washington, USA	812	SHORT
Seattle, Washington, USA	730	10	Portland, Oregon, USA	817	SHORT
Medford, Oregon, USA	735	5	Portland, Oregon,	831	SHORT

			USA		
Seattle, Washington, USA	740	5	Boise, Idaho, USA	909	SHORT
Seattle, Washington, USA	800	60	Bellingham, Washington, USA	851	SHORT
Portland, Oregon, USA	815	15	Spokane, Washington, USA	919	SHORT
Spokane, Washington, USA	845	30	Portland, Oregon, USA	956	SHORT
Sacramento, California, USA	846	1	Portland, Oregon, USA	1026	SHORT
Eugene, Oregon, USA	847	1	Portland, Oregon, USA	928	SHORT
Boise, Idaho, USA	850	3	Portland, Oregon, USA	1010	SHORT
Seattle, Washington, USA	850	0	Spokane, Washington, USA	948	SHORT
Seattle, Washington, USA	855	5	Portland, Oregon, USA	950	SHORT
Oakland, California, USA	902	47	Portland, Oregon, USA	1050	SHORT
Reno, Nevada, USA	927	25	Portland, Oregon, USA	1106	SHORT
Seattle, Washington, USA	930	3	Portland, Oregon, USA	1019	SHORT
Seattle, Washington, USA	955	25	Wenatchee, Washington, USA	1036	SHORT
Seattle, Washington, USA	955	0	Spokane, Washington, USA	1053	SHORT
Spokane, Washington, USA	1000	45	Portland, Oregon, USA	1111	SHORT
Seattle, Washington, USA	1000	0	Portland, Oregon, USA	1052	SHORT
Santa Rosa, California, USA	1003	3	Portland, Oregon, USA	1144	SHORT
Sacramento, California, USA	1022	19	Portland, Oregon, USA	1202	SHORT
Boise, Idaho, USA	1022	0	Portland, Oregon, USA	1142	SHORT
Seattle, Washington,	1045	23	Yakima, Washington,	1130	SHORT

USA			USA		
Redmond, Washington, USA	1055	10	Portland, Oregon, USA	1135	SHORT
Seattle, Washington, USA	1055	0	Spokane, Washington, USA	1158	SHORT
Portland, Oregon, USA	1057	2	Medford, Oregon, USA	1154	SHORT
Portland, Oregon, USA	1058	1	Bellingham, Washington, USA	1157	SHORT
San Jose, California, USA	1115	57	Portland, Oregon, USA	1316	SHORT
Portland, Oregon, USA	1137	22	Spokane, Washington, USA	1241	SHORT
Seattle, Washington, USA	1140	3	Medford, Oregon, USA	1312	SHORT
Vancouver, British Columbia, Canada	1140	0	Portland, Oregon, USA	1250	SHORT
Seattle, Washington, USA	1208	68	Pasco, Washington, USA	1307	SHORT
Reno, Nevada, USA	1227	19	Portland, Oregon, USA	1404	SHORT
Bellingham, Washington, USA	1235	8	Portland, Oregon, USA	1336	SHORT
Portland, Oregon, USA	1309	74	Eugene, Oregon, USA	1347	SHORT
Redmond, Washington, USA	1310	1	Portland, Oregon, USA	1348	SHORT
Portland, Oregon, USA	1315	5	Spokane, Washington, USA	1419	SHORT
Spokane, Washington, USA	1320	5	Portland, Oregon, USA	1431	SHORT
Seattle, Washington, USA	1338	18	Eugene, Oregon, USA	1447	SHORT
Seattle, Washington, USA	1349	11	Wenatchee, Washington, USA	1440	SHORT
Medford, Oregon, USA	1350	1	Portland, Oregon, USA	1445	SHORT
Boise, Idaho, USA	1400	50	Portland, Oregon, USA	1519	SHORT
Seattle, Washington, USA	1420	20	Bellingham, Washington, USA	1505	SHORT

Eugene, Oregon, USA	1425	5	Portland, Oregon, USA	1504	SHORT
Santa Rosa, California, USA	1445	20	Portland, Oregon, USA	1624	SHORT
Portland, Oregon, USA	1516	71	Eugene, Oregon, USA	1554	SHORT
Seattle, Washington, USA	1530	14	Pasco, Washington, USA	1621	SHORT
Portland, Oregon, USA	1549	19	Spokane, Washington, USA	1652	SHORT

Table 56: Proposed Seattle Desk Schedule during AM and PM shifts with Estimated Times of Departure (ETD) and Arrival (ETA) and Computed Interarrival Times (IATs)

ORG	*ETDPDT*	*ETD IATs*	*DST*	*ETAPDT*	*HAUL*
Wenatchee, Washington, USA	600		Seattle, Washington, USA	642	SHORT
Vancouver, British Columbia, Canada	600	0	Seattle, Washington, USA	654	SHORT
Bozeman, Montana, USA	600	0	Seattle, Washington, USA	803	SHORT
Redmond, Washington, USA	615	15	Seattle, Washington, USA	720	SHORT
Portland, Oregon, USA	635	20	Seattle, Washington, USA	727	SHORT
Bellingham, Washington, USA	700	65	Seattle, Washington, USA	741	SHORT
Portland, Oregon, USA	700	0	Seattle, Washington, USA	748	SHORT
Boise, Idaho, USA	700	0	Seattle, Washington, USA	843	SHORT
Calgary, Alberta, Canada	700	0	Seattle, Washington, USA	847	SHORT
Reno, Nevada, USA	700	0	Seattle, Washington, USA	904	SHORT
Great Falls, Montana, USA	710	10	Seattle, Washington, USA	900	SHORT
Spokane, Washington, USA	715	5	Seattle, Washington, USA	822	SHORT

Victoria, British Columbia, Canada	728	13	Seattle, Washington, USA	826	SHORT
Eugene, Oregon, USA	745	17	Seattle, Washington, USA	857	SHORT
Seattle, Washington, USA	747	2	Redmond, Washington, USA	853	SHORT
Portland, Oregon, USA	800	53	Seattle, Washington, USA	853	SHORT
Portland, Oregon, USA	815	15	Seattle, Washington, USA	908	SHORT
Spokane, Washington, USA	820	5	Seattle, Washington, USA	932	SHORT
Portland, Oregon, USA	830	10	Seattle, Washington, USA	923	SHORT
Pasco, Washington, USA	903	73	Seattle, Washington, USA	1002	SHORT
Portland, Oregon, USA	910	7	Seattle, Washington, USA	1006	SHORT
Spokane, Washington, USA	910	0	Seattle, Washington, USA	1021	SHORT
Medford, Oregon, USA	923	13	Seattle, Washington, USA	1050	SHORT
Bellingham, Washington, USA	929	6	Seattle, Washington, USA	1018	SHORT
Portland, Oregon, USA	935	6	Redmond, Washington, USA	1013	SHORT
Redmond, Washington, USA	935	0	Seattle, Washington, USA	1049	SHORT
Vancouver, British Columbia, Canada	940	5	Seattle, Washington, USA	1043	SHORT
Portland, Oregon, USA	945	5	Vancouver, British Columbia, Canada	1053	SHORT
Seattle, Washington, USA	948	3	Kelowna, British Columbia, Canada	1054	SHORT
Boise, Idaho, USA	950	2	Seattle, Washington, USA	1128	SHORT
Missoula, Montana, USA	958	8	Seattle, Washington, USA	1137	SHORT
Portland, Oregon, USA	1030	72	Seattle, Washington, USA	1123	SHORT

Spokane, Washington, USA	1039	9	Seattle, Washington, USA	1147	SHORT
Wenatchee, Washington, USA	1115	76	Seattle, Washington, USA	1204	SHORT
Reno, Nevada, USA	1117	2	Seattle, Washington, USA	1321	SHORT
San Luis Obispo, California, USA	1128	11	Seattle, Washington, USA	1351	SHORT
Lewiston, Idaho, USA	1130	2	Seattle, Washington, USA	1249	SHORT
Spokane, Washington, USA	1135	5	Seattle, Washington, USA	1245	SHORT
Kelowna, British Columbia, Canada	1140	5	Seattle, Washington, USA	1251	SHORT
Seattle, Washington, USA	1145	5	Victoria, British Columbia, Canada	1235	SHORT
Portland, Oregon, USA	1151	6	Redmond, Washington, USA	1230	SHORT
Yakima, Washington, USA	1208	57	Seattle, Washington, USA	1253	SHORT
Seattle, Washington, USA	1219	11	Vancouver, British Columbia, Canada	1309	SHORT
Portland, Oregon, USA	1230	11	Seattle, Washington, USA	1320	SHORT
Medford, Oregon, USA	1234	4	Seattle, Washington, USA	1400	SHORT
Edmonton, Alberta, Canada	1234	0	Seattle, Washington, USA	1438	SHORT
Spokane, Washington, USA	1240	6	Seattle, Washington, USA	1348	SHORT
Victoria, British Columbia, Canada	1320	80	Seattle, Washington, USA	1407	SHORT
Santa Rosa, California, USA	1332	12	Seattle, Washington, USA	1540	SHORT
Seattle, Washington, USA	1335	3	Redmond, Washington, USA	1438	SHORT
Pasco, Washington, USA	1345	10	Seattle, Washington, USA	1447	SHORT
Vancouver, British Columbia, Canada	1355	10	Seattle, Washington, USA	1452	SHORT

Kalispell, Montana, USA	1400	45	Seattle, Washington, USA	1534	SHORT
Missoula, Montana, USA	1435	35	Seattle, Washington, USA	1610	SHORT
Helena, Montana, USA	1453	18	Seattle, Washington, USA	1644	SHORT
Spokane, Washington, USA	1500	47	Seattle, Washington, USA	1606	SHORT
Wenatchee, Washington, USA	1520	20	Seattle, Washington, USA	1607	SHORT
Redmond, Washington, USA	1520	0	Seattle, Washington, USA	1625	SHORT
Eugene, Oregon, USA	1525	5	Seattle, Washington, USA	1631	SHORT
Seattle, Washington, USA	1530	5	Victoria, British Columbia, Canada	1616	SHORT
Bellingham, Washington, USA	1548	18	Seattle, Washington, USA	1630	SHORT
Portland, Oregon, USA	1600	52	Seattle, Washington, USA	1649	SHORT

Table 57: Proposed Multiregional Desk Schedule during AM and PM with Estimated Times of Departure (ETD), Arrival (ETA), Computed Interarrival Times (IATs)

ORG	*ETDPDT*	*ETD IATs*	*DST*	*ETAPDT*	*HAUL*
Portland, Oregon, USA	625		Sacramento, California, USA	806	SHORT
Portland, Oregon, USA	625	0	Oakland, California, USA	822	SHORT
Portland, Oregon, USA	631	6	Santa Rosa, California, USA	812	SHORT
Seattle, Washington, USA	700	69	Fresno, California, USA	914	SHORT
Seattle, Washington, USA	720	20	Albuquerque, New Mexico, USA	1019	SHORT
San Jose, California, USA	730	10	Reno, Nevada, USA	828	SHORT
San Jose, California,	730	0	Burbank, California,	842	SHORT

USA			USA		
Seattle, Washington, USA	750	20	Missoula, Montana, USA	919	SHORT
Los Angeles, California, USA	753	3	Santa Rosa, California, USA	925	SHORT
Boise, Idaho, USA	757	4	San Jose, California, USA	950	SHORT
Boise, Idaho, USA	800	43	Sacramento, California, USA	939	SHORT
Seattle, Washington, USA	815	15	San Luis Obispo, California, USA	1048	SHORT
Seattle, Washington, USA	824	9	Missoula, Montana, USA	1011	SHORT
Portland, Oregon, USA	830	6	Boise, Idaho, USA	944	SHORT
Seattle, Washington, USA	833	3	Santa Rosa, California, USA	1042	SHORT
Seattle, Washington, USA	840	7	Reno, Nevada, USA	1037	SHORT
Boise, Idaho, USA	842	2	Reno, Nevada, USA	1000	SHORT
Santa Rosa, California, USA	850	8	Santa Ana, California, USA	1028	SHORT
Burbank, California, USA	922	72	San Jose, California, USA	1033	SHORT
San Francisco, California, USA	926	4	Santa Ana, California, USA	1112	SHORT
Portland, Oregon, USA	927	1	Albuquerque, New Mexico, USA	1209	SHORT
Seattle, Washington, USA	940	13	Lewiston, Idaho, USA	1047	SHORT
Seattle, Washington, USA	948	8	Edmonton, Alberta, Canada	1147	SHORT
Portland, Oregon, USA	1017	69	Reno, Nevada, USA	1149	SHORT
Albuquerque, New Mexico, USA	1105	88	San Diego, California, USA	1312	SHORT
Santa Ana, California, USA	1110	5	Santa Rosa, California, USA	1254	SHORT
Santa Rosa, California, USA	1120	10	Los Angeles, California, USA	1259	SHORT

Seattle, Washington, USA	1155	35	Kalispell, Montana, USA	1318	SHORT
Portland, Oregon, USA	1204	49	Boise, Idaho, USA	1318	SHORT
Seattle, Washington, USA	1225	21	Missoula, Montana, USA	1352	SHORT
Portland, Oregon, USA	1225	0	Santa Rosa, California, USA	1407	SHORT
Seattle, Washington, USA	1231	6	Helena, Montana, USA	1414	SHORT
Seattle, Washington, USA	1248	17	Santa Rosa, California, USA	1506	SHORT
Albuquerque, New Mexico, USA	1249	1	Santa Ana, California, USA	1449	SHORT
Portland, Oregon, USA	1405	156	Santa Barbara, California, USA	1620	SHORT
San Francisco, California, USA	1418	13	Santa Ana, California, USA	1604	SHORT
Portland, Oregon, USA	1420	2	Boise, Idaho, USA	1534	SHORT
Los Angeles, California, USA	1434	14	Missoula, Montana, USA	-697	SHORT
Seattle, Washington, USA	1438	4	Reno, Nevada, USA	1635	SHORT
Portland, Oregon, USA	1440	2	Sacramento, California, USA	1620	SHORT
Seattle, Washington, USA	1445	5	Edmonton, Alberta, Canada	1644	SHORT
San Jose, California, USA	1520	75	Boise, Idaho, USA	-695	SHORT
Seattle, Washington, USA	1530	10	Bozeman, Montana, USA	-687	SHORT
Seattle, Washington, USA	1530	0	Missoula, Montana, USA	1656	SHORT
Santa Rosa, California, USA	1544	14	Los Angeles, California, USA	-679	SHORT
Los Angeles, California, USA	1559	15	Santa Rosa, California, USA	-663	SHORT
Portland, Oregon, USA	1603	44	Reno, Nevada, USA	-661	SHORT

Portland, Oregon, USA	1604	1	San Jose, California, USA	-596	SHORT
Portland, Oregon, USA	1317		Kansas City, Missouri, USA	1645	LONG

References

110th US Congress. Rail Safety Improvement Act (RSIA) (2008).

Airline Dispatchers Federation. (2014). Early Dispatch History. Retrieved from https://www.dispatcher.org/dispatcher/early-dispatch-history

Aksin, Z., Armony, M., & Mehrotra, V. (2007). Customer Behavior Modeling in Revenue Management and Auctions: A Review and New Research Opportunities. *Society*, *16*(6), 665–688. https://doi.org/10.3401/poms.

Alighanbari, M., & How, J. (2006). Robust Decentralized Task Assignment for Cooperative UAVs. In *AIAA Guidance, Navigation, and Control Conference and Exhibit* (pp. 1–16). Keystone, Colorado: American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/6.2006-6454

Allender, L., Kelley, T. D., Salvi, L., Lockett, J., Headley, D. B., Promisel, D., … Feng, T. (1995). Verification, validation, and accreditation of a Soldier-System Modeling Tool. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1219–1223).

Barnes, C. M., Hollenbeck, J. R., Wagner, D. T., DeRue, D. S., Nahrgang, J. D., & Schwind, K. M. (2008). Harmful help: The costs of backing-up behavior in teams. *Journal of Applied Psychology*, *93*(3), 529–539. https://doi.org/10.1037/0021-9010.93.3.529

Bergum, B. O., & Lehr, D. J. (1962). Vigilance performance as a function of paired monitoring. *Journal of Applied Psychology*, *46*(5), 341.

Berry, K. A., & Pace, J. J. (2011). Examining the Actors and Functions of an Airline Operations Center. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 1412–1416).

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing Communication Sequences for Team Training Needs Assessment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(4), 672–679. https://doi.org/10.1518/001872098779649265

Brecher, A., & Shurland, M. (2015). Study on Improving Rail Energy Efficiency (E2): Best Practices and Strategies. In *Joint Rail Conference* (pp. 1–8).

Bureau of Transportation Statistics. (2016). National Transportation Statistics. Retrieved

July 26, 2017, from http://www.bts.gov/publications/national_transportation_statistics/

Burgard, W., Moors, M., Fox, D., Simmons, R., & Thrun, S. (2000). Collaborative multi-robot exploration. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on* (Vol. 1, pp. 476–481).

Burns, T., & Stalker, G. M. (1994). Mechanistic and Organic Systems of Management. In *The Management of Innovation*. https://doi.org/10.1093/acprof

Burton, R. M., & Obel, B. (2018). The science of organizational design : fit between structure and coordination. *Journal of Organization Design*, *7*(5).

Carrel, A., Mishalani, R. G., Wilson, N. H. M., & Attanucci, J. P. (2013). A framework for evaluating operations control on a metro line: Integrating multiple perspectives and automatically collected train and passenger movement data. *Public Transport*, *5*(3), 149–176. https://doi.org/10.1007/s12469-013-0065-9

Chandler, P. R. (2004). Cooperative control of a team of UAVs for tactical missions. In *Proceedings of the AIAA 1st Intelligent Systems Technical Conference*.

Choi, H.-L., Brunet, L., & How, J. P. (2009). Consensus-based Decentralized Auctions for Robusttask Allocation. *IEEE Transactions on Robotics*, *25*(4), 912–926. https://doi.org/10.1109/TRO.2009.2022423

Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C., & Cannon-Bowers, J. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, *7*(3), 179–199. https://doi.org/10.1037/1089-2699.7.3.179

Crandall, J. W., & Cummings, M. (2007a). Attention Allocation Efficiency in Human-UV Teams. *AIAA Infotech@Aerospace 2007 Conference and Exhibit*, (May). https://doi.org/10.2514/6.2007-2704

Crandall, J. W., & Cummings, M. L. (2007b). Identifying Predictive Metrics for Supervisor Control of Multiple Robots. *IEEE Transactions on Robotics*, *23*(5), 942–951. https://doi.org/{10.1109/TRO.2007.907480

Crandall, J. W., Goodrich, M. A., Nielsen, C. W., & Olsen, D. R. (2005). Validating Human-Robot Interaction Schemes in Multi-Tasking Environments. *IEEE Transactions on Systems, Man, and Cybernetics -Part A: Systems and Humans*, *35*(4), 438–449. Retrieved from http://scholarsarchive.byu.edu/facpub%5Cnhttp://scholarsarchive.byu.edu/facpub/363

Cummings, Bertucelli, L. F., Macbeth, J., & Surana, A. (2014). Task versus vehicle-based control paradigms in multiple unmanned vehicle supervision by a single operator. *IEEE Transactions on Human-Machine Systems*, *44*(3), 353–361. https://doi.org/10.1109/THMS.2014.2304962

Cummings, M. L. (2004). *Human Supervisory Control of Swarming Networks*. Cambridge, MA.

Cummings, M. L. (2015). Operator interaction with centralized versus decentralized UAV architectures. In *Handbook of Unmanned Aerial Vehicles* (pp. 977–992). Springer.

Cummings, M. L., Gao, F., & Thornburg, K. M. (2016). Boredom in the workplace: A new look at an old problem. *Human Factors*, *58*(2), 279–300. https://doi.org/10.1177/0018720815609503

Cummings, M. L., & Guerlain, S. (2007). Developing operator capacity estimates for supervisory control of autonomous vehicles. *Human Factors*, *49*(1), 1–15. https://doi.org/10.1518/001872007779598109

Cummings, M. L., Mastracchio, C., Thornburg, K. M., & Mkrtchyan, A. (2013). Boredom and distraction in multiple unmanned vehicle supervisory control. *Interacting with Computers*, *25*(1), 34–47. https://doi.org/10.1093/iwc/iws011

Cummings, M. L., & Mitchell, P. J. (2005). Operator scheduling strategies in supervisory control of multiple UAVs. *Aerospace Science and Technology*, *11*(4), 339–348. https://doi.org/{10.1016/j.ast.2006.10.007

Cummings, M. L., & Mitchell, P. J. (2008). Predicting Controller Capacity in Supervisory Control of Multiple UAVs. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, *38*(2), 451–460. Retrieved from http://dx.doi.org/10.1016/j.ast.2006.10.007

Cummings, M. L., & Nehme, C. E. (2010). Modeling the impact of workload in network centric supervisory control settings. *Neurocognitive and Physiological Factors during High-Tempo Operations*, 23–40. Retrieved from http://web.mit.edu/aeroastro/labs/halab/papers/Cummings_Ashgate_Chapter.pdf

Cummings, M. L., Nehme, C. E., Crandall, J., & Mitchell, P. (2007). Predicting operator capacity for supervisory control of multiple UAVs. In J. S. Chahl, L. C. Jain, A.

Mizutani, & M. Sato-Ilic (Eds.), *Studies in Computational Intelligence: Innovations in Intelligent Machines 1* (Vol. 70). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-72696-8_2

Cummings, M. L., & Ryan, J. (2013). Shared Authority Concerns in Automated Driving Applications. *Journal of Ergonomics*, *3*, 1–4. https://doi.org/10.4172/2165-7556.S3-001

Devoe, D. B. (1974). *An Analysis of the Job of Railroad Train Dispatcher*. https://doi.org/10.1051/ita

Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission control of multiple unmanned aerial vehicles: a workload analysis. *Human Factors*, *47*(3), 479–487. https://doi.org/10.1518/001872005774860005

Donmez, B., Nehme, C., & Cummings, M. L. (2010). Modeling Workload Impact in Multiple Unmanned Vehicle Supervisory Control. *IEEE Transactions on Systems, Man, and Cybernetics*, *40*(6), 1180–1190.

Dorrian, J., Roach, G. D., Fletcher, A., & Dawson, D. (2007). Simulated train driving: Fatigue, self-awareness and cognitive disengagement. *Applied Ergonomics*, *38*(2), 155–166. https://doi.org/10.1016/j.apergo.2006.03.006

Drews, C., Laughery, K., Kramme, K., & Archer, R. (1985). *LHX Cockpits: Micro SAINT Simulation Study and Results*. Dallas, TX.

Driskell, J., & Salas, E. (1992). Collective Behavior and Team Performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *34*(2), 277–288. https://doi.org/10.1177/001872089203400303

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

FAA. Responsibility for operational control: Domestic operations (1964). United States. Retrieved from https://www.law.cornell.edu/cfr/text/14/121.533

Federal Aviation Administration. Advisory Circular: Dispatch Resource Management Training (2005). United States of America.

Federal Aviation Administration. Aircraft Dispatcher Practical Test Standards (2013). USA: Flight Standards Service.

Federal Railroad Administration. (1995). *National Train Dispatcher's Follow-up Review*. Washington, D.C.

Federal Railroad Administration. Appendix A to Part 228-- Requirements of the Hours of Service Act: Statement of Agency Policy and Interpretation (2012).

Fontenelle, G., & Laughery, K. R. (1988). A Workload Assessment Aid for Human Engineering Design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1122–1125). Anaheim, CA: SAGE Publications Inc.

FRA. (2016). Summary of Accident/Incident Counts from January-December (2015 Preliminary). Retrieved January 1, 2016, from http://safetydata.fra.dot.gov/officeofsafety/publicsite/summary.aspx

Fries, I. B. (2008). Frequent Frequencies - AOPA. Retrieved October 11, 2017, from https://www.aopa.org/news-and-media/all-news/2008/august/flight-training-magazine/frequent-frequencies

Gao, F. (2013). Modeling Teamwork of Multi-Human Multi-Agent Teams. In *Computer Supported Cooperative Work* (pp. 47–50). San Antonio, TX: ACM. https://doi.org/10.1145/2441955.2441969

Gao, F., & Cummings, M. (2012). Using Discrete Event Simulation to Model Multi-Robot Multi-Operator Teamwork. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, pp. 2093–2097). Boston: SAGE Publications Inc. https://doi.org/10.1177/1071181312561444

Gao, F., & Cummings, M. (2014). Barriers to Robust and Effective Human-Agent Teamwork. *2014 AAAI Spring Symposium Series*, 36–41. Retrieved from http://www.aaai.org/ocs/index.php/SSS/SSS14/paper/viewPDFInterstitial/7680/7727

Gao, F., Cummings, M. L., & Solovey, E. T. (2014). Modeling Teamwork in Supervisory Control of Multiple Robots. *IEEE Transactions on Human-Machine Systems*, *44*(4), 441–453.

Garnett, O., Mandelbaum, a., & Reiman, M. (2002). Designing a Call Center with Impatient Customers. *Manufacturing & Service Operations Management*, *4*(3), 208–227. https://doi.org/10.1287/msom.4.3.208.7753

Gertler, J., & Nash, D. (2004). *Optimizing Staffing Levels and Schedules for Railroad Dispatching Centers* (DFRA.01035). Washington, D.C.: U.S. Department of Transportation.

Gibson, H. (2012). *Railway Action Reliability Assessment user manual - A technique for the quantification of human error in the rail industry*. Retrieved from http://www.rssb.co.uk/sitecollectiondocuments/pdf/reports/research/T270-guide-final4.pdf

Goodrich, M. A., Quigley, M., & Cosenzo, K. (2005). Task Switching and Multi-Robot Teams. In L. E. Parker, F. E. Schneider, & A. C. Schultz (Eds.), *Multi-Robot Systems. From Swarms to Intelligent Automata Volume III: Proceedings from the 2005 International Workshop on Multi-Robot Systems* (pp. 185–195). Dordrecht: Springer Netherlands. https://doi.org/10.1007/1-4020-3389-3_15

Gopher, D. (1980). On the Training of Time Sharing Skills: An Attention Viewpoint. *Proceedings of the Human Factors Society Annual Meeting*, *24*(1), 259–263. https://doi.org/10.1177/107118138002400170

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, *49*(12–13), 1312–1325.

Graeber, R. C. (1988). Aircrew fatigue and circadian rhythmicity.

Green, L., Kolesar, P., & Svoronos, A. (1991). Some Effects of Nonstationarity on Multiserver Markovian Queueing Systems. *Operations Research*, *39*(3), 502–511. https://doi.org/10.1287/opre.39.3.502

Green, L., Kolesar, P., & Whitt, W. (2007). Coping with Time-Varying Demand When Setting Staffing Requirements for a Service System. *Production and Operations Management*, *16*(1), 13–39. https://doi.org/10.1111/j.1937-5956.2007.tb00164.x

Hall, R. (1990). *Queueing Methods—For Services and Manufacturing*.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index). *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Holden, J., & Goel, N. (2016). *Fast-Forwarding to a Future of On-Demand Urban Air Transportation* (Series No. Number). San Francisco.

Howard, A., Mataric, M. J., & Sukhatme, G. S. (2002). Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem. *Distributed Autonomous Robotic Systems*, *5*, 299–308. https://doi.org/10.1.1.12.8990

Huang, L., Cummings, M. L., & Nneji, V. C. (2018). Preliminary Analysis and Simulation of Railroad Dispatcher Workload. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, pp. 691–695). Philadelphia: SAGE Publications Inc. https://doi.org/10.1177/1541931218621156

Humphreys, P. (1998). *RAMS Simulation of Romainian Airspace*.

Hungerford, E. (1946). *Men of Erie, a story of human effort,*. New York: Random House.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., … Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*, *75*(3, Suppl.), A44-53.

International Association of Public Transport. (2016). *World Report on Metro Automation*. Retrieved from http://www.uitp.org/sites/default/files/cck-focus-papers-files/UITP_Statistic Brief_World Metro Automation 2016_LQ+_0.pdf

Johannsen, G. (1979). Workload and Workload Measurement. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 3–11). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4757-0884-4_1

Kirby, N. H., & Nettelbeck, T. (1991). Speed of information processing and age. *Personality and Individual Differences*, *12*(2), 183–188.

Koetse, M. J., & Rietveld, P. (2009). The impact of climate change and weather on transport: An overview of empirical findings. *Transportation Research Part D: Transport and Environment*, *14*(3), 205–221. https://doi.org/10.1016/j.trd.2008.12.004

Lam, K., & Lau, R. S. M. (2004). A simulation approach to restructuring call centers. *Business Process Management Journal*, *10*(4), 481–494.

Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team Mental Models: Techniques, Methods, and Analytic Approaches. *Human Factors*, *42*(2), 242–271. https://doi.org/10.1518/001872000779656534

Laughery, R. (1999). Using Discrete-Event Simulation To Model Human Performance in Complex Systems. In *1999 Winter Simulation Conference* (pp. 815–820).

Law, A. (2007). Statistical analysis of simulation output data: the practical state of the art. In *Simulation Conference, 2007 Winter* (pp. 77–83). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4419590

Law, & Kelton, W. (2000). Simulation Modeling and Analysis. *ACM Transactions on Modeling and Computer Simulation*. https://doi.org/10.1145/1667072.1667074

Lawrence, J. A., & Pasternack, B. A. (2002). *Applied management science: modeling, spreadsheet analysis, and communication for decision making*. Wiley New York.

Lewis, M., Polvichai, J., Sycara, K., & Scerri, P. (2006). Scaling-up human control for large UAV teams. In *Human factors of remotely operated vehicles* (pp. 237–250). Emerald Group Publishing Limited.

Lewis, M., Wang, H., Chien, S. Y., Velagapudi, P., Scerri, P., & Sycara, K. (2011). Process and Performance in Human-Robot Teams. *Journal of Cognitive Engineering and Decision Making*, *5*(2), 186–208. https://doi.org/10.1177/1555343411409323

Linden, D. Van Der, Frese, M., & Meijman, T. F. (2006). Mental Fatigue and Cognitive Control. *Acta Psychologica*, 1–38.

Loft, S., Sanderson, P. M., Neal, A., & Mooij, M. (2007). Modeling and Predicting Mental Workload in En Route Air Traffic Control: Critical Review and Broader Implications. *Human Factors*, *49*(3), 376–399. https://doi.org/10.1518/001872007X197017.

Lysaght, R. J., Hill, S. G., Dick, a O., Plamondon, B. D., Linton, P. M., Wierwille, W. W., … Wherry, R. J. (1989). Operator workload: Comprehensive review and evaluation of operator workload methodologies. *United States Army Research Institute for the Behavioral Sciences, Technical Report*, *851*, 903–986.

Macmillan, J., Entin, E. E., & Serfaty, D. (2004). Communication Overhead: The Hidden Cost of Team Cognition. *Team Cognition: Process and Performance at the Inter- and Intra-Individual Level*, (August 2016), 61–82. https://doi.org/10.1080/03637759309376288

Majumdar, A., & Ochieng, W. Y. (2002). The factors affecting air traffic controller workload : a multivariate analysis based upon simulation modelling of controller workload . *Transportation Research Record*, 1–41. https://doi.org/10.3141/1788-08

Majumdar, A., & Polak, J. (2001). Estimating Capacity of Europe's Airspace Using a Simulation Model of Air Traffic Controller Workload. *Transportation Research Record*, *1744*(1), 30–43. https://doi.org/10.3141/1744-05

Martland, C. D. (1982). Workload Measurement and Train Crew Consist Adjustments: The Boston & Maine Experience. *Transportation Journal*, *21*(3), 34–57.

Mazzuchi, T. A., & Wallace, R. B. (2004). Analyzing skill-based routing call centers using discrete-event simulation and design experiment. In R. G. Ingalls, M. D. Rossetti, J. S. Smith, & B. A. Peters (Eds.), *36th Conference on Winter Simulation* (pp. 1812–1820). Washington, D.C.

Mekdeci, B., & Cummings, M. L. (2009). Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles. *Proceedings of the 9th Workshop on Performance Metrics for Intelligent Systems*, 1–8. https://doi.org/10.1145/1865909.1865911

Mogford, R. H., Guttman, J. A., Morrow, S. L., & Kopardekar, P. (1995). *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature*.

Mogford, R. H., Murphy, E. D., & Guttman, J. A. (1994). Using knowledge exploration tools to study airspace complexity in air traffic control. *The International Journal of Aviation Psychology*, *4*(1), 29–45.

Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, *22*(2), 89–106. https://doi.org/10.1002/job.86

Mueller, E. R., Kopardekar, P. H., & Goodrich, K. H. (2017). Enabling Airspace Integration for High-Density On-Demand Mobility Operations. In *17th AIAA Aviation Technology, Integration, and Operations Conference* (pp. 1–24). https://doi.org/10.2514/6.2017-3086

National Research Council (U.S.). (1993). *Workload Transition: Implications for Individual and Team Performance*. (B. M. Huey & C. D. Wickens, Eds.). EBSCO. Retrieved from http://books.google.com/books?hl=fr&lr=&id=rhz5mUWYwJ4C&pgis=1

National Transportation Safety Board. (1999). *Evaluation of U.S. Department of Transportation Efforts in the 1990s to Address Operator Fatigue*. Washington, D.C. https://doi.org/Safety Report NTSB/SR-99/01

Naylor, J. C., & Dickinson, T. L. (1969). Task structure, work structure, and team performance. *Journal of Applied Psychology*, *53*(3p1), 167.

Nehme, C. E. (2009). *Modeling human supervisory control in heterogeneous unmanned vehicle systems*. MIT. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA530930

Nehme, C. E., Mekdeci, B., Crandall, J. W., & Cummings, M. L. (2009). The Impact of Heterogeneity on Operator Performance in Future Unmanned Vehicle Systems. *The International C2 Journal*, *2*(2), 1–30. https://doi.org/10.1007/s13398-014-0173-7.2

Neth, H., Khemlani, S. S., Oppermann, B., & Gray, W. D. (2006). Juggling multiple tasks: A rational analysis of multitasking in a synthetic task environment. *Human Factors*, 1142–1146. https://doi.org/10.1177/154193120605001106

Nneji, V. C., Cummings, M. L., & Stimpson, A. J. (2019). Predicting Locomotive Crew Performance in Rail Operations with Human and Automation Assistance. *IEEE Transactions on Human-Machine Systems*, in production.

Nneji, V. C., Cummings, M. L., Stimpson, A. J., & Goodrich, K. H. (2018). Functional Requirements for Remotely Managing Fleets of Personal On-Demand Aircraft. In *AIAA Aerospace Sciences Meeting, Science and Technology Forum*. Orlando: American Institute of Aeronautics and Astronautics. https://doi.org/doi:10.2514/6.2018-2007

Nneji, V. C., Stimpson, A., Cummings, M. (Missy), & Goodrich, K. H. (2017). Exploring Concepts of Operations for On-Demand Passenger Air Transportation. In NASA (Ed.), *17th AIAA Aviation Technology, Integration, and Operations Conference* (pp. 1–12). Denver, CO: American Institute of Aeronautics and Astronautics AVIATION Forum.

Olsen, D. R., & Goodrich, M. A. (2003). Metrics for evaluating human-robot interactions. *Proceedings of PERMIS*, *2003*(4), 507–527. https://doi.org/10.1016/j.intcom.2005.10.004

Olsen, D. R., & Wood, S. B. (2004). Fan-out: Measuring Human Control of Multiple Robots. In *CHI* (Vol. 6, pp. 231–238). Vienna, Austria.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parker, L. E. (2008). Distributed Intelligence: Overview of the Field and its Application in Multi-Robot Systems. *Journal of Physical Agents*, *2*(1), 5–14.

Pavone, M., Bisnik, N., Frazzoli, E., & Isler, V. (2007). Decentralized Vehicle Routing in a Stochastic and Dynamic Environment with Customer Impatience. In *1st International Conference on Robot Communication and Coordination* (pp. 1–8). Athens, Greece. https://doi.org/10.4108/ICST.ROBOCOMM2007.2220

Peters, G. (2016, September 20). Are driverless freight trains safe? *Railway Technology*.

Retrieved from http://www.railway-technology.com/features/featureare-driverless-freight-trains-safe-5008616/

Pfleiderer, E. M. (2005). *Relationship of the Aircraft Mix Index With Performance and Objective Workload Evaluation Research Measures and Controllers ' Subjective Complexity Ratings*. Washington, D.C.

Pinedo, M. (2012). *Scheduling*. Springer.

Plott, B. M., Scott-nash, S., Hallbert, B. P., & Sebok, A. L. (1995). Computer Modeling of a Nuclear Power Plant Operating Crew to Aid in Analysis of Crew Size Issues. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1214–1218). San Diego.

Pollard, J. K., Sussman, E. D., & Stearns, M. (1990). *Shipboard Crew Fatigue, Safety and Reduced Manning: Final Report*.

Popkin, S., Gertler, J., & Reinach, S. (2001). *A Preliminary Examination of Railroad Dispatcher Workload, Stress, and Fatigue*. Washington, D.C. Retrieved from http://ntl.bts.gov/lib/12000/12800/12876/ord0108.pdf

Porter, C. O. L. H., Gogus, C. I., & Yu, R. C. (2010). When Does Teamwork Translate Into Improved Team Performance ? A Resource Allocation Perspective. https://doi.org/10.1177/1046496409356319

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, *83*(2), 157.

Rahimi, M., Dessouky, M., Gounaris, I., Placencia, G., & Weidner, M. (2000). *A Task Decomposition Model for Dispatchers in Dynamic Scheduling of Demand Responsive Transit Systems*. Los Angeles.

Rand, W., & Wilensky, U. (2006). *Verification and Validation through Replication: A Case Study Using Axelrod and Hammond's Ethnocentrism Model*. *North American Association for Computational Social and Organization Sciences*. Retrieved from http://ccl.northwestern.edu/papers/naacsos2006.pdf

Rasmussen, J. (1983). Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics*, *13*(3), 257–266.

Roach, G. D., Fletcher, A., & Dawson, D. (2004). A Model to Predict work-related fatigue based on hours of work. *Aviation Space and Environmental Medicine*, *75*(SUPPL.1), 61–69. https://doi.org/10.1016/S1369-8478(01)00015-8

Robinson, S. (1997). Simulation Model Verification and Validation: Increasing the Users' Confidence. In S. Andradottir, K. J. Healy, D. H. Withers, & B. L. Nelson (Eds.), *Winter Simulation Conference* (pp. 53–59). Retrieved from https://www.informs-sim.org/wsc97papers/0053.PDF

Robinson, S. (2004). *Simulation: the practice of model development and use*. *Journal of Simulation*. https://doi.org/10.1057/palgrave.jos.4250031

Rosenhand, H., Roth, E., & Multer, J. (2012). *Cognitive and Collaborative Demands of Freight Conductor Activities Results and Implications of a Cognitive Task Analysis*. Washington, D.C. https://doi.org/10.1177/1071181311551392

Roth, E., Malsch, N., & Multer, J. (2001). *Understanding How Train Dispatchers Manage and Control Trains: Results of a Cognitive Task Analysis*. Washington, D.C.

Roth, E., Malsch, N., Multer, J., & Coplen, M. (1999). Understanding How Railroad Dispatchers Manage and Control Trains: A Cognitive Task Analysis of a Distributed Team Planning Task. In *43rd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1–6). Houston, TX: Human Factors and Ergonomics Society.

Rouse, W. B. (1983). Models of human problem solving: Detection, diagnosis, and compensation for system failures. *Automatica*, *19*(6), 613–625. https://doi.org/10.1016/0005-1098(83)90025-0

Ruff, H., Narayanan, S., & Draper, M. H. (2002). Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. *Presence: Teleoperators and Virtual Environments*, *11*(4), 335–351. https://doi.org/10.1162/105474602760204264

Rusnock, C., Borghetti, B., & McQuaid, I. (2015). Objective-Analytical Measures of Workload – the Third Pillar of Workload Triangulation? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9183*, 124–135. https://doi.org/10.1007/978-3-319-20816-9

Salas, E., Diazgranados, D., & Lazzara, E. H. (2009). Promoting Teamwork When Lives Depend on It: What Matters in the Railroad Industry? In *Transportation Research Circular E-C159: Teamwork in U.S. Railroad Operations* (pp. 10–26). Irvine, CA.

Sargent, R. (2005). Verification and validation of simulation models. *Proceedings of the*

37th Conference on Winter Simulation, (mimic), 130–143. https://doi.org/10.1109/WSC.2000.899697

Schmidt, D. K. (1978). A queuing analysis of the air traffic controller's work load. *IEEE Transactions on Systems, Man and Cybernetics*. https://doi.org/10.1109/TSMC.1978.4310003

Schruben, L. W. (1980). Establishing the credibility of simulations. *Simulation*, *34*(3), 101–105. https://doi.org/10.1177/003754978003400310

Sebok, A., Wickens, C. D., & Laux, L. (2015). *Human-Automation Interaction : Lessons Learned from the Aviation Industry and Recommendations to the Rail Industry for Regulatory Guidance*. Washington, D.C.

Senders, J. W. (1964). The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems. *IEEE Transactions on Human Factors in Electronics*, *HFE-5*(1), 2–5. https://doi.org/10.1109/THFE.1964.231647

Sharples, S., Millen, L., Golightly, D., & Balfe, N. (2010). The impact of automation in rail signalling operations. *Journal of Rail and Rapid Transport*, *225*(F), 179–191. https://doi.org/10.1177/09544097JRRT380

Sheridan, T. B. (1989). Telerobotics. *Automatica*, *25*(4), 487–507. https://doi.org/10.1016/0005-1098(89)90093-9

Sheridan, T. B., & Tulga, M. K. (1978). A model for dynamic allocation of human attention among multiple tasks (Vol. 3, pp. 569–592).

Singleton, W. T. (1989). *The mind at work: Psychological ergonomics. The mind at work: Psychological ergonomics.* https://doi.org/http://dx.doi.org/10.1017/CBO9780511525407

Squire, P., Trafton, G., & Parasuraman, R. (2006). Human control of multiple unmanned vehicles: effects of interface type on execution and task switching times. *… /SIGART Conference on Human- …*, 2–9. https://doi.org/10.1145/1121241.1121248

Steinfeld, A., Fong, T., & Kaber, D. (2006). Common metrics for human-robot interaction. *… -Robot Interaction*, (1), 33–40. https://doi.org/10.1145/1121241.1121249

Sterman, J. D. (2000). *Systems Thinking and Modeling for a Complex World*. *Management* (Vol. 6). https://doi.org/10.1108/13673270210417646

Sukthankar, G., Sycara, K., Giampapa, J. A., & Burnett, C. (2009). Communications for Agent-Based Human Team Support. In *Organizations in Multi-Agent Systems* (pp. 284–312). IGI Global. https://doi.org/10.4018/978-1-60566-256-5.ch012

Sycara, K., & Lewis, M. (2004). Integrating Intelligent Agents into Human Teams. In E. Salas & S. Fiore (Eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance* (pp. 1–36). Washington, D.C.: American Psychological Association. https://doi.org/10.1177/154193120204600342

Tewes, J. (1999). *Model Simulation of the Bulgarian Airspace*.

The MITRE Corporation. (2014). *MITRE Systems Engineering Guide*. Retrieved from www.mitre.org

Thommesen, J., & Andersen, H. B. (2012). *Human Error Probabilities (HEPs) for generic tasks and Performance Shaping Factors (PSFs) selected for railway operations*. Retrieved from http://findit.dtu.dk/en/catalog/2185752564

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *59*(236), 433–460. https://doi.org/10.1093/mind/fz061

U.S. Congress, & Office of Technology Assessment. (1988). The Human Factor in Commercial Aviation. In *Safe Skies for Tomorrow: Aviation Safety in a Competitive Environment* (OTA-SET-38, pp. 117–131). Washington, DC: U.S. Government Printing Office.

Verwey, W. B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. *Ergonomics*, *43*(2), 187–209.

Waag, W. L., & Halcomb, C. G. (1972). Team Size and Decision Rule in the Performance of Simulated Monitoring Teams. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *14*(4), 309–314. https://doi.org/10.1177/001872087201400403

Wang, J., & Lewis, M. (2007). Human control for cooperating robot teams. *Proceeding of the ACM/IEEE International Conference on Human-Robot Interaction - HRI '07*, 9. https://doi.org/10.1145/1228716.1228719

Wang, J., Wang, H., & Lewis, M. (2008). Assessing cooperation in human control of heterogeneous robots. In *3rd ACM/IEEE International Conference on Human-Robot Interaction* (pp. 9–15). Amsterdam, Netherlands: Association of Computing Machinery. https://doi.org/10.1145/1349822.1349825

Wickens, C. D. (1979). Measures of Workload, Stress and Secondary Tasks. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement* (pp. 79–99). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4757-0884-4_6

Wickens, C. D. (2008). Situation Awareness : Review of Mica Endsley ' s 1995 Articles on Situation Awareness Theory and Measurement, *50*(3), 397–403. https://doi.org/10.1518/001872008X288420.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, *25*(2), 227–248. https://doi.org/10.1177/001872088302500209

Wiener, E. L. (1964). The Performance of Multi-Man Monitoring Teams. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *6*(2), 179–184. https://doi.org/10.1177/001872086400600207

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for Mental Workload Measurement in a Test and Evaluation Environment. *Human Factors*, *35*(2), 263–281. https://doi.org/10.1177/001872089303500205

Wilson, K. A., Salas, E., Priest, H. A., & Andrews, D. (2007). Errors in the Heat of Battle: Taking a Closer Look at Shared Cognition Breakdowns Through Teamwork. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(2), 243–256. https://doi.org/10.1518/001872007X312478

Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1037/h0073415

Young, M. S., & Stanton, N. A. (2005). Mental Workload. In *Handbook of Human Factors and Ergonomics Methods* (pp. 391–399). CRC Press LLC.

Yow, A., Walters, B., Plott, B., Laughery, R., & Persensky, J. (2005). Predicting nuclear power-plant operator performance using discrete event simulation. *Cognition, Technology & Work*, *7*(1), 29–35. https://doi.org/10.1007/s10111-004-0167-x

Biography

Victoria Chibuogu Nneji thanks God for this moment. Her story is highly improbable, but with God it has all been possible. Victoria earned her B.S. in applied mathematics (2014) at Columbia University. She returned to Durham and earned her M.Eng. in engineering management (2015) at Duke University and is graduating with her Ph.D. in mechanical engineering-robotics. Her peer-reviewed articles include:

Nneji, V. C., Cummings, M. L., & A. Stimpson. (2019). Predicting Locomotive Crew Performance in Rail Operations with Human and Automation Assistance. *IEEE Transactions on Human-Machine Systems*.

Huang, L., Cummings, M.L., & V.C. Nneji. (2018, Oct). Preliminary Analysis and Simulation of Railroad Dispatcher Workload. *Human Factors & Ergonomics Society 62nd International Annual Meeting*.

Nneji, V. C., Cummings, M. L., Stimpson, A., & K.H. Goodrich. (2018, Jan). Functional Requirements for Remotely Managing Fleets of Personal On-Demand Aircraft. *AIAA Infotech@SciTech Conference*.

Nneji, V. C., Stimpson, A., Cummings, M. L., & K.H. Goodrich. (2017, Jun). Exploring Concepts of Operations for On-Demand Passenger Air Transportation. *AIAA Aviation Technology, Integration, Operations Conference*.

Martelaro, N., Nneji, V. C., Ju, W., & P. Hinds. (2016, Mar). Tell me more: Designing HRI to encourage more trust, disclosure, and companionship. *ACM/IEEE International Conference on Human Robot Interaction*.

Victoria's honors since her B.S. degree include:

- Matthew Isakowitz Fellowship (2019), Commercial Spaceflight Federation
- Best Paper Award (2018), Human Performance Modeling Technical Group, Human Factors & Ergonomics Society
- Women in Aerospace Symposium Scholar (2018), Stanford-MIT-CU Boulder
- Board of Trustees (2018-), American Institute of Aeronautics & Astronautics
- Stanford University Engineering Graduate Summer Research Grant (2015)