# A Process for Improving Transit Service Management During Disruptions

by

Angela Mary Moore

B.S., Urban and Regional Studies
Cornell University (1997)

Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Transportation

at the

Massachusetts Institute of Technology

February 2003



Signature of Author: ______________________________
Department of Civil and Environmental Engineering
September 5, 2002

Certified by: ______________________________
Nigel H.M. Wilson
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by: ______________________________
Oral Buyukozturk
Chairman, Departmental Committee on Graduate Students

# A Process for Improving Transit Service Management During Disruptions

by
Angela Mary Moore

Submitted to the Department of Civil and Environmental Engineering on September 5, 2002 in partial fulfillment of the requirements for the degree of Master of Science in Transportation

## ABSTRACT

This thesis develops a three-part framework and process for managing disrupted transit service. It formalizes the approach to disruption response, disruption response support methods prior to disruption occurrence, and evaluation of policy areas affecting disruption management. The processes are designed to improve the transit provider's ability to avoid and respond to service disruptions, thus improving passengers' experiences. The framework and process can be applied to any transit service disruption case on high or low frequency bus and rail service.

Pre-planning and policy changes were identified as the primary means of improving disruption response. Hold-ins, scheduled bus route blockages, and rail blockages should be entirely pre-planned, while unpredictable disruptions, such as standing vehicles and unplanned bus route blockages, should have response guidelines developed by route segment, direction, and time of day.

Additionally, agencies cannot effectively respond to disruptions on low frequency service, so it must be protected and scheduled service replaced as soon as possible. To limit passenger impacts during disruptions, service managers must be able to respond quickly by knowing as much information as possible about the route in advance and having a limited set of promising responses from which to select a response. In addition to having prepared service managers, agencies can assign reliable vehicles to essential runs, develop higher pre-determined headway and operations plan to accommodate common disruptions on high frequency routes, and develop a list of candidate runs to borrow during disruptions, which could be supplementary runs in a base schedule where the remaining service provides enough capacity.

The processes were applied to the Chicago Transit Authority (CTA) Route 77 Belmont. While there are several disruption response options available under ideal conditions for a standing bus disruption, CTA practice and operational constraints make most responses infeasible, leading them to either add service or do nothing during the initially disrupted trip. Assuming service managers can intercept and instruct operators in time, the ideal responses are available for the remainder of the disrupted trips. The primary agency practices and system elements constraining CTA bus disruption response are the communications system, vehicle availability, and operator and service management training.

Thesis Supervisor: Nigel H.M. Wilson, Ph.D.
Title: Professor of Civil and Environmental Engineering

## Acknowledgments

First and foremost, thanks to my advisor, Nigel H.M. Wilson, who always understood the conceptual problems and the proper operations management approach before I did, was patient with my writing, supported my research trips to CTA, and was always confident that I was on the verge of finishing. Thanks also to Fred Salvucci and Mikel Murga who helped me with my approach to the case studies, kept me focused on passenger concerns, always saw the value of this research, and did not give up on me finishing the thesis. Your many insights have made this a better thesis.

Thanks to the Tren Urbano-UPR/MIT Professional Development Program and CTA-MIT-UIC Joint Research Collaborative for funding my Masters degree and research. My CTA summer internship at CTA and my sustained contact with the helpful operations and planning staff was essential to this thesis. Thanks to the CTA Planning and Development and Operations staff, Daniel Shurz, Robert Smith, Mersija Besic, and Hugh Muller who made my research possible, answered my questions, provided operating data and insight on operations problems, as well as made my summer in Chicago very enjoyable. Thanks also to Jeff Sriver and Jessica Vargas for organizing our visits and helping us find people in the agency to advance our research.

Thanks to CTA General Manager Bill Platt, Transportation Managers Eugene Thurmond and Walter Thomas, and Supervisors Kevin Majors and Fred Young who opened my eyes to the world of service management, the problems and challenges, and the creativeness needed to make field management work. A special thanks to the 77th Garage radio shop foreman and radioman staff for teaching me about the MDT units, doing all of the garage disk swapping, and helping me in the field from morning pull-out to late at night. Your help was essential to my data collection efforts. Thanks to Adam Rahbee for helping me focus and bring all the disparate pieces together.

Thanks also to my fellow CTA research students. David Barker provided C/PC data and answered all of my communications systems questions. Thanks for finishing on-time so I could cite your communications work. Thanks to Yoosun Hong, who provided the 77 Belmont running time and boardings by trip data. Cordy Crockett entertained my transit connectivity questions and discussed transit management ideas with me.

In addition to the technical support I received, my family and friends played an essential role in my work at MIT and this thesis. A tremendous thanks to my family -- the most important people in my life – Mom and Paul, Dad and Merilyn, and Meghan, Kelly, and Laura. Thanks for listening patiently while I explained my thesis topic and why I hadn't finished yet. You have made my life so much fuller and have always supported me.

Thanks to my friends – Erik Wile, Pilar Rodriguez, Meredith Coley, Dan Morgan, Mark Schofield, Ali Davis, Frances Switkes, Ryan Tam, and Margaret Cortes – for all the fun and support over the past two tremendous years. Thanks to Mark Hayden for kicking my butt to get into graduate school, warning me about the pitfalls of graduate school, and encouraging me during my thesis writing.

## Table of Contents

## List of Tables

## List of Figures

# Chapter 1: Introduction

This research develops a framework and process for transit agencies to evaluate and improve their response to service disruptions. The framework suggests a methodology for how service managers and agency policymakers should think about service restoration, the information service managers need to make better decisions, the criteria used to select among feasible disruption responses, and how the pre-planning and management structure needs to be organized so that reasonable options are available to managers when crises arise. The process outlines the steps that should be taken to evaluate options and arrive at the desired agency balance between service quality and cost. The frameworks and processes are applied to the Chicago Transit Authority (CTA) bus system.

## 1.1. Motivation

Managing disrupted transit service properly is critical for transit agencies because this is when passengers have the greatest potential to experience long delays and vehicle crowding problems. In addition, disruptions will never be completely avoidable, regardless of the agency's best efforts to properly schedule operators, maintain its vehicles and right-of-way, and respond to street closures. Operators will take unanticipated sick leave, vehicles and right-of-way equipment will break down while in service, and route blockages will occur. Once the agency has done its best to reduce the incidence of disruptions, it must have plans to respond to disruptions to minimize their negative impacts.

There are several critical ways in which agencies may mismanage service disruptions. The first is that service managers may spend all of their time responding to the problem itself, instead of managing the remaining service. For example, they may focus on getting a vehicle back into service or investigating the accident, while ignoring the effectiveness of the remaining service and the resulting passenger experience. This may lead to a minimal or even non-existent service recovery response, allowing the route or line to operate with a large gap until the missing vehicle would have pulled-in or another vehicle is available.

A second problem may be that service managers misunderstand key service restoration concepts or lack sufficient information to evaluate the options and choose among them. These include misunderstanding service goals, the expected impacts of different disruptions, feasible service restoration responses, the need to protect critical service, or not knowing the typical operating characteristics of the affected route. The first part, misunderstanding service restoration concepts, is more critical to disruption response because it precludes even knowing the proper operations goals or right approach to achieve those goals, whereas not having enough information can be managed even if at some uncertainty and loss of effectiveness.

A third problem is the lack of feasible restoration options when disruptions occur and options diminish as time passes. This means that even if the service manager can identify the best service restoration approach quickly, it may be infeasible given the current situation. While the specific situation dictates which options are ideally available, the available information and communications options are also likely to preclude many ideally feasible options. Contingency planning diminishes the time needed to develop and implement a response.

A final problem with service management's disruption response process is a lack of means for evaluating the benefits of their actions. The processes developed in this research are based on evaluation criteria that serve the passengers' primary concerns, yet are simple enough to be estimated in real-time. The ease of estimation allows agencies to close the feedback loop. Thus, once the agency has applied the process and evaluated how it currently does and how it should manage different disruption responses, it will serve to validate current techniques or suggest new approaches to the problem. It can also serve as a springboard for suggesting areas of concern needing additional analysis. Lastly, adopting a single process for determining the proper disruption response should lead service managers to respond more quickly and uniformly to disruptions, particularly in agencies with many inexperienced service managers.

## 1.2. Objectives

This research has three primary objectives. The first is to structure an approach to the problem of transit service management during a disruption. This is accomplished by developing a framework and process for how service managers should approach disruption management, including how to define the problem, identify the available options to mitigate passenger impacts, and outline the information service managers will need to choose among them. This approach provides service managers and agency policymakers with an understanding of the problem and a uniform process with which to respond to disruptions.

The second is to develop appropriate evaluation criteria, or passenger impact and agency cost indicators, that properly value the primary passenger concerns and weigh agency cost during a service disruption, without overwhelming service managers with too many indicators.

The last is to define and evaluate policy level issues that can either assist the service manager in mitigating negative passenger impacts during disruptions or reduce the number of disruptions occurring. Agencies can use the policy evaluation process to assess the cost-effectiveness of changing different agency policies that exclude otherwise feasible service restoration options or affect the frequency of disruptions. Changing agency policies is a very powerful tool because of the wide-reaching effects of policies, so it is important to determine whether specific policy changes are worthwhile. This process helps agencies understand how their policies affect the agency's ability to respond to service disruptions and what the costs and benefits of policy changes would be. Additionally, service managers can improve the agency's response to these disruptions by understanding which disruptions are amenable to mitigation or avoidance and how they should be managed.

## 1.3. Service Management Overview

Service management is the process of, and staff involved in, ensuring that service operates safely and according to the intent of the operations plan. Service managers manage all operators and equipment involved in service delivery, respond to vehicle and operator problems, and balance the needs of passengers, the agency, and operators. Most agencies have both centrally- based staff and field-based staff to direct operations. Historically, agencies have placed most service management decisions in the hands of the field-based supervisors, while centrally-based dispatchers have been responsible for radio operations. Centrally-based managers (typically called dispatchers or controllers) are primarily responsible for monitoring and managing messages on the radio and other communication systems. Field-based supervisors work directly with operators to maintain and restore service. Service management can be implemented in a variety of ways, with responsibilities distributed across these two primary positions (Levinson, 1991).

The framework and process developed in this research is not designed with a specific service management approach in mind and does not rely on having a manager centrally located or available in the field. Rather it refers to a single entity, a service manager, who is responsible for the entire response process. The service manager gathers information, develops a disruption response plan and implements the plan through instructing the operators.

Another aspect of service management that differs among agencies is the availability and use of automatically collected passenger demand and vehicle location information. The primary systems that provide this information are automatic vehicle location systems (AVL), which provides vehicle location information, and automatic passenger counters (APC) or automatic fare collection (AFC), which provides passenger demand and load information. This information can be provided in real-time or saved in a database to be reviewed at a future date. These technologies can improve disruption response because of the reduction of uncertainty in knowing where vehicles are and how full they are.

Real-time information provides service managers with current information, from which they can make decisions. It also allows them to implement options that would be too risky without accurate vehicle location or passenger demand information. If this is unavailable, electronically stored historic vehicle location and passenger data is an excellent tool for understanding how the service normally operates and disruption responses should be developed from this information in the absence of real-time information. Without this wealth of automatically collected and electronically available information, service managers have relied on manual point or ride checks that record vehicle locations and passenger loads, as well as their own understanding of how the route operates. While improved data quality will assist service managers in making better decisions for passengers, the framework and process developed in this research can be applied within the constraints of the information available. However agencies collect their information, the framework and process is not constrained by the availability of information from technology. The less information available, the more estimates service managers are required to make, which should lead them to a more conservative approach to disruption response.

## 1.4. Prior Approaches

While other studies have touched on areas of this research, there has not been a comprehensive evaluation of service management during disruptions or thought given to how to create a situation where the best disruption response is available to service managers and how they should select among service restoration techniques in a given disruption situation. The literature has focused on three areas: how to identify and address problems in rail transit service, service management studies, rail and bus operations control strategies, and setting policies, such as extraboard and spare ratio size. The first two topics are discussed below and operations control strategies are discussed in Chapter 2. Extraboard and spare ratio sizing are not discussed extensively.

### 1.4.1. Rail Systems Analysis

Rahbee (2001) developed a process to identify and understand recurring problems in rail systems. The study describes three areas of concern: the line characteristics, operations plan and service management. While the framework was developed with rail service in mind, these analysis areas work for both bus and rail. This research modifies these areas to adapt them to bus service and disruption response applications.

### 1.4.2. Service Management Studies

Levinson (1991) documented service management practices, reviewed service restoration literature, and developed recommendations for improving service management. This study focused specifically on the job of the typical supervisor and outlined responses for disruptions and other problematic situations. While it provides an excellent overview of response strategies, it does not identify which supervision strategies best serve different situations. It also lists events that can lead to disruptions, such as inclement weather, traffic congestion, and fires, but does not categorize disruption types or information service managers need to select the proper service restoration technique.

Froloff, Rizzi, and Saporito (1989) developed a comprehensive bus service management manual that described service goals, basic service restoration techniques, and route-level service management. It was developed on the premise that while models exist to describe how transit service is assumed to operate under different conditions, the models have moved forward without a solid base in service management theory. This created two problems. The first is route managers learning to use new technology to manage service without any formal criteria for what is good service, and the second is it impedes the design of future systems because the industry does not know what it should be measuring. It describes a decision context for service management, which lists the primary factors affecting service management and discusses the specifics of how a route manager would address generic service problems, such as gaps, bunching, and delays. Last it develops a training simulator for route managers.

Barker (2002) focused on the distribution of service management responsibilities between centrally- and field-based decision-makers and where different types of decisions should be made. This study concludes that accident and equipment failure responses should be coordinated centrally through dispatchers, because other parties, such as tow or mobile repair trucks, may need to be dispatched to the scene. Service management and restoration decisions, however, should be made by field-based supervisors because they will always have a better understanding of current conditions and can work more effectively with operators.

### 1.4.3. Literature Gaps

Topics of interest that were not adequately covered in the literature were valuing passenger impacts during disruptions, induced agency impacts created by using operators for extra trips and extended periods of overtime, and an assessment of when to use different disruption response options. There have been studies on how passengers value their time waiting for service vs. in-vehicle travel time, but there have not been studies showing how passengers value expected vs. unexpected wait time or other impacts. For example, perhaps the first half headway of waiting is not terribly irritating, since passengers anticipated waiting at least that long. Waiting between a half headway and one headway is a little more anxious, since the vehicle should arrive any time now. Waiting longer than a headway may be really irritating because now the passenger has no idea if the vehicle is a little late, very late, or if the service is severely disrupted. Other impacts not reported in the literature are the irritation of being passed by a full, or empty, bus or being forced to alight a bus in the middle of the route and catch the following vehicle.

Secondly, the eventual agency impacts of having a culture of lots of overtime or a policy of forced overtime have not been explored. This information would help agencies understand how operators respond after constantly working overtime and the agency impacts of those responses.

Lastly, operations control strategy studies use evaluation measures to compare the effectiveness of the operations control strategies considered for their selected cases, but they are not the proper measures for a broader assessment of managing service disruptions. For example, most operations control studies have focused on disruption response from the moment the disruption occurs through the end of the trip. While they have focused on understanding the impacts of specific operations control strategies, the proper way to implement them, and how to select among control strategies during a given situation, an overall evaluation of available options and general situations in which they are useful has not been made.

## 1.5. Methodology

The framework and process developed in this research are designed to guide agencies' disruption response practices and provide them with a means of evaluating the impacts of their policies on service restoration effectiveness. First, a framework is developed; then, a process is outlined and applied in a specific context.

### 1.5.1. Framework Development

The framework describes how service managers could approach real-time disruption management and the evaluation criteria used to select among service restoration options, identifies disruptions that could be mitigated or avoided, and describes how agency policymakers could review policies that impact disruption response. It introduces both the general types information and specific knowledge needed to make disruption response decisions, service restoration concepts, and operations goals. It also describes how service managers can support real-time disruption response in advance and discusses how changing agency policies can benefit service restoration, which are concepts used in the disruption response processes.

### 1.5.2. Process Development

Once the framework of how to think about the problem has been established, a process that applies the framework findings is developed. It outlines the steps service managers should take to identify and evaluate service restoration options, respond to disruptions, and monitor the effectiveness of the response, as well as the steps agency policymakers can take to evaluate the effectiveness of agency policies which can affect the incidence of disruptions or the possible responses.

### 1.5.3. Process Application

Lastly, the service management process is applied to the CTA bus system, to illustrate how the real-time disruption response processes can be used and show the benefits of using the process compared with a the existing disruption response process. A case applying the agency policy evaluation process is also presented.

## 1.6. Thesis Organization

Chapter 2 provides an overview of operations management, introducing the elements of an operations plan, disruption types, and restoration techniques. Chapter 3 develops disruption response and agency policy evaluation frameworks, which outline how service managers and agency policymakers should approach disruption response and evaluate agency policies influencing service restoration, as well as describing the evaluation criteria used to select among service restoration options. Chapter 4 presents the disruption response processes for service managers and agency policymakers to follow to reduce the number of disruptions and negative passenger impacts during disruptions. Chapter 5 applies the processes developed in Chapter 4 to the CTA bus system. Chapter 6 summarizes the concepts, process and major findings, and suggests areas for further research.

# Chapter 2: Fundamentals of Operations Management

In order to understand the specifics of service management during disruptions, it is important to have a basic understanding of operations and service management. This chapter reviews the basic concepts of operations management, including the operations plan, disruption types, and service restoration techniques needed to understand the framework developed in Chapter 3. It also summarizes prior operations control strategy studies for both bus and rail systems.

## 2.1. Operations Plan

Understanding the components and purpose of an operations plan, as well as how it functions is essential to understanding service management during disruptions. The operations plan, typically developed by the service planning and scheduling departments, is the ideal service plan crafted in advance. Broadly, it is all the scheduled service planned for the service period. Ideally, it is what the agency has determined best serves passengers, given existing service policies, cost and work rule constraints. Then, it is the responsibility of the operations department to implement it: every day operators, service managers, mechanics and other operations staff strive to meet the operations plan.

### Service Components

Before describing the components of the operations plan, one must understand the elements needed to provide service. There are two required elements of service, the operator and the vehicle (Froloff, Rizzi, and Saporito, 1989). Each operator is assigned to one run per day, which is a daily assignment of work that complies with all work rules. Vehicles are assigned to vehicle blocks, which describe the operations of each vehicle from scheduled pull-out to scheduled pull-in. Vehicle blocks are much less constrained than operator runs because they are not subject to work rules. Throughout the day several runs may operate different pieces of a block.

Agencies are typically required to provide operators with a break sometime during their 8-hour shift. To give operators a break, the operations plan includes reliefs, which is where one operator starting his shift or finishing his break takes over for, or relieves, the one who just drove up who ends his shift or takes a break. Reliefs can occur at the start or end of work shifts or to give operators a mid-shift break, and can be made at a terminal or mid-route. This allows agencies to use operators and vehicles more efficiently. Having operators start and end their shifts and take breaks on the route reduces time spent ferrying vehicles to and from the garage.

### Operations Plans Components

There are several components of an operations plan, which is typically built at a route level and then integrated into a system of service for service coordination and ease of transfers. Route timetables are developed from the cycle time and desired headway, which jointly determine the number of vehicles necessary to operate the plan. Cycle time is the sum of running time, the time needed to make a round trip, and recovery time, extra time taken at the terminals which allows operators to take a short break and start their return trip on time. Recovery time is the time between when the trip is scheduled to end and when its next trip is scheduled to begin. This time is provided to increase both service reliability and operator comfort. Without sufficient running and recovery time, late trips will affect future trips and the operator will fall farther and farther behind schedule.

The headway is the primary service decision because this is the service level provided to passengers. This drives the number of vehicles needed because the route length and speed along the route (which are set once the route is developed) determine the running time. Headways can be set in a two ways. One is a policy headway determined by the agency, which is a statement of minimum service level that passengers can expect. Policy headways assume there is enough capacity on each vehicle to serve all passengers with the first arriving vehicle. The second is a demand-based headway, which is determined from the route's passenger demand and the agency's statement of a maximum number of passengers per vehicle.

### Operations Plan References

Lastly, operators and service managers use printed references with operations plan information. The first is the schedule book or supervisor guide, which include standard schedule information provided to the public, such as timepoints and specific times indicating where vehicles should be when, as well as relief, run and block information. The second is a set of paddles; a paddle is a run synopsis and there is a paddle for each run. Paddles indicate when and where the run starts, its timepoints, recovery time, relief times and locations, which other run is making or getting relieved, and when and where the run ends. Typically, they also provide a summary of pay hours.

## 2.2. Service Disruptions

Every day as the operations plan is implemented, there are many events that can make it infeasible, ranging from traffic congestion and uneven passenger loads to equipment failures and not having enough operators report to work. This section introduces common service problems and discusses the various types of disruptions. Disruptions are categorized by their impacts on service. The three general disruption categories are service gaps, insufficient capacity, and route blockages. The various sources of each of these disruptions are described in greater detail below.

### 2.2.1. Common Service Problems

In this section, common service problems are described, distinguishing between degraded and disrupted service. Service degradation results from off-schedule vehicles or minor operating irregularities caused by vehicles not being scheduled at proper intervals. Controllable sources of service degradation include vehicles leaving the garage (or terminal) late (or early), inaccurate scheduled running time, operator inability to maintain the schedule, and passenger boarding procedures. Service degradation sources typically outside the transit agency's immediate control include variable traffic conditions and passenger boarding demand. Service degradation is of greatest concern on high frequency routes where headway and passenger load distributions can rapidly deteriorate in the absence of close monitoring and effective operations supervision.

Disrupted service is distinct from degraded service because it results from service failures that make meeting the operations plan infeasible. The operations plan becomes unattainable when there is a missing vehicle, increased running time such that the new running time is greater than the cycle time, insufficient capacity to serve all passengers, or a route blockage. Disruptions can also be caused by severe passenger demand spikes, service hold-ins, accidents, mechanical or other vehicle failures, severe weather, severe traffic conditions, construction, emergency services, or other events or activities that create route blockages.

Disruption management is the process of managing the actual problem (if there is a specific event causing the disruption), as well as managing the remaining service until the disruption has concluded and the agency is able to meet the original operations plan. This research focuses entirely on managing transit service during disruptions in an effort to restore normal service. Because of this focus, "service restoration" will be used to mean interventions in the face of disrupted service and "service management" is a more general term describing how service is managed or the process of managing service. Service management also includes route management interventions when service is degraded, but not disrupted.

### 2.2.2. Service Gap

Service gaps are created either when the running time increases to more than the scheduled cycle time or when there is a missing vehicle. Both make the operations plan and headway unattainable.

#### Insufficient Running Time

Insufficient running time disruptions occur when vehicles do not have enough running time to complete their current trip before the start of their next trip. This can result from inclement weather, construction delays, traffic congestion, and severe demand spikes, as well as other disruptions, such as route blockages. Scheduling a sufficient amount of cycle time will help reduce the likelihood of this type of disruption occurring.

#### Missing Vehicle

A service gap can also be created by a missing vehicle, which results from hold-ins or standing vehicles. Hold-ins are created by an insufficient number of operators and/or vehicles to operate the schedule. Standing vehicles are created from in-service vehicle failures due to mechanical problems, accidents, or missed or late reliefs. Train failures create route blockage disruptions, as well, because trains cannot pass on the same track.

### 2.2.3. Insufficient Capacity

The second type of disruption results from not having enough capacity to serve all passengers. This often creates problems for both elements that determine the feasible headway: the number of vehicles and cycle time. If there is insufficient capacity, the number of vehicles with the current cycle time is insufficient to carry the passengers waiting for service. Assuming the vehicles cannot speed up, which is usually the case, the route will need more vehicles to serve all passengers in a timely manner. While the need for vehicles is clear, this type of disruption will likely also increase the necessary running time because of the additional dwell-time delay created from passengers boarding and alighting the crush-loaded vehicles. This impact is more often associated with bus service than rail service because marginal and average boarding times per passenger are much higher for bus than for rail.

Passenger demand spikes vary in location and length of delay. Severe passenger demand spikes can be created by special events, disruptions on parallel or connecting high capacity service, such as trunk bus lines or heavy rail systems, and inclement weather if travelers switch from other modes to transit. The least disruptive case is when there is only one location with an increase in demand and it occurs only briefly. It is more difficult to respond to cases with a sustained increase in passenger demand along the route.

### 2.2.4. Route Blockage

The last type of disruption is a route blockage, which requires re-routing service around the blockage. Route blockages also typically create other disruptions, such as service gaps and insufficient capacity. Often when service is first reestablished, there is a large pent-up passenger demand because service has been missing for several headways. In addition, the new routing typically takes longer than the regular routing, which may create an insufficient running time disruption that needs to be managed.

Service blockages manifest themselves differently for bus and rail because of their infrastructure. One major difference between rail and bus transit operations is that while buses can generally pass each other during both normal and disrupted operations, trains cannot ever overtake each other on the same track. It would be extremely rare for a disabled bus to be the cause of a bus route blockage, as the street network offers many other options for the following buses and other traffic. However, when trains fail in-service they create a route blockage and service must be managed around the blockage.

Rail blockages can result from a mechanical failure of the right-of-way systems, such as the track or signalling systems or from other disruptions, such as a vehicle failure that strands a standing vehicle in the middle of the right-of-way. Rail typically has fewer reroute options than bus, but while the infrastructure is quite complex, the agency generally retains control of its infrastructure and typically owns, maintains, and operates its exclusive right-of-way, which includes all tracks and signaling equipment. Thus, the agency can be proactive about reducing track blockages through track and rolling stock maintenance programs.

Bus right-of-way includes the street network available to buses. While bus right-of-way is much simpler than rail systems and is mostly maintained by other agencies, it is typically shared with many other users and bordered by many private and public uses. Users are competing for the same space and the transit agency does not control the right-of-way. Many blockages are not created or managed by the transit agency; however, the agency can and should coordinate closely with other involved agencies. It may not be able to reduce the number of route blockage disruptions, but it can improve its response.

## 2.3. Managing High Frequency and Low Frequency Service

Service management for high frequency and low frequency service is fundamentally different. While the passenger satisfaction goals are the same, how service is provided to achieve them is quite different. This is because of the different passenger arrival patterns and the severity of passenger delay when a vehicle on a low frequency route is missing.

For high frequency service, defined as service with headways at or less than every 10 minutes, most passengers arrive randomly (Turnquist and Blume, 1980). They expect to have to wait at most one headway and, on average, half the headway. Thus, providing even headways minimizes the average wait time. If there is one missing vehicle without any service management, a gap of two headways, or, up to 20 minutes, is created and passengers are assumed to wait 10 minutes, on average.

For low frequency service (i.e. routes with headways greater than 10 minutes), particularly in the evenings and weekends when passengers understand service is less frequent, most knowledgeable passengers will time their arrival based on the printed schedule. This is a substantial behavior difference compared with passengers using high frequency service. Their arrival behavior, while intended to minimize wait time under normal operating circumstances, will make them wait at least an entire headway for the following if a bus is missing as a result of a disruption.

Waiting an additional headway is problematic because the scheduled headway is already quite high. For example, if the service is scheduled every 15 minutes, passengers just missing the previous bus will wait 30 minutes when they expected their wait to be less than 5 minutes and at most 15 minutes. Additionally, the following bus will be carrying twice the normal load, which will likely also delay it. Thus, as service is less frequent, each vehicle becomes more critical to successful operations. Instead of reducing passenger wait time. spacing vehicles to reduce the gap size could actually increase passenger wait time and could be more disruptive to passengers. Therefore. service managers cannot simply reduce the passenger impacts of a service gap disruption on low frequency routes and should generally not respace the headways.

## 2.4. Service Restoration Techniques

This section describes the principal service restoration techniques available to service managers. The primary service restoration goal is to minimize negative passenger impacts. If the resulting headway is intolerable or the disrupted service is low frequency. the agency must deploy additional service. As discussed previously. disruptions make the operations plan unattainable. which results in compensating by increasing the headway. if the cycle time and number of vehicles do not change. The cycle time cannot be modified because it is set by the right-of-way or traffic conditions.

The first option to consider is adding a vehicle to the operations plan. in order to replace missing service or provide more vehicles to maintain the scheduled headway. the second is to borrow resources from other routes. and the third is to manage within the existing resources on the route. The disruption type and nature determine which options are available to service managers. as well as their cost and benefit.

### 2.4.1. Adding Service

Service managers may respond to the disruption by adding, or replacing. service. Sources of new service are an extra vehicle already stationed in the field or on a rail siding (a "gap" bus or train), a pull-in operator and/or vehicle that has finished his run and is returning to the garage or rail yard. or an extra pull-out from the garage or rail yard. While the benefits of providing additional service with any of these options are similar. the cost of each source is different. Once scheduled, gap vehicles and operators are a fixed cost. Using them for a disruption does not create additional cost. The only constraint is that service managers should try to use them to mitigate the disruption creating the greatest passenger impact. Pull-in vehicles and extra pull-outs cost the agency the operator's hourly or overtime wage. Additionally, pull-in operators may have work rule constraints that preclude them from working additional hours.

### 2.4.2. Reallocating Service

A second option is to reallocate service from one route to another. This requires the route from which the service was taken, which is now disrupted, to be managed. While moving service from another route provides extra service on the disrupted route, it does not add more service to the system and passenger impacts and agency cost must be carefully considered across both routes.

### 2.4.3. Managing Service

The last option is to manage the remaining service, which is usually necessary even if additional or reallocated service is added to properly incorporate the new service. Service managers can modify the route's operations plan by applying operations control strategies (service restoration techniques) to satisfy the goal of minimizing negative passenger impacts.

The three basic control strategies for managing in-route service are expressing, holding, and short-turning. Expressing is any action that increases the vehicle operating speed. It includes several variations such as on- or off-route repositioning, limited stop service, drop-off only service, and leapfrogging vehicles. Holding is the slowing or temporary stopping of a vehicle. Short-turning is only done while in-route and involves taking a vehicle traveling in one direction and putting it in-service in the opposite direction without completing its current trip. Short-turning can only be used if the passenger demand of the portion of the route served is much greater than the demand in both directions over the skipped portion of the route, thus warranting moving one vehicle's worth of capacity from the segment of the route to the trunk. Additionally, the portion of the route losing a vehicle must be managed to the extent possible if a service gap results. In addition to having dominant passenger directions during the peaks, rail also develops capacity problems more quickly than bus when there is a service blockage, which may prompt the agency to operate single track service around the blockage.

From the route terminal, the service manager can reschedule the trips leaving the terminal (respacing), thus respacing them with the new headway in mind. This generally works well for high-frequency service, when providing even headways is the proper operations goal. This involves having vehicles leave early (moving up) or late (spacing back) from their scheduled dispatch time. Another option is to use short-turning or deadheading to redistribute service from the ends of the route to provide the capacity needed on the route's trunk section.

In-route interventions can be used to manage both short- and long-term disruptions, including degraded service and unexpected disruptions (such as service gaps) and maintaining service under a new operating plan. Terminal adjustments move service to a modified operations plan and are made for anticipated or longer duration disruptions.

Lastly, there are several options for re-routing rail or bus service. Rail service restoration options include waiting for the disruption to clear, using single track operations, setting up a rail shuttle, or operating a bus bridge. Buses have more right-of-way options and can almost always re-route around the blockage.

## 2.5. Operations Control Literature

### 2.5.1. General

This section reviews studies whose findings apply to both bus and rail service. One of the main differences between bus and rail control strategies is the ability of vehicles to pass each other. Early studies established the motivation and primary objectives of transit vehicle control and generally assumed only one or two vehicles that did not bunch and, therefore, were not faced with the decision to have the following vehicle pass its leader. Studies focusing on passenger-based objectives, such as reducing passenger delay time and those that do not encounter the decision to pass or not to pass apply to both bus and rail. Later studies explored specific control strategies that could be used in specific contexts.

Abkowitz (1978) identified three areas in which to improve service reliability: operational, priority, and control. Operational methods are long-term schedule changes, training efforts, and route restructuring. Priority methods are intended to provide more control over bus travel time and reduce travel speed variability due to traffic resulting from their shared right-of-way. Examples are signalized intersection priority and roadway design, such as exclusive and queue jump lanes. Lastly, control methods are interventions made by service managers in real-time, given the current conditions, passenger loads and demand, and vehicle locations.

Welding (1957) was the first to identify that the primary way to reduce average passenger wait times where passengers arrive randomly was to reduce vehicle arrival variance. Holroyd and Scraggs (1966) studied bus arrivals and passenger wait times in central London to find the relationship between passenger wait times and bus arrival times. Barnett (1974) used a simple single control point system to explore the benefits of holding vehicles to a predetermined threshold headway to achieve more even headways. He found the optimal holding location was just before the highest loading point or when the number of passengers on the bus is less than the downstream passengers. Barnett (1978) developed a case with one infinite capacity vehicle and found different optimal holding strategies under different passenger cost structures. Turnquist and Blume (1980) found under what conditions of headway variation and percent of passengers delayed that holding strategies would reduce passenger wait times.

Later studies evaluated the effectiveness of stop skipping and short-turning with findings that can be applied in general to transit operations. Deckoff (1990) developed guidelines to determine under which conditions field-based inspectors should choose to short-turn trains on the Massachusetts Bay Transit Authority (MBTA) Green Line, a light rail system with four branches that converge into the downtown trunk subway section. The field-based inspector making the decision has limited vehicle location, passenger load and passenger demand information, so Deckoff developed a set of short-turning guidelines based on the two previous headways of each train considered for short-turning. He then compared the benefits of short-turning under these conditions with the benefits if the inspector had accurate automatic vehicle location (AVL) information which would calculate passenger delay statistics directly and make short-turning decisions. In general, if there were limited information available, he recommended short-turning trains if both were bunched, if there was a long headway followed by two trains or if there was a long headway and the second train is not in sight. While his guidelines developed for field-based decision-makers with limited information reduced passenger dumping ratios and delay compared with the then current decision-making process, passenger delay could be reduced further if the inspector had AVL information.

Wilson et al (1992) studied the use of a broader range of operations control strategies on the MBTA Green Line. They found that a lack of vehicle location, vehicle passenger load, and passenger demand information reduced the field-based inspectors' abilities to select and implement control strategies properly. They also found that the most effective use of stop skipping was for vehicles with a long preceding and short following headway and high passenger demand beyond the skipping segment.

### 2.5.2. Bus

This section describes the literature on control strategies used in bus applications. These studies frame how service restoration techniques are used in the disruption response framework and process. Each explores the effectiveness of using a particular service restoration technique in a given degraded service circumstance, but does not develop examples showing the effectiveness of the technique when used to respond to a service disruption.

Koffman (1978) developed a simulation model to evaluate the effectiveness of holding, skip stop expressing, short-turning, signal priority, and even spacing from the terminal. He found that holding, expressing, or short-turning if used without considering the passenger loads and demand could worsen passenger impacts, rather than improve their service. Signal priority and dispatching vehicles more evenly both improved passenger experience. Turnquist (1978) found that a "prefol" headway strategy, or making control decisions based on a vehicle's position with respect to its leader and follower, reduced passenger delay better than making control decisions based solely on the vehicle's leader.

Abkowitz and Engelstein (1984) and Abkowitz, Eiger, and Engelstein (1986) used simulation to find holding threshold headway values, develop a threshold-based holding control strategy, and determine the desired holding control point. They found that running time variation propagates along a route and the location of the optimal control location and threshold value depended on the route's passenger demand and loading profile. Abkowitz and Tozzi (1986) demonstrated the effects of boarding/alighting profile on the effectiveness of headway-based control.

Abkowitz and Lepofsky (1990) implemented headway-based holding on the MBTA Route 1, with service along Massachusetts Avenue between Dudley and Harvard Squares, and found that selection of the control point depends on the boarding/alighting profile. Specifically, the control point must be where on-board delays due to holding actions are not too burdensome and yet early enough that passengers waiting downstream can still benefit from more uniform headways. The terminal is the optimal single control point only when both the variation in dispatching headway and the terminal boarding demand are high (Koffman 1978; Turnquist and Blume 1980; Eberlein 1995).

Fu and Yang (2001) developed a simulation model to describe benefits of regularizing headways with preceding and following headway information for the selected control point. They found the optimal location of the control point to be at a stop with high boarding demand near the middle of the route. They also found diminishing returns when adding control points and recommended using two control points because it improved significantly over a single control point and obtained most of the benefits in the extreme case of being able to hold vehicles at every stop. They used a threshold headway for holding control and found that even a small threshold headway realized most of the waiting time reduction benefits. Lastly, they looked at the benefits of real time information, which would allow headway management based on both its leading and following headways and found that while it would reduce in-vehicle travel time, it required a number of control points and there were limited benefits to passenger wait times.

### 2.5.3. Rail

Eberlein (1995), O'Dell (1997), Shen (2000), Ortiz (2000) and Puong (2001) have all formulated analytical models with the primary objective being to reduce passenger waiting time. Shen (2000), Ortiz (2000), and Puong (2001) consider in-station waiting time, in-vehicle travel time, and extra wait time for passengers passed up due to train capacity constraints. They also used weighted values for in-station waiting time because passengers value this time greater than that of travel time once in the vehicle.

Eberlein (1995) evaluated the effectiveness of different operations control strategies. Eberlein et al (1997) considered the case of deadheading an empty train from the terminal to put the trip back into service along the line to shorten a service gap. They found that this only works well if passenger demand is low at the terminal and the skipped stations. It creates more passenger delay when implemented on lines with heavy initial boardings.

O'Dell (1997), Shen (2000), and Puong (2001) evaluated the use of operations control strategies to respond to rail blockage disruptions. O'Dell (1997) showed that passenger delay is reduced to a greater extent when controlling trains ahead of the blockage, rather than those behind the blockage.

While prior research sought to minimize passenger impacts only through the completion of the trip in the direction of the disruption, Puong (2001) developed a model that considered the passenger impacts for one round trip. It included terminal operation constraints at the beginning of the trip with the disruption, the terminal at the end of the trip with the disruption, and the return trip. This is an important improvement when considering the broader, downstream impacts of control strategies. It recognizes that trains cannot be dispatched in the return direction if they have not arrived at the terminal and spent the minimum necessary time to turn around, and that trains cannot get into terminals, if the platforms are already occupied with trains.

One study evaluated the passenger impacts of providing schedule- vs. headway-based performance measures for high frequency rail service. Ortiz (2000) reviewed the schedule adherence- and missed trip- based penalty clauses in San Juan, Puerto Rico's rail line, Tren Urbano, contract with the contracted operator and compared the operator's induced behavior to the desired passenger-based behavior of headway-based service management. She found that schedule-based performance incentives would lead the transit operator to implement operations control in a way that would be better than doing nothing, but would not provide the optimal situation for passengers, particularly during long disruptions.

An important consideration for operations control in rail systems is evaluating the ability of an agency to rely on single track operations during rail blockage disruptions. Song (1998) outlined the options available to agencies for providing service around a rail blockage. He described techniques for providing single track operations around a disruption and identified that a bus bridge may be needed where disruptions involved both tracks or where single track operations would not provide enough capacity to serve all passengers. He found that a line's most constrained segment depended on passenger load and the line's characteristics. Specifically, it depends on the time needed to implement and maintain single track operations, as well as the passenger demand and resulting vehicle loads determined for each track segment. He also explored the use of control strategies to increase capacity. Once this information is known, service managers can determine whether single track operations will provide enough capacity to serve all passengers or if other service is needed, such as a bus bridge. He developed a capacity constraint profile and control strategies plan for the MBTA Red Line.

## 2.6. Operations Management and Control Strategies Summary

This chapter has covered the key operations management concepts and real time service management for degraded or disrupted transit service. Prior studies have identified that providing even headways for high frequency service is the best way to reduce passenger wait times and all have used a variation of minimizing passenger wait time as its primary objective. The benefits of specific control strategies have been reviewed for both bus and rail service, but there has not been a review of how service managers should consider and select among options, which is the focus of this study.

# Chapter 3: Disruption Response Framework

This chapter presents an overview of the disruption response problem and agency operations environment, and develops the evaluation criteria used to assess the passenger impacts and agency costs created during disruptions and mitigated by disruption responses. First, the service management problem, real-time disruption management support, and agency policy assessment elements, concepts, motivation and mechanisms are described. Then service goals, evaluation criteria, and secondary impacts are discussed. The concepts developed in this chapter are applied in the disruption response process described in Chapter 4.

## 3.1. Overview

Service disruptions will occur despite the best efforts of agencies. Service managers must respond to any problems that occur, from managing service during snow storms and major emergencies to late or missing service. Therefore, the focus of any effort to improve disruption response must be on improving the effectiveness of real-time disruption management. The framework is built with this concept in mind. The framework is divided into three parts that describe distinct areas of disruption response. The first part, referred to as the core service restoration problem, assumes a disruption has occurred and focuses on how service managers should assess it and select a response to best mitigate passenger impacts balanced against agency costs. The second part supports real-time disruption response by identifying what disruption response elements can be pre-determined to reduce passenger impacts. The last part identifies and prioritizes which agency policies should be considered for changes in order to reduce the number of disruptions or to minimize passenger impacts when disruptions do occur.

Once a disruption occurs, the service manager must do the best he can to reduce passenger impacts, with inevitably limited time, information, and resources. Thus, the real-time service restoration process should be as simple and streamlined as possible. This leads to having disruption responses developed as much as possible beforehand, through pre-planning responses given the agency's current policy constraints, which is the purpose of the second part of the framework. These two frameworks are designed to improve the agency's disruption response capability by reducing the impacts of disruptions.

The third part provides a way for an agency to evaluate its policies affecting disruption likelihood and response. It is aimed at reducing the impacts of disruptions by both reducing the number of disruptions and the impacts of disruptions once they occur. While policies cannot be changed when managing disruptions in real-time, agencies should try to understand the full range of options service managers could have to manage disruptions if policies were different. This should lead the agency to understand which policies are critical to disruption response effectiveness, the role the current policies and constraints play in the disruption response and on which policies it should focus to improve disruption response.

## 3.2. Core Service Restoration Problem

The central service restoration problem is determining how service managers should respond to different disruptions. The primary concern should be to keep the rest of the service operating effectively, while attempting to minimize passenger inconvenience and complying with agency work rules and other policies and constraints. The first part of the framework provides service managers with information on how to develop their disruption response, with the service goals, passenger impacts, and agency work rules in mind.

This section describes the information service managers need to have or be able to estimate reasonably well to properly respond to disruptions and explains how it helps them respond effectively. First, the influence of agency policy is discussed. Then the elements of service managers' core service restoration problem are discussed. The first part of the core service restoration problem is gathering information on the disruption and the route affected. The second is determining the feasible service restoration techniques given the disruption information and route characteristics and developing a disruption response plan from the set of options.

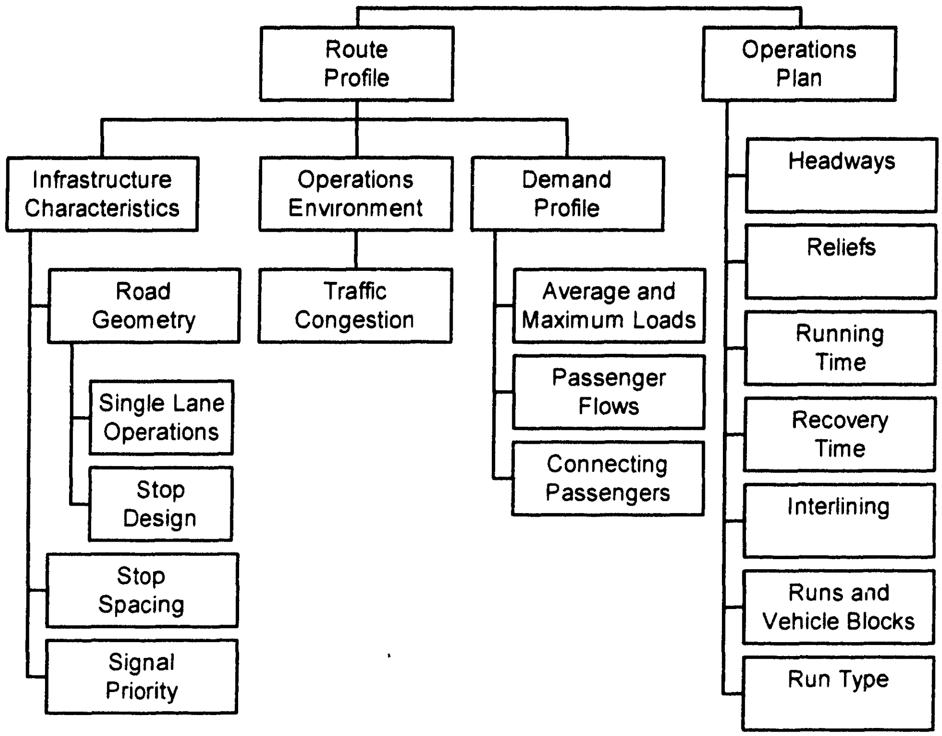
### 3.2.1. Agency Policy

Before discussing the specific information needed to address disruptions, it is important to understand that service managers formulating a disruption response must work within the context of existing agency policies. Agency policies play a key role in determining how effective disruption response can be. They determine the availability of additional resources, the level of staff training, and the existence of other disruption response support. While they are not part of the core service restoration problem because they are determined well in advance of the disruption and are typically in place for a long period of time, they facilitate and constrain disruption response. Therefore, service managers take policies as given when restoring service. Agency policies affecting disruption response are described in detail in Section 3.4.

### 3.2.2. Route Characteristics

Service managers need extensive information about both the affected route and the disruption itself to make effective service restoration decisions. This subsection and the next discuss the primary information of interest to service managers when responding to disrupted service: route and disruption characteristics. Service managers need to understand the affected route's characteristics because they influence which runs and vehicle blocks are critical, the appropriate service goals during disruption response and, therefore, what service restoration techniques are likely to be most helpful. Figure 3-1 below lists the route information needed for effective service management.

**Figure 3-1 Route Characteristics of Interest during Disruption Response**




## a) Route Profile

The route profile corresponds to what Rahbee (2001) describes as line characteristics. It describes the infrastructure. operations environment and passenger demand patterns of a bus route or rail line. It includes the infrastructure characteristics. vehicle operating environment, and passenger demand profile. Infrastructure characteristics and operating environment include road or track geometry, stop/station design, stop/station spacing, travel speeds and congestion, signal priority for bus, and track signalling for rail.

### 1. Infrastructure Characteristics and Operations Environment

The concept is similar for bus and rail transit, but the characteristics of concern are clearly distinct. Train movements depend on the track layout and protection systems in place, such as blocks and signals. Rail service managers are concerned with complex signalling and minimum safe separation, as well as possible problems with that infrastructure.

Bus operations depend on the street infrastructure and traffic operations characteristics. The lane geometry, including number of lanes, turning lanes, and on street parking can all constrain operations and service restoration options. Transit-specific elements, such as signal priority, stop spacing and stop design are also important considerations when evaluating the feasibility and effectiveness of service restoration techniques. Additionally, while the rail operations environment is highly controlled, the bus operations environment is not controlled by the agency and is heavily influenced by traffic variability and congestion.

### 2. Demand Profile

The demand profile describes the pattern of passenger boarding and alighting, and typical vehicle loads along the route by direction and time of day. This information is necessary to understand where passengers are and where they want to go, in order to provide sufficient vehicle capacity and minimize passenger delay. For example, some routes, such as crosstown routes may experience a sustained, multiple, or moving peak load point, while other routes, such as downtown commute-based routes may have an increasing load along the route with most passengers alighting toward the end of route. The disruption response should be different in each of these cases. Service managers also need to determine whether the route is experiencing its usual passenger demand or if there is a significant deviation from the typical demand profile.

## b) Operations Plan

As discussed in Chapter 2, the operations plan is the plan developed by the scheduling department and executed by operations on a daily basis. If there were no disruption, this should be the ideal plan for operations. It also serves as a reference for what operators and vehicles should be doing, thus highlighting service recovery constraints and opportunities.

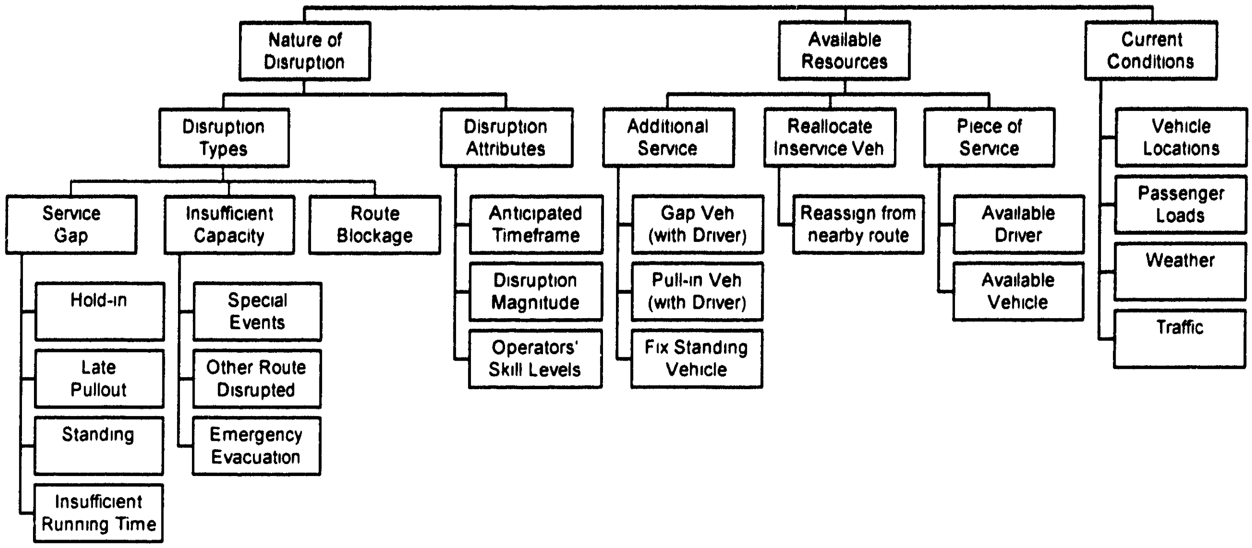
Service managers must be familiar with the operations plan and paddle of each run involved because they provide essential information about the future course of the disrupted and remaining runs. Critical operations plan information when managing a disruption are the headways, running times, recovery times, pull-out and pull-in times, interlining, and reliefs. Paddles, which describe each run in detail, provide additional information to service managers, such as run type, pull-out and pull-in times, and hours worked, which helps service managers identify work rule compliance and additional agency cost issues when changing their duties.

Pull-out and pull-in information gives service managers a picture of when the number of vehicles on the route changes. Specifically, pull-out information lets service managers know when new service should be on the street and pull-in times identify the potential for having an additional vehicle, driver or both. Interlining information flags potential impacts on other routes. Reliefs, run type, run start and end times, and hours worked provide the service manager information about the operator that is necessary to comply with work rules, which is not typically included in the operations plan or schedule summaries.

### 3.2.3. Disruption Information

As shown in Figure 3-2 below, disruption information includes the nature of the disruption, available resources and current conditions. This highlights what service managers would like to know about the specific disruption to which they are responding. While they generally need to be aware of this information, it is not necessary for them to have an exact answer for each of the items listed. Service managers can and should make service restoration decisions without all the information listed when the information they have indicates the response they are considering is better than doing nothing. The additional information may simply allow a better solution, but they do need to be careful not to implement service restoration techniques when missing critical information because of the potential to worsen passenger impacts (Turnquist, 1978). Generally, this information assists in selecting among feasible service restoration options.

**Figure 3-2 Disruption Information Needed for Disruption Response**




## a) Nature of Disruption

Service managers must identify the nature of the problem and other characteristics that will help structure the initial service goals and frame the response, as well as know specific disruption attributes, such as the timeframe, magnitude, and operator skill levels, which frame the service restoration options' costs and benefits.

### 1. Disruption Types

Disruption types were discussed in Chapter 2. There are three categories of service disruptions: service gaps, insufficient capacity, and route/rail blockages. Service gaps can be created by hold-ins, late pullouts, standing vehicles, or insufficient running time. Insufficient capacity is likely to occur with special events, evacuations, disruptions on other routes that lead many passengers to transfer all at once, and any other situation that creates a passenger demand spike. Lastly, route blockages typically occur from emergency response activity such as the paramedics, police or fire department, construction, special events, and accidents. Rail blockages are usually created from equipment problems or other incidents involving the track right-of-way.

## 2. Disruption Attributes

### Anticipated Disruption Timeframe

The disruption's anticipated timeframe influences what service restoration options service managers consider and the response selected. This is generally the first attribute of interest and can be estimated in minutes or hours for short-term or well-understood timeframes, or, more generally by service time period (am peak, midday, pm peak, etc.) if the disruption is anticipated to last longer or its likely length is unknown.

The key disruption timeframes are less than a trip, more than a trip, more than a time period, all day, and more than a day because of the implications on how the disruption is managed and service restoration options available. Disruptions that are less than a trip are difficult to respond to in time, and their impacts finite because they are so short, but they often create the greatest passenger impacts for any given trip. Disruptions spanning several trips must be managed to limit the passenger impacts. Disruptions should not span longer than a time period because at the end of the peaks, service should be available to end the disruption. The only disruptions lasting a day or longer should be unavoidable route blockages that are typically known about in advance and coordinated with the party creating the disruption.

Service managers also need an estimate of the time needed to change service to the modified operations plan, as well as switching back to the original plan. For example, for degraded service or disruptions that are likely to last less than one trip's running time, service managers may make some in-route service adjustments to get around the problem, but will not add service or change the dispatching instructions for future trips. For disruptions lasting longer than a trip, additional service must be considered and service managed through terminal dispatch adjustments until the disruption concludes.

### Disruption Magnitude

Timeframe estimation leads the service manager into the next attribute of interest, disruption magnitude. The service manager needs to estimate the likely scope and impacts of the disruption, given its nature and anticipated timeframe. In general, as the estimated time period grows, the number of passengers, runs, and reliefs impacted increases. For longer disruptions or those with wider impact, the service manager will need to look for a more sustainable disruption response that can be maintained for the duration of the disruption, whereas with shorter or limited impact disruptions, service managers may not want to stray far from the operations plan, knowing they will need to move the route back to the original plan shortly.

### Operator Skill Levels

Operator driving skill levels and knowledge of service restoration techniques are important considerations. They permit or prevent the use of specific service restoration techniques, determine the amount of time needed to explain their instructions to operators, and serve as an indicator of the likelihood that the service manager's instructions will be implemented correctly. For example, a service manager may ask only a skilled operator to try to make up some time or leapfrog its leader to pick up passengers and assist in speeding up the leader because he knows a skilled operator will make safe driving decisions and has the best chance at being able to carry out the instructions.

## b) Available Resources

The resources available to deal with the disruption can include unassigned extraboard operators, operators available to work overtime, the number of operating vehicles available for service, "gap or extra buses" (operators and vehicles scheduled to be at the disposal of service managers) and equipment that can be borrowed from nearby, parallel routes. These are the opportunities for which service managers are looking. While they must use other service restoration techniques to manage routes without enough equipment or running time, they want to find and use the least disruptive and least costly resource available as soon as possible to restore the scheduled service and return to the operations plan.

## c) Current Conditions

Current conditions highlight aspects of the operations environment which may differ from normal or change during the disruption, and which may affect the preferred response. Vehicle locations, passenger demand, weather, and traffic are the most variable items that can aggravate disruptions, but this category also includes any other relevant characteristics that have deviated from their normal condition.

### 3.2.4. Feasible Service Restoration Options

Service restoration options are tools used to reduce passenger impacts, not avoid disruptions. In this section, these techniques are first linked to disruption type and then combined to create a set of responses for each disruption type. Then, personnel management restoration techniques are introduced.

#### a) Feasible Operations Control Techniques

This section describes which service restoration strategies are appropriate for each of the three general disruption types: service gap, insufficient capacity and route/rail blockage. Table 3-1 shows the feasible restoration techniques for each disruption type and indicates which are likely to produce service improvements. The amount of benefit provided and the agency cost to implement the service restoration technique will be discussed in Chapter 4.

**Table 3-1: Service Restoration Techniques by Disruption Type**

| | **Disruption Type** | | |
|---|---|---|---|
| **Service Restoration Technique** | **Service Gap** | **Insufficient Capacity** | **Route/Rail Blockage** |
| Replacement/Extra Service | + | + | + |
| **From Terminal** | | | |
| Respacing[1] | + | 0 | + |
| **In-Route** | | | |
| Limited Stop/Drop-off Express | + | + | 0 |
| Holding | + | 0 | 0 |
| Short-turn | + | + | 0 |
| Re-route[2] | N/A | N/A | + |

+ Improves service. 0 No service effect.

[1] High frequency service only.
[2] Route/rail blockages only.

In general, extra service can always be integrated into the disrupted route so as to reduce passenger impact during any type of disruption. Respacing the route can improve service with a service gap or route blockage. While service is in-route, the expressing options and short-turning can be used when there is a service gap or insufficient capacity and holding only improves service when there are service gaps. Thus, service managers should first look for no-cost replacement service, such as a gap vehicle and operator or an operator with a vehicle who is not scheduled to work the full number of hours to reach the guaranteed pay minimum. The only constraint on using these resources is that there is not another problem elsewhere in the system which would be a better use of these resources.

### Use of In-Route Techniques

Respacing service from the terminal has a straightforward implementation. In-route techniques are a little more difficult to use properly and there are more variations, as shown in Table 3-2 below. They are applied to different vehicles together to create one coordinated response. For example, to provide more capacity along one section of the route, the service manager may hold or slow a vehicle within the section, while he may express vehicles to the beginning of the section.

**Table 3-2: In-route Service Restoration Techniques**

| |
|---|
| Holding/Slowing |
| Limited Stop Express |
| Drop-off Express |
| Limited Stop/Drop-off Express |
| Off-route Express |
| Leapfrogging[3] |
| Short-turn |

While the general in-route service restoration techniques were described previously, there are variations on how expressing is implemented that have not been explained. Limited

---

[3] Leapfrogging is typically used as a passive service restoration technique once vehicles are bunched. As a matter of agency policy, operators may be instructed to pass their leader, if he has fallen behind his schedule. It allows the capacity in both vehicles to be better used and allows them to serve the route a little faster. In that sense it is a service restoration technique, but service managers would not choose to bunch vehicles just to use it.

stop expressing skips stops in order to increase the vehicle's speed. Vehicles should still serve major boarding and alighting stops because of the high demand. Drop-off only expressing only allows passengers to alight, but does not stop to board any passengers. This may be a good option when the vehicle is completely full. One note is that it is difficult to stop to allow passengers to get off the bus, but not allow passengers standing there to board, so often agencies choose to allow passengers to board when there is space on the vehicle if it has stopped to let passengers get off the bus. Limited stop/Drop-off expressing is a combination of the two. Vehicles stop both to allow passengers to alight and at major boarding locations to allow passengers to board. Lastly, off-route expressing is where the operator is instructed to take another, faster route, typically a parallel arterial or expressway for buses, and start service at a certain point along the route. Passengers must be notified of where the vehicle's next stop will be, since they may need to get off the vehicle and wait for the next one. Expressing without passengers is considered deadheading and is used to quickly put a vehicle in service somewhere along the route.

#### Personnel Management Techniques

Personnel management techniques are tools that should have a negligible effect on passengers, but move operators where they need to be for their run schedule, to avoid extra agency cost, or to comply with operator work rules. While they do not improve the passenger experience, they can be used to improve operator satisfaction and reduce agency cost. The two primary techniques are swapping reliefs when operators arrive late with the vehicle at the relief point and exchanging drivers to get operators back to a route terminal faster.

### b) Service Restoration Plan Options

It is important to understand how the various service restoration techniques can be used together to improve service. This section describes the available combined operations control strategies that create disruption response options for both the initial in-route response and sustained terminal-based service management. The options discussed below assume that adding no-cost service has already been explored and is not available.

## 1. Service Gap

### Adding or Reallocating Service

After the no-cost additional service option is explored, service managers should determine whether there is additional service available at a cost. If there is, he has two options. He can add service at a cost to the agency or manage the route with the existing resources. In order to decide whether additional service is warranted, he must estimate the negative passenger impacts if he does not add the service. Basically, he must determine what the achievable headway is with and without the added service and, given the passenger demand, what is the difference in passenger impacts between those two headways.

If there is no additional service available, even at a cost, the service manager must investigate his service reallocation options. Typically, it is cost-effective to reallocate service from a high frequency route to a low frequency route when the low frequency route experiences a service gap because of the large passenger impact on the low frequency route and relatively low passenger impact on high frequency service with missing service. Nearly all of the passengers waiting during the service gap on the low frequency routes will be forced to wait an additional headway. Lastly, he can choose to manage the remaining service if additional service is not necessary or if there are no additional or cost-effective service reallocation options available.

### Managing Service

If a service gap disruption is discovered before the gap passes the major passenger boarding locations, in-route techniques, such as holding and expressing should be considered. Specifically, the gap leader should be held into the gap and the gap follower should be expressed if possible. Successfully expressing the vehicle following the gap is very unlikely because of the additional passenger demand placed on it. There are more passengers at each stop because passenger arrival rates are steady and there is a larger than scheduled service gap, making the vehicle fall farther behind its schedule. This makes holding the leader the only dependable in-route option. Holding the leader also shortens the gap and reduces the passenger load on the gap follower. Short-turning is an option if capacity is needed along the main portion of the route and one end of it has sufficient capacity in both directions with one vehicle missing.

Once the gap has passed the boarding portion of the route, only terminal-based options are available for managing subsequent trips. Routes with service gaps should either have extra service added or the service should be re-spaced from the terminal. As discussed in Chapter 2, re-spacing should only be considered for high frequency service where passengers arrive randomly to achieve the best possible headway given the running time and number of vehicles operating on the route. Low-frequency service must have the missing service replaced to keep to the schedule as much as possible.

### 2. Insufficient Capacity

In-route techniques available to address insufficient capacity disruptions are adding more service, expressing and short-turning. Terminal-based management does not address the disruption for future trips for the same reasons: re-spacing is not likely to add enough capacity. Therefore, the service must be provided by either adding service (extra service or short-turning) or reducing the running time (expressing). Additionally, short-turning would only be effective if capacity were needed along the main portion of the route and one end of it had a much smaller demand in both directions when compared with the trunk section.

For this type of disruption, extra service is usually the only solution that will properly address the problem because re-spacing service along the route is unlikely to provide enough capacity to serve all passengers, making the only acceptable solution to add more service. Additional service may be available from the garage, pull-ins, and nearby routes.

### 3. Bus Route Blockage

The primary restoration technique used is re-routing the service. Once this is accomplished, the service manager must address other disruptions that have been created from the blockage. Most blockages will create an initial service gap and future trips will have insufficient running times. Lengthy disruptions may create an insufficient capacity disruption. If available, extra or reallocated service may be used to serve passengers past the blockage before buses are stopped behind the blockage for more than a headway, particularly if there will be capacity problems.

To the extent possible, the initial service gap impacts should also be managed with in-route techniques, such as slowing vehicles in front of the blockage or expressing vehicles once vehicles are at capacity after they pass the blockage. Sustained disruptions, such as insufficient running time, may develop, which will create the need to use additional service or manage the additional disruption as long as the re-route is in place.

#### 4. Rail Blockage

Re-routing rail service is more complex and the blockage with any additional disruptions created is a unified problem because of the constrained track capacity and the inability for vehicles to pass. Buses use a standard re-routing procedure regardless of other impacts, but a rail blockage re-route option is selected based on the other anticipated disruptions, such as insufficient capacity or running time disruptions.

Assuming a two-track alignment, if both tracks are blocked rail service cannot pass the blockage. The agency must stop service or look for service alternatives, such as a bus bridge. If only one track is blocked, service managers can hold vehicles ahead of the blockage, short-turn in front of, or behind, the blockage, single-track around the blockage, or institute alternative service. The selected response depends primarily on passenger demand, capacity provided, and anticipated length of blockage, but is also affected by track layout, terminal operations, and the time needed to single-track or short-turn trains (Puong, 2001 and Song, 1998).

## 3.3. Disruption Response Support

This second part of the framework discusses how pre-planned support can help increase the effectiveness of disruption responses and how to determine the proper mix of pre-planned and real-time decision-making. Pre-planning disruption responses is one of the primary ways to achieve better service restoration management because it helps the service manager to identify the best disruption response quickly. The support framework presents the motivation for pre-planning, the criteria for determining which disruption responses should be pre-planned, and to what extent they can be reasonably pre-planned, and which disruption responses are poorly suited to pre-planning.

### 3.3.1. Motivation to Pre-plan Responses

Time is the primary concern for service managers when dealing with service disruptions because passenger impacts are growing and response options diminishing as time progresses. Reducing the amount of time needed to develop an effective response allows service managers to focus on implementation and monitoring, rather than developing the plan itself. Senior service managers and other staff can develop pre-planned disruption response guidelines and SOP's for service managers. Pre-planning also requires them to sit down and think about the important issues and possible opportunities beforehand. Pre-planned responses may need to be modified in real-time, but they should identify the likely service goals and service restoration techniques to be used for common or major disruptions. This process may also assist senior service managers in identifying the key policies that currently constrain effective service management.

In addition to an insufficient amount of time to make sound decisions, a second concern is that service managers do not have the necessary tools to properly assess the correct service goals and/or the effectiveness of various service restoration techniques when responding to a disruption. Even if tools were made available to service managers, they would still encounter the first problem, insufficient time to use them effectively in real-time. A related motivation is to develop guidance on difficult value judgments or choices which must be made while responding to disruptions. One example of this is determining which situations merit a short-term fix while looking for a longer-term solution.

Preplanning can also be used as a training tool for service managers, which should lead to more uniform service restoration responses. While many service managers will immediately know what approach they will take to a given disruption once they have the basic information, the effectiveness of this approach depends on their expertise and knowledge, understanding how the route operates, and whether the service manager has encountered this type of disruption before. Pre-planned service restoration responses should improve junior service managers' responses due to their limited experience. Senior service managers' abilities to manage unfamiliar service types should improve and the pre-planning process should reinforce the importance of understanding the affected route and disruption, as well as the use of service goals and decision-making criteria. Pre-planning should also help eliminate learned service restoration responses that do not work well and support the use of already-used responses that do work well.

### 3.3.2. Disruption Response Pre-planning Criteria

The first issue facing service managers when pre-planning is which and to what extent should disruptions be pre-planned. While pre-planning can offer service managers clear-cut, well thought-out solutions designed to improve service delivery during disruptions, these cases need to be worth pre-planning and have route and disruption characteristics that rarely change. There are a few basic criteria for determining which disruptions are useful to pre-plan and to what extent they should be pre-planned. Disruptions that justify pre-planning must be common, critical, costly, and difficult to manage. These characteristics can be broadly categorized into three areas, frequency of occurrence, passenger impact, and disruption characteristics.

Rather than establishing specific thresholds for pre-planning their disruptions, the agency should prioritize disruptions for pre-planning based on these criteria and then move down the list, pre-planning all worthwhile scenarios. For disruptions not meeting these criteria, the agency's response might still be improved with pre-planned decision-making support. This is where determining the appropriate mix of pre-planning and real-time decision-making, discussed below, is helpful.

### 3.3.3. Pre-planning / Real-Time Decision-Making Mix Overview

While a limited number of disruption responses should have fully pre-planned or entirely real-time response. most will have specific elements that can be pre-planned with the rest of the response needing to be developed in real-time based on the specific disruption characteristics. This subsection describes the different levels of pre-planning and develops criteria for determining the right mix of pre-planning and real-time development. Figure 3-3 below shows the various pre-planning levels available.

**Figure 3-3: Levels of Pre-planning and Real-Time Decision-Making**

| | | | |
|---|---|---|---|
| All Real-Time Decision-Making | Mostly Real-Time | Limited Real-Time | No Real-Time Decision-Making |
| ◄ | ——— | ——— | ► |
| No Pre-planning | Restoration Guidelines | Pre-plan Elements | Pre-plan Entirely |

For disruptions that do not warrant having the entire response pre-planned. service managers can evaluate the proper mix of real-time and pre-planned decisions. as described in the paragraphs below. The appropriate mix by disruption type will be discussed in Section 3.3.4.

#### Entirely Pre-planned Responses and Pre-planning Elements

Disruption responses appropriate for being entirely pre-planned are limited because of the difficulty of anticipating disruptions and the difference in the appropriate approach with changing disruption and route characteristics. Full pre-planning includes instructions for all runs from disruption start to finish. The next level of pre-planning. pre-planning elements of the disruption response. can include two aspects: pre-determining how to approach specific aspects of a particular disruption and making specific decisions in advance.

Although the operations environment or passenger demand may deviate from its typical characteristics. this cannot be included when pre-planning for disruptions. Large deviations would likely be caused by some other problem that would require more effort to manage, making these disruption scenarios less likely to be useful to pre-plan. Service managers should have at least a general idea of this information before there is a disruption.

### Response Guidelines

The next level is response guidelines, where the reliance on real-time decision-making is much greater. Only the basic underlying approaches to the restoration response are addressed. In general, they are recommendations for how service managers should approach service restoration on different types of services, such as for crosstown vs. feeder services.

General guidelines include all the general concepts service managers need, such as potential resources and their costs, knowledge about managing different services with different passenger demands and how to respond to different disruption types. They should be used during uncommon disruptions with service restoration options for which it is difficult to estimate passenger impacts. This is because of the limited information available in advance about the specific disruption and the clear need for support.

### Entirely Real-Time Decision-Making

The last level is full dependence on real-time decision-making. None of the disruption has been pre-planned and the service manager must develop the entire plan and implement it as best he can with the information available. All decisions are based on his personal experience subject only to the standard operating procedures and work rules. This is what occurs by default when service managers do not have pre-planned options on hand or in their heads.

## 3.3.4. Determining the Pre-planning / Real-Time Decision-Making Mix

A summary of the general approach to the pre-planning / real-time decision-making mix by disruption type is shown in Table 3-3 below. Common hold-ins and rail blockages are the primary candidates for developing "off-the-shelf" responses, while standing vehicles are the most dependent upon real-time decision-making because of their highly variable nature. The following sections describe how to determine the appropriate mix of real-time and pre-planned decision-making for each disruption type. Critical runs and vehicle blocks are those involved in reliefs or assigned to low frequency routes. Non-critical runs and vehicle blocks are school trippers or other short time length service providing extra capacity or independent part-time pieces of work that pullout and pull back into the garage.

**Table 3-3: Real-Time / Pre-planned Decision-Making Mix**

| | | Disruption Type | | | |
|---|---|---|---|---|---|
| Frequency of Occurrence[4] | Passenger Impact[5] | Hold-in and Rail Blockage | Insufficient Capacity or Running Time | Bus Route Blockage | Standing Vehicle |
| Common | Severe | F | F | P | P |
| Common | Moderate | F | P | P | G |
| Uncommon | Severe | P | P | P | G |
| Uncommon | Moderate | P | G | G | R-T |
| | | **F** – Fully Pre-plan | **P** – Partial Pre-plan | **G** – Guidelines | **R-T** – Real-Time |

## Hold-ins

Besides scheduled disruptions, which are discussed below, agencies have the most advanced warning and resource allocation flexibility with hold-ins. Common hold-ins, especially hold-ins on non-critical runs and vehicle blocks, should be entirely pre-planned. Less common hold-ins can be planned in part by pre-identifying the service goals and pre-selecting the service restoration techniques to use. This will be described in greater detail in Section 3.3.5.

## Rail Blockage

Rail blockages create such a severe impact on operations and to passengers that entirely pre-planning disruptions of different timeframes, during different time periods, and in

[4] Common disruptions occur at least weekly.
[5] Moderate passenger impacts are a maximum wait time of 1.5 – 2.5 headways or 15 – 20 minutes and severe passenger impacts are a maximum wait time greater than 2.5 headways or more than 20 minutes

each of the track segments is likely to be worthwhile. Models like Puong's (2001) can be used to identify how to respond to different blockage scenarios in-advance or in real-time. While uncommon or complex rail blockages may not need an entirely pre-planned response, service managers would still be able to use the disruption timeframe information highlighting capacity problems, the pre-planned single tracking solutions, the pre-identified service goals, and the pre-selected service restoration techniques.

### Insufficient Running Time and Capacity

Common insufficient running time and capacity disruptions creating severe impacts should have disruption responses entirely pre-planned because of the difficulty of developing a plan in real-time. For less common running time or capacity disruptions, many of the responses' elements can be pre-determined, such as established expressing segments or short-turning points along the route or on parallel arterials with high operating speeds. Another useful rule of thumb for high frequency routes facing a running time problem is pre-determining a new headway that accommodates most disruptions and developing a response for that case. For example, moving from a seven-minute headway to an eight- or nine-minute headway. Another pre-identified element for routes that develop capacity problems quickly should be parallel routes from which to borrow service.

While service managers could probably successfully determine items such as the proper service goals and service restoration techniques, it would be more difficult to determine other things in real-time. For example, pre-determining elements such as where it is appropriate to express or short-turn vehicles is important because it is very difficult for service managers to properly establish these locations in real-time and the effectiveness of these techniques is highly dependent upon where along the route it is implemented, which depends on the detailed route profile information.

### Bus Route Blockage

Because buses generally operate on shared, multi-purpose rights-of-way, most bus route blockages are not created or controlled by the agency. They are typically due to road or adjacent construction, emergency services, special events, or accidents. Thus, the focus is on impact mitigation. Many bus route blockages, such as those due to construction or special events, can be anticipated, however, and the agency should coordinate with the appropriate public and private entities to obtain this information in advance. For anticipated disruptions, the agency should entirely pre-plan the new service by developing a new operations plan, if necessary, and providing new routing instructions, passenger information, and new paddles for each run for the duration of the blockage.

Unanticipated blockages, such as fire or police activity, will need to be managed more in real-time, but re-route plans should be developed in advance. Service managers need to know which re-route paths are legal, which are not a good idea but can be used, and which are illegal. Because this information rarely changes and would be difficult for service managers to identify in real-time, re-route guidelines should be developed for each route by major route segment. Critical blockage timeframes and lengths of additional running time created should be pre-determined and service managers should identify parallel bus routes from which service can be borrowed when needed for unanticipated disruptions. However, establishing service goals and restoration techniques will depend on the anticipated disruption timeframe and the additional running time needed.

Route blockages may also lead to another disruption occurring on the route, such as insufficient capacity or running time. In which case, service managers will need to select from those pre-planned elements to solve those problems.

### Standing Vehicle

Standing vehicles occur in a variety of ways, making entirely pre-planning these responses impossible. However, many elements can be pre-planned, such as the likely service goals for different anticipated disruption timeframes and disabled vehicle locations, service restoration techniques, timeframes in which the service must be managed or replaced, and parallel routes from which to borrow service for low frequency routes.

### 3.3.5. Disruption Mitigation Pre-planning

The prior sections have discussed what agencies can pre-plan to improve disruption responses once a disruption has occurred. This last section on disruption pre-planning develops ways to redefine disruptions in order to mitigate their impacts. The idea is that, although the disruption is not avoided altogether, the new disruption's impacts are substantially smaller than those of the original disruption facing the agency. This can be achieved by reallocating resources in the garage before runs pull-out or once in the field to protect critical, high impact service. Mitigation cases are discussed below. Disruption avoidance and mitigating passenger impacts through policy changes are discussed in Section 3.4.

#### Prioritizing Hold-ins

When a garage is faced with hold-ins, the critical services must be protected and resources assigned to them first. If there is a vehicle shortage, it is important to fill runs pulling out with vehicles having all day vehicle blocks, so the equipment is on the street. This will help limit the impacts on different runs throughout the day. Similarly, in the case of an operator shortage, the pull-out staff should cover long runs because finding an operator later to cover the run would be unlikely, particularly in the peak when all service is needed. Additionally, runs should not be held-in on low frequency routes because of the difficulty of managing the disruption in the field. If necessary, vehicles or operators on high frequency routes should be reassigned to the missing low frequency runs before leaving the garage, which is discussed below.

Further, resources should be assigned such that the easiest-to-manage disruption is the one that makes it to the field. For example, pre-planned responses for service hold-ins should lead the agency to fill critical runs first, such as full-day runs, and hold-in easier to manage disruptions that have shorter timeframes or smaller impacts, such as runs that are not full shifts, school trippers, or those that only operate during the peak. These priorities are explained in detail below.

### Hold-in Guidelines for Operator and Vehicle Assignment Clerks

Agencies should create hold-in guidelines for operator and vehicle assignment clerks to follow. For example, the Chicago Transit Authority (CTA) created draft guidelines to standardize practices in their eight bus garages. Their approach focuses on prioritizing missing runs, rather than protecting vehicle blocks and is designed to limit the impacts of the service reduction on passengers. They identified both what the garage should assign in advance and how to prioritize holding in runs. This can be used as a model from which other agencies can develop their own hold-in guidelines.

Primary candidates to hold-in include short trippers on heavy routes, routes with scheduled short-turns, routes with headways of less than 10 minutes, and routes without other missing service. Critical runs to fill are the first and last trips of the day, the first and last trips of any given service (such as rush hour expresses), trips whose leader or follower were held-in, back-to-back trips on any branch or varying terminal service, night owl trips, and routes with only one vehicle. As a second tier, it recommends protecting off-peak trips, headways at or above 15 minutes, line-hauls that interline with downtown shuttles, routes with timed connections, and routes transporting relief operators to their relief points.

### Assigning Reliable Vehicles to Critical Blocks

Another way to use garage assignment to mitigate the impacts of vehicle failures is to assign the most reliable vehicles to critical vehicle blocks. While this will not eliminate disruptions on critical service, it should reduce them.

### Low Frequency Service Disruption Avoidance

Another type of disruption that should be avoided as much as possible is a service gap on low frequency routes because of the difficulty of adequately managing these disruptions. As discussed in Chapter 2, the schedule-based passenger arrival pattern makes respacing the service to provide even headways ineffective at best and creates longer passenger wait times at worst.

Thus, it is necessary to provide all scheduled vehicles because otherwise it results in poor service quality and makes successful operations (making reliefs and maintaining the schedule) extremely difficult. Disruptions in low frequency service are, therefore, essentially unmanageable. Missing service must bc replaced as soon as possible to provide reasonable service and not create operations problems on the route. Garages should make every effort to eliminate disruptions to these routes by not holding-in runs and service managers must replace missing service created from standing buses as soon as possible from neighboring routes.

## 3.4. Agency Policy Assessment

Agency policy is the overall environment, standard operating procedures, guidelines, and constraints under which service managers operate normal service and respond to disrupted service. First, this section identifies the relevant policy areas affecting disruption avoidance and mitigation. Then it discusses the evaluation of the effectiveness of agency policies, including the operations environment, staff availability and locations, the communications system, and disruption response procedures and guidelines. The assessment develops a framework for how agencies should think about the impacts of their policies and a way for them to evaluate where they are in the range of options and where they should be. Policy changes may result in new options that are better for passengers and/or the agency.

There are two ways to reduce the impacts of disruptions. The first is to reduce the number of disruptions and the second is to reduce the impacts of disruptions by improving disruption response. Previous sections have discussed service restoration responses and disruption mitigation options to choose from once the disruption has occurred, given the agency's existing policy constraints.

Section 3.4.1 presents the primary reasons to evaluate the impacts and cost-effectiveness of current agency policies. Section 3.4.2 describes the policy areas affecting service during disruptions. Section 3.4.3 provides a summary of which policy areas affect each disruption types and how amenable they are to avoidance and mitigation. Section 3.4.4 outlines mechanisms available to agencies to avoid disruptions and discusses policy areas and mechanisms to improve disruption response.

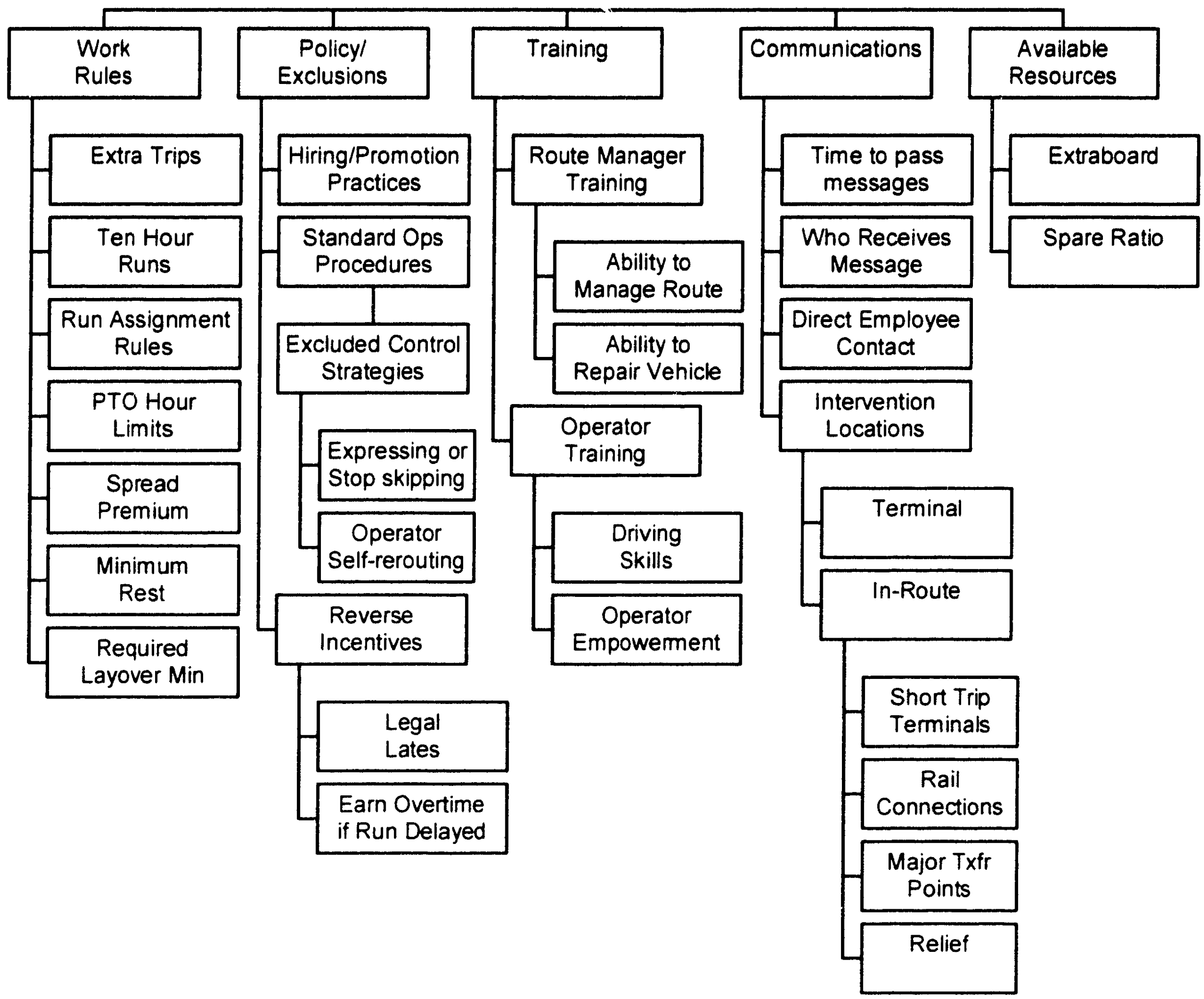
### 3.4.1. Policy Assessment Motivation

While agency policy provides options to service managers, it can also serve to constrain the service restoration options available. Agencies should evaluate the cost-effectiveness of potential agency policy modification because it is likely that some disruptions could be better managed if agency policies were changed to provide service managers with service restoration options that would have fewer negative passenger impacts than those currently available. Policies of interest include resource allocation and technology-based service management support systems. Additionally, some disruptions create an unacceptable level of negative passenger impacts and cannot be managed well enough to sufficiently reduce those impacts, even with pre-planned support, and, hence, should be avoided when possible. These disruptions may be avoided by providing enough resources to cover these identified critical services and by properly allocating the available resources to these critical services.

### 3.4.2. Service Management Environment and Constraints

This section describes the policy areas relevant to disruption response. Figure 3-4 shows the major agency policy areas that can affect service management's ability to manage service disruptions effectively. These areas include agency work rules, operations and hiring policies, training, policy exclusions, the communications system and service management staffing, and available resources. While the agency should be concerned with evaluating each of these areas because of their broader effects on service delivery, it is also important specifically to consider their impacts on disruption management.

**Figure 3-4: Agency Service Management Environment Elements**




These topic areas frame how operations management and staff think about and can respond to disruptions. More specifically, they determine whether disruptions can be avoided altogether and which service restoration options are available to service managers as they manage service disruptions. For example, the agency's extraboard, or operator reserve, and size and number of available vehicles determine whether or not hold-in disruptions can be avoided, and if a problem develops while a vehicle is already in service, the likelihood of having a replacement vehicle or operator available.

In the short term, the existing rules in these areas represent constraints on service managers' decisions on how to respond to disruptions. In the longer term, changes can be made, including work rule changes during contract negotiations, acquiring and implementing new technologies to support agency functions, changing hiring practices to better support agency and operator goals, redistributing service managers' physical locations or responsibilities, changing the extraboard size, or changing the vehicle spare ratios.

The primary reason for agency policymakers to be concerned with each of these areas' effects on disruptions is that they determine the internal decision-making and implementation environment service managers face. Day-to-day service management decisions, service managers' and operators' experiences, and their disruption responses are all framed within this context. Disruption responses that violate these constraints are not considered when making real-time service management decisions or developing pre-planned responses. Thus, service managers dealing with disruptions on a daily basis may develop myopia when considering how to improve service. They may plan responses based on the agency's current environment, not considering what could be a better solution given a changed environment.

The agency should evaluate the opportunities available, but not considered by service managers, because of these existing constraints. The agency environment, while set on a day-to-day basis for service managers, can be changed over time. There may be some unnecessary or costly constraints that impact passengers, operators and the agency, which could and should, be changed. Agency policymakers can make a large impact on service delivery by changing any of these areas. The assessment process should determine which areas have the greatest potential for improvement. The following six sections discuss each of these policy areas in greater detail.

### a) Work Rules

Work rules include contract- and agency-based rules governing the work of operators, service managers, dispatchers and others involved in providing or restoring service. As shown in Figure 3-4, the specific areas of concern are the use of 8 hour vs. 10 hour runs and part-time operators, having operators make extra trips, work assignment rules, spread limits, minimum rest between runs, and required layover time. Work rule constraints are typically laid out in the agency's labor contract, making the structure of these constraints highly dependent upon the contract. Some may prohibit certain uses of operator time, while others cost the agency if applied.

The agency should continually assess this part of the operating environment because this may be a primary constraint in finding reasonable disruption responses. Determining what types of operator flexibility service managers generally need to manage service problems should show policymakers where the greatest benefits can be found. Additionally, evaluating prohibited options, such as having operators work different runs than they selected during the bidding process during a disruption, and overtime penalty options may be useful policy options to evaluate to determine which provide the greatest benefit. This may include reviewing the workforce size, percentage of part-time, 8 hour and 10 hour operators, prohibited operator uses, part-time operator limits on a daily, weekly, or monthly basis, and the flexibility of operator layover time. Work rule concerns and evaluations may also extend to support functions, such as vehicle maintenance or garage administrative staff, if there are inefficiencies here that keep good equipment or operators out of service.

#### b) Policy/Exclusions

Agency policy and excluded practices include both field management and other agency departments' support practices. The major interest areas in this section are hiring and promotion practices, operator, service manager or service manager behavior penalties and incentives, and excluded service control strategies, as shown in Figure 3-4 above. The first, agency hiring and promoting practices, is generally in the domain of Human Resources, with input from Operations, Service Planning, and Scheduling. These policies are linked to some of the work rule constraints mentioned above because the work rules will likely have an impact on attrition and new hiring rates.

### 1. Hiring Practices and Use of Part-Time Operators

The agency benefits of having part-time operators in the workforce include being able to cover smaller pieces of work without paying an operator for a full 8 hour day, reducing benefit costs, and having more flexibility in operator assignments. In return, part time operators have time to pursue other interests or even other careers.

One concern in the industry has been who is attracted to work as an operator and what their career goals are (MacDorman, 1986). There are a variety of practices in employing part-time operators in the workforce. Some agencies keep separate hiring pools, while others hire their full-time operators exclusively from the pool of current part-time operators. This is a concern because the goals of part-time and full-time employees are often different and they may not be satisfied serving in the other capacity. Potential employees interested in full-time work may not want to work part-time until a full-time position opens and those interested in part-time work may only want to work part-time and not transition into a full-time position.

Promotion practices are also important in the transit industry because, in most agencies, most operations managers start as operators and are promoted within Operations. This has been a traditional practice because agencies value having field management, dispatchers, and senior operations management know about operations as well as the different problems operators may encounter. Regardless of the merits of this system, while it is still dominant in most agencies, agencies should try to attract operators who can eventually become an effective part of operations management.

### 2. Excluded Control Strategies

The second area of interest is policy-excluded control strategies, which is where agencies decide a specific restoration technique is so egregious to passengers or operators that they prohibit its use. Common exclusions are skip-stop or express service because the vehicle passes passengers waiting for service. There are several key issues, which may aggravate passengers, including how the agency performs the skip-stop service, what information is provided to customers and how close is the following vehicle. Many of these negative impacts can be properly managed, perhaps allowing the techniques' positive effects to dominate.

### 3. Inadvertent Incentives

The last area of interest is of greater concern during regular service operations. It is the incentives inadvertently given to operators that may be contrary to the agency's interests. Two examples are the acceptability of being late, or early, and the ability of operators to earn overtime when they are delayed and return to the garage later than scheduled. Being early is generally not tolerated because operators can control this. However, particularly in the case of bus operations, being late is generally seen as not being the operator's fault because there are so many circumstances beyond his control, which may contribute to lateness. Management's tolerance for being late will likely depend on the circumstances, including the schedule, road and passenger demand conditions, and the operator's experience. Rail operations management may be much less tolerant of tardiness because of the lack of external influences, such as traffic congestion, as well as the severe impacts of slow operations on following trains.

Another inadvertent incentive can exist in how runs are structured. Operators earn overtime when they work more than 8 hours, which may encourage some operators whose runs are close to 8 hours to drag the street and return late. These reverse incentives are general service management concerns, as opposed to service restoration concerns.

## c) Training

Training plays an important role in getting the best service restoration solution implemented properly. Both operator and service manager training are important to service delivery and restoration. Training affects service managers' abilities to use restoration techniques properly and operators' abilities to implement them. It dictates what most employees understand and consider to be important factors, and their available skills. Although some individuals may have a particular skill, if it is not addressed in training, most employees are unlikely to have the skill.

In addition to service restoration training, vehicle operating skills training can assist with disruption management because better operating skills should reduce the number of in-service mechanical failures and accidents, thereby reducing the number of disruptions.

### d) Communications

The communications system plays a large role in the constraints and effectiveness of managing or restoring service. The two primary aspects of interest in the communications system, with respect to managing disrupted service, are the ability and amount of time needed to pass messages among service management staff, including field-based management, dispatchers, operators and repair trucks. If done properly, it can support the agency's service management structure in providing them more information and intervention options. This is important because who receives information and when they receive that information directly affects the magnitude of the service disruption created for passengers and operators and the restoration options available to service managers.

Two important parts of communications are the service management structure and automatic route management capabilities. Service management staffing and distribution of responsibilities is an important impact area because it determines which problems get managed and how. Automatic route management allows service management to manage more critical events, while the smaller impact service problems, such as service degradation, are addressed with the system. These relationships are explored in more detail in Barker (2002).

### e) Available Resources

This is the last element of the service management environment. It refers to the availability of operators and vehicles, as well as the availability of operators to work overtime assignments. This area is concerned with the overall match or mismatch of extra resources and the cost-effectiveness of the mix of extraboard and overtime operators. Availability of resources during disruptions, including reassigning equipment from nearby routes to the disrupted route, is addressed in the pre-planning phase.

In addition, the timing of each type of resource availability is critical because the supply and demand of operators must be matched temporally. Unanticipated critical hold-in runs at the end of the pull-out period are the most difficult to address because of the inability to reassign resources to those runs, since most of the service has already left the garage, and unassigned operators and vehicles are likely to have been assigned to earlier pullouts. One possibility is having some operators report later in the pull-out period to cover any last-minute hold-ins. Another option is to explore the efficiency of having a couple of part-time runs scheduled to pull out late in the pull out period in order to have non-critical runs whose operators and vehicles can be reassigned to critical service. However, scheduling part-time runs later in the pull-out period may be less efficient than having additional or rescheduled extraboard operators due to the peak demand nature of providing service.

### 3.4.3. Policy-based Disruption Avoidance and Impact Mitigation

This subsection shows how changes in different policy areas can reduce the likelihood of disruptions or facilitate disruption response. The changes' potential benefits, costs, and mechanisms are also described. Different disruption types have different potential for disruption avoidance and impact mitigation.

Policy makers need to understand how they can reduce the occurrence of disruptions and service managers need to understand how to reduce the impacts of disruptions. Hold-ins have the greatest potential for direct avoidance because the agency has more advance notice and options to avoid or mitigate the impacts of hold-ins. For example, the garage can pre-assign extraboard operators to open runs, whereas they cannot know in advance which vehicles will break down and assign those to less critical runs. Avoiding standing buses is more difficult, but there are some steps agencies can follow to reduce the incidence of standing disruptions. Similarly, rail blockages can be avoided with internal steps to improve equipment reliability. However, bus route blockages are largely outside the control of agencies, so agencies have few options to avoid these disruptions. Table 3-4 below outlines the different mechanisms for each disruption type.

**Table 3-4: Disruption Avoidance and Enhanced Response Strategies by Disruption Type**

| | Service Gap | | | Insufficient Capacity | Bus/Rail Blockage |
|---|---|---|---|---|---|
| | Hold-in | Standing Vehicle | Insufficient Running Time | | |
| Avoid | Increase resource availability through improved operator and vehicle management | Improve overall vehicle and operator reliability. Improve operations plan reliability. | Improve operations plan. | Reduce average maximum load per trip. | Rail: Improve track and signalling reliability.<br><br>Bus: Difficult to avoid. |
| Mitigate | Prioritize filling critical runs. | Target improved reliability on critical runs by assigning most reliable vehicles and operators to critical runs. | Protect low frequency service. | Make gap vehicle available.<br><br>Closely monitor service quality. | Prior communication with party creating blockage.<br><br>Pre-plan recovery options. |

In addition to understanding which disruption types can be avoided or mitigated in advance, it is also important for agencies to know which policies can contribute to disruption avoidance and reducing disruption impacts and to assess the feasibility and cost-effectiveness of changing these policies. Table 3-5 shows the costs and potential benefits of modifying different agency policies. This can help agencies target the key areas of concern. In addition, the mechanism for changing the policy is noted to give an idea of the ease of changing the policy. For example, work rules are governed by the labor union contract, so these changes would have to result from contract negotiations. However, levels of overall operator staffing, the size of the extraboard, and extraboard report times can be changed at the agency's discretion over time or at the beginning of each pick.

**Table 3-5: Costs and Potential Benefits of Agency Policy Changes**

| Policy Area | Change | Benefit | Cost of Policy Change | Potential Benefit | Type of Change |
|---|---|---|---|---|---|
| **Work Rules** | Flexibility of Operator Use | Improve Response | Negotiated up to value of benefits | Reduce cost per unit of service<br>Increase disruption response options | Contractual Agreement |
| **Policy/ Exclusions** | Hiring/ Promotional Practices | Improve Response | Negotiated up to value of benefits | Increase operator satisfaction<br>Increase disruption response options | Contractual Agreement |
| | Allow Service Restoration Techniques | Improve Response | Retraining | Improve service management | Agency Policy |
| **Training** | Improve Training | Improve Response/ Reduce Disruptions | Program Development and Retraining Staff | Reduce accidents<br>Improve disruption response | Agency Policy |
| **Communi-cations** | More Efficient Use of Existing System | Improve Response | System Development and Retraining Staff | Increase disruption response options | Agency Policy |
| | Increase Capabilities | Improve Response | New Equipment | Increase disruption response options | Technology |
| **Available Resources** | Increased Extraboard, Operator Work Force, or Fleet Size | Improve Response/ Reduce Disruptions | Increased Staff Cost or New Vehicles | Increase service reliability<br>Decrease overtime | Agency Policy |
| | Reallocation of Extraboard or Spare Vehicles | Improve Response/ Reduce Disruptions | None | Increase Service Reliability<br>Decrease overtime | Resource Allocation |

### 3.4.4. Disruption Avoidance and Mitigation

Anticipating, addressing and avoiding disruptions cost-effectively is critical to the satisfaction of transit riders and, therefore, the success of the transit agency. Generally, agencies should aim to eliminate all potential disruptions subject to a cost-effectiveness assessment. Most of the policy areas discussed earlier have the potential for change to reduce the number of disruptions. For example, better training can lead to a reduction in accidents, greater resource availability can reduce the number of hold-ins, work rules could be redesigned to reduce absenteeism and communications could be improved to reduce the number of hold-ins. This section discusses policies that affect the agency's ability to avoid disruptions.

The first line of disruption defense is having extra resources available, including spare vehicles, extraboard operators, and operators available to work overtime for service gaps, insufficient capacity, and route blockages creating either of these problems. Having a sufficient amount of each at the necessary time can reduce hold-ins, and improved maintenance programs and proper vehicle operation can reduce the number of in-service vehicle failures. However, if employing additional resources is not cost-effective, the disruption should be managed once it occurs, following the hold-in priorities discussed in Section 3.3.5 above to mitigate its impacts.

### a) Avoiding Hold-ins

Policies which can affect the incidence of hold-ins include workforce and equipment planning and pre-assigning drivers and vehicles to fill critical runs, as well as managing the schedule, extraboard, and spare ratio to cover open service. The dilemma in service provision is that while the agency needs a certain number of operators and vehicles each day for the morning and afternoon pullouts, carrying more operators than necessary costs the agency money, while not having enough to provide the scheduled service creates service problems for passengers.

#### 1. Managing Operator and Vehicle Availability

The agency would like to cover its scheduled service, without having idle operators or idle vehicles, however, the daily variability of absences and equipment availability makes this difficult. While the ideal situation may be unattainable, there is a window for which the agency can aim. It must cover a minimum number of critical runs, but not have chronically unproductive operators to pay or vehicles to maintain. Proper operator and vehicle management and scheduling can avoid or mitigate the negative impacts of disruptions, while limiting agency cost.

There have been several comprehensive studies on how to determine and manage the workforce and extraboard planning, scheduling, and size that have incorporated service reliability and agency cost. Koutsopoulos and Wilson (1987) developed a framework and set of methods to assist agencies in determining the appropriate operator workforce size and how to assign operators. Shiftan and Wilson (1993) developed a workforce planning model that incorporated overtime, absence, and reliability concerns.

There are five key times at which service managers should review their scheduled operators and vehicles for a given service day. Each of these reviews can help the agency get closer to having all service covered, without scheduling more operators than needed. The first three are well in-advance of disruptions without a particular work day in mind, the fourth is prior to each work day, and the last is at the beginning of the pull-out period.

There are three important resource planning decisions in the long-term. The first is the workforce size decision, which can be reviewed quarterly or annually by the Human Resources or Personnel department with input from the Budgeting, Service Development and Operations departments. The second is the annual assignment of operator vacations and vehicle overhaul scheduling. Agencies often provide different levels of service by season based on varying demand levels. For example, ridership dips at most agencies during the summer because school is not in session and families take vacations, reducing the number of students and commuters making school and work trips, respectively. Therefore, this may be a good time to have operators take vacations and to overhaul vehicles. The third long-term decision occurs after operators have selected their runs for the upcoming service season. At this time, management can see where there will be chronically unstaffed runs and assign extraboard operators accordingly.

In the short-term, specific open runs must be filled. A day in advance, all scheduled absences for vacation, training and other purposes are known and vehicle maintenance schedules have been determined; remaining unknown absences are limited to sick outs and misses. Prior to each pull-out, garage staff will know the anticipated short falls in service coverage and can decide to call additional operators, defer vehicle maintenance, or identify candidates for hold-ins.

Agencies should determine how well each of these key workforce planning and assignment decisions are being made. They should also monitor their workforce costs and usage to ensure they have the proper workforce size and are allocating their hours properly.

### 2. Extraboard Pre-Assignment

Pre-assignment allows the agency to identify unfilled critical runs sooner, which can allow for better extraboard and spare ratio management in real-time when operators or vehicles are needed. The agency should pre-assign runs to improve service quality, reduce the number of unproductive paid hours for operators, increase vehicle utilization and improve vehicle maintenance scheduling.

As mentioned above, most types of operator and equipment shortages are known in advance and garage management can make plans for the entire service period, each week and day to avoid or limit the shortages. At the beginning of each service season, garages can pre-assign extraboard drivers to critical runs that are unfilled for the entire pick and develop a list of runs that are hold-in candidates. At the beginning of each week or one day in-advance, they can review the list of temporarily unfilled runs, such as those for operators on vacation, in training, or other scheduled leaves, and assign available operators to the priority runs. The morning of service, as unscheduled absences are called in or operators miss their report times, any remaining operators can be juggled among the newly unfilled runs.

### 3. Service Reassignment Flexibility

Labor contracts typically prohibit or tightly constrain operator reassignment from their chosen run, but are likely flexible enough to allow having the operator cover trips on a critical run, then return him to his regular duty. For example, it may violate the contract to reassign the operator's entire run, but the service manager may be able to borrow him during the peak period after which time service managers expect other resources to become available. When other resources are found, the service manager can send the operator back to his assigned run and have the new operator cover the disrupted run.

Typically, part-time operators can be reassigned to full-time runs, but are limited in the number of hours they can work. Vehicle assignments, however, are very flexible and can be made according to the priority list of runs. Therefore, operator shortages create more difficult disruption problems because these are primarily full-time runs. The inflexibility of reassigning operator run assignments makes properly sizing the extraboard more critical to providing good service than having the right number of vehicles. Additionally, operators must be paid if they are scheduled to work, regardless of whether they provide service, but vehicles do not cost the agency if they are parked in the garage.

**b) Avoiding Standing Vehicles**

While standing disruptions are very costly to passengers and their elimination would greatly improve passengers' experiences and operator utility and reduce field-based maintenance costs, they stem from problems that develop for operators, vehicles or the right-of-way while in-service. Thus, they cannot be avoided in the same manner as hold-ins, although there are still several strategies to reduce the frequency of this problem. The primary disruption avoidance option is to develop initiatives to improve maintenance in the area of equipment failures and safety training for operators to decrease accident rates. However, these efforts may be expensive, particularly if the agency is already filling the critical runs and already has a low accident rate.

**c) Rail Blockage Avoidance**

For rail service with exclusive right-of-way, vehicle or right-of-way equipment failures typically create most line blockages. Thus, the primary way to avoid rail blockages is to properly maintain vehicles and right-of-way equipment. Rail equipment reliability is more critical than bus reliability because the passenger impacts of a train failure are much greater. When trains or track equipment fails, it affects more people for longer because following trains cannot pass standing vehicles, whereas properly managed bus failures typically only affect passengers on the disabled and immediately following buses. Training can also help operators respond to equipment problems before they create track blockages that result in lengthy passenger delays.

For systems without exclusive right-of-way, such as light rail, the agency will also need to monitor blockages created by external parties and determine to what extent they can be avoided. These impacts may be more like the bus case, where the agency cannot avoid them, but they can take steps to pre-plan how to respond to the disruption.

### d) Improve Operations Plan

While the prior disruption avoidance sections have all described how to avoid disruptions of different types, this section describes a mechanism for avoiding disruptions for multiple disruption types. Recurring disruptions may be a sign to review the operations plan and consider making changes to improve reliability. More robust scheduling solutions should be considered if recurring disruptions develop because of a general inability to execute the operations plan properly, such as in the case of insufficient capacity or running time, chronically late or missing interlining service and reliefs. Although the change may cost more on paper, the true passenger impacts and cost is based on the actual performance, not what is scheduled. If the proposed plan is more reliable or less expensive when compared to the actual performance, the agency should change the operations plan.

One caution is that service managers may want the additional slack time or capacity in the schedule, whether or not changing the operations plan is a cost-effective way to avoid the disruption. Changing the operations plan is a structural change that often comes at substantial cost and should be made only if there are clear passenger and/or agency benefits.

### e) Mitigate Passenger Impacts of Standing Vehicles

Although disruption impact mitigation was discussed in Section 3.3.5, this subsection pertains specifically to the reduction of disruption impacts achievable through agency policy change. As discussed in Chapter 2, unfilled runs and standing vehicles both create essentially the same operational problem: missing service. Therefore, the same types of runs that are difficult to manage if unfilled are also difficult with standing vehicles. However, there is little agency discretion over which runs fail while in service and thus become standing vehicles.

### 1. Assign Reliable Equipment to Critical Service

There are two ways to try to mitigate the impacts of standing vehicles. The first is improving the reliability of critical runs, for which disruptions are the most difficult to manage in real-time. Additionally, there may be problems in the operation plan, such as problematic reliefs and interlines, which agencies should review for more reliable solutions. The second approach is similar to prioritizing critical runs, where the essential runs are filled first. If the common source of standing vehicles is vehicle failures, agencies should assign the most reliable vehicles to critical vehicle blocks. This concept has very limited use, however, if the problem is due to operators not making on-time reliefs or other operator-based problems.

### 2. Initial Operator Disruption Response

The time immediately after the disruption is critical in determining the negative impacts to the passengers on the affected vehicle. Once the disruption occurs, the agency's first response is not through its service managers, it is through the operators. Two policy areas, training and SOP's, as well as pre-planning can work together to improve the initial disruption response.

For example, for mechanical failures, operators could be instructed to first attempt to fix the vehicle and given some very basic training to troubleshoot common or easy to fix problems. If this does not fix the vehicle, the operator must protect the passengers' interests and should alert service management (typically by using the radio to contact the control center or dispatch) and flag down his follower to move passengers to that bus in order to get them on their way, unless otherwise instructed. Another example of using training and SOP's is when facing a route blockage. Operators cannot reroute themselves without the SOP's giving them the authority to do so. These simple steps do not need to be managed by a service manager; clear instructions should be sufficient and will save passengers time.

## 3.5. Service Goals

This section describes the various concerns or service goals service managers should have in mind when responding to disruptions. Froloff, Rizzi, and Saporito (1989) proposed four service goals to guide service delivery during disruptions in a way that mitigates their negative impacts: capacity management, even headways, schedule adherence and personnel management. However, the middle two are both related to minimizing passenger delay, the first being relevant to high frequency service and second being relevant to low frequency service. Thus, in this research three service goals will be used.

Service managers need operations criteria they can calculate quickly and accurately that reflect simple, pre-determined service objectives that best serve passengers and the agency, based on the route's operating plan, passenger demand, and operator needs In order of priority the service goals are: capacity management, passenger delay, and personnel management. The first, capacity management, aims to provide enough capacity on the first arriving vehicles to serve all waiting passengers. While ensuring sufficient capacity is provided should generally be the service managers' primary concern, managing passenger delay concurrently is essential. It is likely that at one point along the route one service goal applies, while in another section of the route, another goal applies. For example, if there is a service gap and the first bus is passing up passengers, the proper service goal is to provide enough capacity. However, there may be other situations along the route where passenger delay or personnel management are the priorities.

The last, personnel management, strives to ensure work rule compliance and get operators to the locations they need to be to make reliefs within work rule constraints. These service goals represent the most important service goals when responding to a disruption and are the basis for the primary measures chosen for evaluating the effectiveness of service restoration techniques.

Schedules are typically written to provide even headways, optimal operator reliefs and hours, and extra trips at the right times to serve large loads, such as trippers serving school start and end times, under normal operating conditions. Thus, schedule adherence is generally the proper goal for non-disrupted service, however, when there is a disruption, it is often beneficial for passengers to have the service spaced at even headways, particularly on high frequency routes. It is also the proper approach to minimizing passenger delay for disrupted low frequency service.

### 3.5.1. Capacity Management

The aim of capacity management is to minimize the number of missed passengers. Being served by the first vehicle is the primary passenger concern because it is inconvenient and annoying to be forced to wait for more than one vehicle. When passengers cannot board the first vehicle and the following service is on time according to the schedule, they may wait an entire headway, in addition to the time already spent waiting for the first vehicle. This severe penalty incurred by not being served by the first vehicle leads to the primary objective, maximizing the number of passengers served by the first vehicle (or group of vehicles).

Capacity problems can be classified as minor or major. There are two types of minor capacity problems. One is not having enough capacity behind a service gap and the second is not having enough capacity to serve all passengers for several vehicles. In the first case, the gap should be narrowed, so that there is enough capacity offered to passengers arriving during the (reduced) gap. In the second case, more service must be provided to carry passengers sooner. Major capacity problems include not being able to serve all passengers for longer periods of time, such as an anticipated demand spike from a large event or being chronically underserved in the schedule.

### 3.5.2. Passenger Delay

The second goal is to minimize the time passengers wait for a vehicle to arrive. There are two ways of serving passengers to minimize their wait time with the appropriate service management strategy depending on passenger arrival behavior.

For low-frequency service, experienced passengers typically arrive at stops based on the schedule in order to minimize their expected wait time. In order to meet passengers' service expectations, schedule adherence becomes the primary service goal because passengers wait less if vehicles arrive at stops and stations according to the schedule. However, if service is frequent, passengers will arrive at stops randomly. In this case, passengers wait less if vehicles arrive at even headways, such that on average passengers are served in half a headway and at worst passengers are served within one headway.

#### Schedule Adherence

When passengers expect to be served at a specific time and plan their arrival based on the scheduled arrival time, maximizing schedule adherence is generally the best approach to minimizing passenger wait times. The schedule adherence objectives are to maximize on-time performance and minimize the number of missed trips. Accounting for the number of missed trips is important because they create long waits for affected passengers.

#### Even Headways

When passengers arrive randomly, as in the case of high frequency service, and more or less evenly within a time period, schedule adherence becomes less important to passengers. In this case, providing passengers with even headways will minimize passenger delay. The objective here is to minimize headway variance and eliminate all service gaps greater than 1½ headways. If vehicles are missing or the running time has been increased, vehicles should be evenly spaced at a higher interval, rather than following the schedule as much as possible, but allowing a gap to develop.

### 3.5.3. Personnel Management

The last service goal is personnel management. This goal, a response to the constraints involved in using operators, addresses the need to comply with all agency work rules and, thus, is to minimize late or missed reliefs and agency cost. It can operate in two ways. The first is that it operates in conjunction with the prior two, such that service managers are thinking about getting operators where they need to be for on-time reliefs. The second is that the need to stay within operator work rule constraints causes the service manager to override one of the passenger-based service goals in order to allow the operator to leave service, service managers should try to avoid these situations because of the resulting passenger impacts.

## 3.6. Evaluation Criteria

The evaluation criteria measure the primary impacts of different disruptions and the effectiveness of different service restoration options. While the evaluation criteria should ideally account for all passenger impacts and agency costs, including impacts on future service, the three primary service goals discussed above emerge as the primary criteria. These measures are the number of missed passengers, the number of passengers with an excessive wait time, maximum passenger wait time, and agency cost, which correspond to capacity management, passenger delay and personnel management. They are used to select among disruption responses.

The evaluation process should consider both passenger impact and agency cost because the goal of the agency is to best serve passengers, within a budget constraint. Therefore, the proper criterion for evaluating service restoration options is centered on the tradeoff between passenger benefits and agency cost. The first subsections identify and quantify the primary impacts to passengers and the agency. Then, secondary passenger impacts are discussed and, last, agency and operator impacts are described.

### 3.6.1. Primary Passenger Impacts

In past studies, the most widely used measure of passenger impact for determining the effectiveness of service restoration techniques has been average passenger delay time over several trips or a time period. However, this averaging process can mask the full impacts of service gaps. The suggested service restoration evaluation criteria in this study focus on mitigating the most severe negative passenger impacts, rather than finding the expressing, holding or short-turning strategy which minimizes expected passenger wait times.

While passengers will not welcome being served a few minutes late with a vehicle more crowded than usual, the most severe passenger impacts that agencies will want to avoid are long passenger waits that include full vehicles passing by waiting passengers. Service should be managed such that the number of passengers not served by the first group of vehicles is minimized both because of the irritation this creates directly and because of the added passenger delay of waiting for the following vehicle, which may be up to a scheduled headway behind. Further, long passenger delays and the uncertainty of when transit service will arrive may encourage passengers to leave transit and find other modes to make their trips. This concept is fundamental, since most disruptions severely impact a small group of passengers and slightly impact many others, as shown in Table 3-6 below.

**Table 3-6: Passengers Affected by Disruption Type**

| Disruption Type | Passengers Impacted Severely | Passengers Impacted Slightly |
|---|---|---|
| Hold-in/ Standing Vehicle | Arrived during gap or following headway. | Others in disruption direction |
| Insufficient Running Time | Uneven headways will create gaps affecting those arriving during gap. | All |
| Insufficient Capacity | At stops where vehicles are full. | Others in disruption direction |
| Route Blockage | Crossing blockage.<br>After blockage. | Prior to blockage |

Estimated primary passenger impacts by disruption type are shown below in Table 3-7. This highlights the likely impacts of each disruption type and the importance of minimizing the number of missed passengers.

**Table 3-7: Passenger Impacts by Disruption Type**

| Disruption Type | Missed Passengers | Max. Passenger Wait Time – No Missed Passengers | Max. Passenger Wait Time – Missed Passengers |
|---|---|---|---|
| No Disruption | No | One Headway | N/A |
| Missing Vehicle | Likely[2] | Two Headways | Three Headways |
| Insufficient Running Time | Likely[2] | One-Two Headways | Two-Three Headways |
| Insufficient Capacity | Yes | N/A | Two or More Headways |
| Bus Route Blockage | Likely[2] | One Headway + Time to Re-route First Bus | Same[1] |
| Rail Line Blockage | Likely[2] | One Headway + Time to Re-route First Train | One Headway + Time to Re-route Two Trains |

[1] Assumes buses bunch at the route blockage point and, therefore, the first group of buses passing the blockage has sufficient capacity.
[2] Depends on average vehicle volume/capacity ratio.

### Missed Passengers

The number of missed passengers depends on the scheduled volume/capacity (V/C) ratio and the actual headway of each bus through the peak load route segments; it becomes an important cushion from missing passengers when disruptions occur. For example, if vehicles are usually 80% full, they can accommodate passengers arriving on a headway 25% greater than average, but if they are only 60% full, on average, they would have enough capacity to accommodate headways 75% greater than average. This ratio is typically highest during the peak service hour, but there may also be capacity problems in the off-peak, if the agency schedules service with demand-based headways and there is less service during this time.

### Passengers with Excessive Wait Times

In addition to determining the longest passenger wait time, it is even more important to know how many people waited an exceptional amount of time. This will gauge how many passengers are severely impacted by the disruption. The agency will need to define what is an excessive wait time, but it might be any wait longer than 2 headways for high frequency service. Passengers using high frequency service can expect to wait up to a full headway during normal operations, but much longer waits should not be expected and may interfere with future transfers or appointments.

### Maximum Passenger Wait Time

Maximum passenger wait time is the longest amount of time any passenger waits to be served. This is the longest headway, assuming a passenger just missed the prior vehicle and no passengers were missed. If passengers could not be served by the first arriving vehicles, the maximum passenger wait time is the longest headway plus however long it takes for a vehicle to arrive that can serve them. For normal service, this should be one headway. For unmanaged service gap disruptions, this is at least two headways. This can be decreased significantly by eliminating missed passengers, as shown in Table 3-7 above by taking the difference between the missed passenger and no missed passengers cases.

Minimizing the maximum passenger wait time works quite well for service with relatively even passenger demand, such as cross-town service, because it should also reduce average passenger wait time. For service with a greater variance in passenger demand along the route or by direction, however, minimizing the maximum passenger wait time may increase the average passenger wait time along the route sections with the greatest number of passengers.

## 3.6.2. Secondary Passenger Impacts

This subsection discusses other passenger impacts that are important, but are not considered primary indicators and, thus, not calculated during the service plan selection process. Although, they are not specifically considered, each are partially reflected in the primary evaluation criteria because they measure similar things.

### In-vehicle Travel Time

While this is considered in the operations control studies of different service restoration techniques discussed in Chapter 2, those studies were concerned with what happened to service while controlling vehicles during a trip, which leads them to evaluate the unintended impacts of service restoration on other passengers. While this is an important consideration for evaluating in-route service restoration techniques, it is not as important for disruptions managed primarily through terminal dispatching, and not relying on spacing the service during the trip. Travel time is assumed to be the typical travel time, given even spacing from the terminal.

### Average Passenger Wait Time

Average passenger wait time is an important measure, but often the difference in average passenger wait time among potential service restoration plans is not compelling. Reducing maximum passenger wait time typically will tend also to reduce average passenger wait time. It is also time- and data-intensive to calculate. Its greatest flaw, however, is that it masks the severe passenger impacts, which are highlighted with the maximum passenger wait time, the number of missed passengers, and number of passengers with excessive wait time measures.

### Crowding

Crowding is another common measure of passenger concern (CTA Customer Satisfaction Survey 1999). It is actually a partially developed case of missed passengers; crowding occurs before passengers are missed. Therefore, this concern is partially reflected in the number of missed passengers measure. The latter measure is more powerful because while crowding is tolerable, being passed up really irritates most passengers. When service is modified to reduce the number of missed passengers, less crowding will also result.

### Passenger Wait Time at Unsheltered Stops during Inclement Weather

The last secondary passenger impact is minimizing passenger wait time at stops with inadequate shelter during inclement weather. Although serving unprotected passengers as quickly as possible is important from a passenger comfort point of view, it is very difficult to accomplish; transit service is fairly inflexible once it has left the terminal.

Once a vehicle has left the terminal and started its trip, it is difficult to express or hard to justify holding.

### 3.6.3. Agency Constraints and Cost

While it is important for service managers to focus on service quality while restoring service, there are agency constraints and costs that should also be recognized in evaluating possible response strategies. This subsection describes agency impacts.

#### Operator Overtime and Regular Wage Costs

Agencies have several sources from which to obtain replacement service. Generally, there is one pay threshold for full-time operators: whether the operator has worked more than 8 hours. Most contracts provide operators with a guaranteed 8 hours of pay each day they work and overtime for any hours worked over 8 per day or 40 per week. The marginal cost of using operators who have not met either of these thresholds is zero. Full-time operators who have met either threshold are entitled to overtime pay, typically 1.5 times their regular wage rate, and are paid accordingly for any extra work. Part-time operator wage structures are typically different. They are not guaranteed 8 hours of pay and earn their regular wage while having worked less than 8 hours of work that day and overtime after they have worked more than 8 hours.

#### Work Rule Compliance

The primary constraints on operator use, however, are not cost, but work rules in the labor contract. These rules typically set limits for shift spread time, minimum rest between days of work, time between reliefs, total hours worked in a day, and the number of hours part-time operators can work. While the agency can choose whether to have operators work extra and pay them accordingly, operators cannot do extra work that violates the work rules.

While part-time operators' work assignments are more flexible than full-time operators' work assignments and their cost is often less for the agency, agencies often have contractual limits on the number of hours they can work, which constrain how they can be used. For example, if there is a cap on weekly hours, operator scheduling staff need to plan their hours carefully. If too many part-time operators' hours are used at the beginning of the week, service at the end of the week may suffer if there are not enough part-time operators available to work.

### Operator Fatigue, Future Absenteeism and Job Satisfaction

There are also operator impacts created when service managers do not manage the disrupted service effectively. When disrupted service is not restored, operators of the remaining service experience increased passenger loads and passenger dissatisfaction with the service. When disrupted service is managed, operators have a modified work schedule, which often is not very different from their regular work schedules, but can be quite different, including making extra trips or working a different route. These impacts are shown in Table 3-8 below.

**Table 3-8: Operator Impacts by Disruption Type**

| Disruption Type | Operator Impact – No intervention | Operator Impact – Service Restoration Intervention |
|---|---|---|
| Missing Vehicle | Service impact for one or two following the gap | Small passenger increase for several.<br>Shifting dispatch and recovery times. |
| Insufficient Running Time | Gaps form and impact following operators.<br>No recovery time. | Passenger increase.<br>Recovery time available. |
| Bus Route Blockage[16] | Gap forms and impacts following operators.<br>Driving on unfamiliar streets. | New driving instructions known.<br>Even passenger loads. |
| Rail Line Blockage[27] | Unfamiliar with operating instructions.<br>Large gap forms. | Familiar with operating instructions.<br>Smaller gap forms. |
| Insufficient Capacity | Full vehicles.<br>Irritated passengers. | Even loading. |

---

[6] Anticipated disruption with re-route planned in advance.
[7] Unanticipated disruption with re-route planned in advance.

Operator fatigue and future absenteeism may be a concern if operators chronically work their days off on overtime, either by choice or as required by the agency. Job satisfaction may decline if full-time operators feel they are constantly called in to work or asked to make extra trips, thus interfering with personal plans and limiting their ability to set their personal schedule. Part-time operators may also feel this way, if they do not have their runs at a consistent time or are working different runs each day.

### 3.6.4. System and Operations Plan Impacts

System impacts are those that propagate from one route to another and operations plan impacts are those that threaten the operations plan. These impacts are important because they are more difficult to track and monitor in real-time and may move from one service manager's district to another's. They are also more complex and require more time to evaluate their full effects. These generally occur with interlined runs, which can result in passenger service problems for other routes. Operations plan impacts include missed or late reliefs and the difficulty of getting back to the regular operations plan once the disruption ends; they become personnel management problems. They do not have their own performance measures, but can be quantified by their effects on the previously defined impact areas.

## 3.7. Summary

As shown in this chapter, the central service restoration problems are the service restoration plan development and implementation processes that occur in real-time. Agencies can improve service by reducing the number of disruptions altogether or reducing the disruption's impacts. In addition, developing the disruption response in-advance to the extent possible allows service managers to spend their time and energy on response implementation.

The primary considerations shaping the disruption response were introduced and frameworks for how to approach disruption avoidance and improved disruption response were developed. A disruption response support framework was developed to select which disruptions are the best candidates for pre-planning. The criteria for determining the proper mix of real-time and pre-planned disruption response decision-making was presented. Much of the decision-making for hold-ins and rail blockages can be done in advance, while standing vehicles responses must generally be determined in real-time. An agency policy evaluation framework was developed to highlight policies that can improve disruption response or avoid disruptions.

In the last two sections, service goals and primary indicators of service restoration effectiveness were discussed in detail and four primary indicators selected. They are the number of missed passengers, the number of passengers with an excessive wait time, maximum passenger wait time, and agency cost.

The service restoration context and evaluation criteria developed in this chapter will be used in Chapter 4 to develop the specific processes for service restoration, pre-planning, and policy evaluation. These processes are then applied to the CTA bus system in Chapter 5.

# Chapter 4: Disruption Response Process

This chapter develops disruption response and agency policy evaluation processes to help agencies avoid disruptions or improve service during disruptions. The first two processes, disruption response and disruption response support processes, provide service managers with the critical passenger and agency impacts of each potential service restoration option from which to develop a disruption response plan. The last, the policy evaluation process, is aimed at identifying which agency policies should be evaluated for change. Policies must be evaluated in terms of the net benefit of specific changes. The specific assessment will vary across policies.

The processes are designed to work for both bus and rail transit service restoration. Where background information of concern and available service restoration techniques are different this has been noted, but the concepts and disruption response planning process are similar. The concepts and evaluation criteria developed in Chapters 2 and 3 serve as the basis for these processes.

## 4.1. Overview

The first two processes provide service managers with a systematic approach both to evaluating the various options available for service restoration and to what should be done in advance to ensure the best decisions are made in real-time. If they are not followed, service managers will likely make decisions based on personal perceptions of the route and perceived passenger and agency impacts, rather than on the real impacts. This can lead to making decisions that do not create the desired result or reflect the agency's priorities. The third process is essential because it gives agencies a way to think about how policy directly affects service provision and the tradeoffs involved in changing a policy are identified.

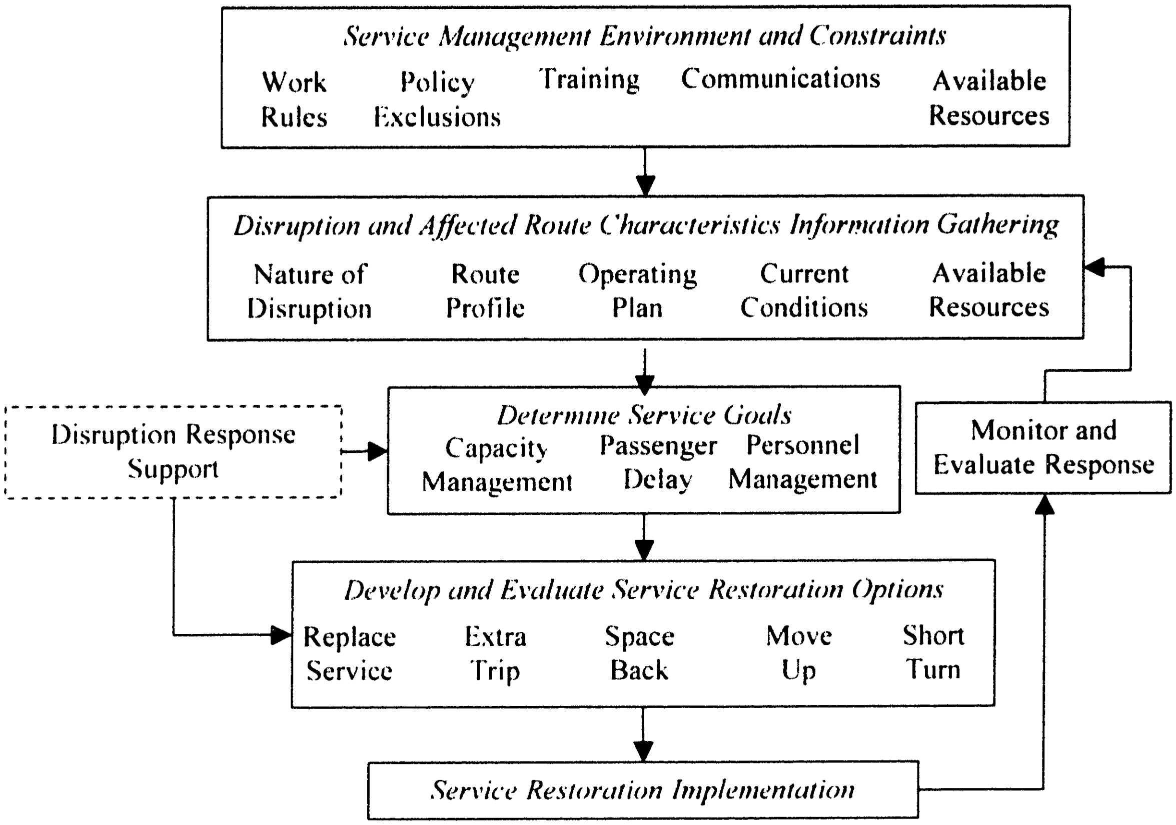
The first process describes the steps involved in selecting the most appropriate response in response to a disruption. The second process helps agencies identify which disruptions are amenable to pre-planning and the proper mix of real-time and pre-planned decision-making for different disruption types. The last outlines the steps of reviewing agency policies, technology and resource allocation to evaluate the cost-effectiveness of changing them in order to reduce disruptions or improve the potential responses.

The third process takes a step back from service restoration and evaluates how agency policies support and constrain the ability to avoid disruptions and service restoration, and what the impacts of changing the policy would be. Policy review is important because service managers do not consider these to be changeable in their day-to-day work when responding to disruptions. Because of the need to consider only responses that work under current policies, they automatically eliminate options that are excluded by current policy. Therefore, service managers, while being intimately familiar with current service restoration techniques, may not consider how service restoration could be improved if some of these policies could be changed. Likewise, agency policymakers may not be in an ideal position from which to evaluate policy impacts on disruption response because they may not have the service restoration expertise to see opportunities for improvement through policy changes or they may not have disruption response quality on their work agendas.

## 4.2. Disruption Response Process

The disruption response process is the generic set of steps that should occur when responding to a disruption. When developing service restoration responses, existing agency policy, such as the communications system, distribution of management responsibility, and work rules, must be viewed as constraints, which limit the feasible range of responses. Therefore, agency policies combined with disruption and route characteristics, current conditions, and available resources determine the service restoration options available to service managers, as shown in Figure 4-1.

**Figure 4-1: Disruption Response Process**




The primary goal is to begin actively managing service as soon as possible to minimize passenger impacts and control agency costs. To this end, service managers must implement a response as soon as possible. They must also identify current and changing conditions that could lead to future problems in the implemented service restoration plan, such as making operator reliefs and not creating additional service problems on the disrupted route or other routes. Lastly, they must develop a plan to return to the scheduled operations plan when possible.

### 4.2.1. Agency Policy Environment

The policy environment frames how service managers respond to disruptions. In this context, policies influence the disruption response decisions that are made, but are not changeable, as reflected in the feedback loop in Figure 4-1. Agency policymakers must ensure that the agency values for balancing passenger impacts and agency cost is clearly conveyed to service managers, so they have a clear picture of how to balance their objectives.

### 4.2.2. Route Characteristics and Disruption Information Gathering

The first step in responding to a disruption is information gathering. As described in Chapter 3, this includes collecting information about the affected route(s), the disruption itself, current conditions, and available resources. Service managers need to familiarize themselves with the particulars of the specific disruption, current conditions, and any historic information known about the route.

The service manager is ultimately interested in determining and implementing the appropriate disruption response. Before he can decide what that is, he must determine the appropriate service goals and the feasible service restoration options from which to choose his response. Route and disruption information are important because they determine the response options available to service managers and are used to identify potential service problems, such as where capacity problems might develop.

Service managers should not need to spend very much time, if any, gathering route information once the disruption has occurred because it does not change dramatically from day to day. Decisions made in advance are discussed in greater detail in decision response support process in Section 4.3. A summary of needed information and when it should be known is shown in Tables 4-1 and 4-2.

**Table 4-1: Route Characteristics**

| Topic | Information | When Known | Determines |
|---|---|---|---|
| Infrastructure and Operating Environment | Signal priority<br>Number of lanes<br>Queue jump lanes<br>Traffic conditions<br>Vehicle speeds | In advance | Service restoration techniques |
| Demand Profile | Load profile<br>Maximum load point<br>Average peak load<br>Average volume/ capacity ratio | In advance | Service goals<br><br>Service restoration techniques |
| Operating Plan | Headways<br>Running Time<br>Recovery Time<br>Relief points<br>Relief times<br>Interlining | In advance | Service goals<br><br>Service restoration techniques |

**Table 4-2: Disruption Characteristics**

| Topic | Information | When Known | Determines |
|---|---|---|---|
| Nature of Disruption | Disruption Type<br>Disruption Timeframe | At Disruption Notification | Service goals<br><br>Service restoration techniques |
| Current Conditions | Vehicle locations<br>Passenger loads<br>Weather | Estimated in advance.<br>Confirmed in real-time. | Service goals<br><br>Service restoration techniques |
| Available Resources | Additional<br>Reallocated<br>Pieces of Service | Estimated in advance.<br>Confirmed in real-time. | Service restoration techniques |

The service goals are determined by the primary passenger needs, which are determined by the nature of the disruption, the demand profile, the operating plan, and current conditions. The feasible service restoration options are determined by the service goals, demand profile, operating plan, nature of the disruption, current conditions, and available resources.

### 4.2.3. Determine Service Goals

Once key route and disruption characteristics have been identified, the next step is to determine the primary service goals given the disruption, route, vehicle location, current conditions, and passenger demand information available. As described in Chapter 3, the three service goals, listed in typical order of priority, are capacity management, passenger delay and personnel management. While providing enough capacity to serve all passengers with the first arriving vehicle is important and generally the overriding concern, it will generally not be the only goal. Minimizing passenger delay is also important and, indeed, at times it may be more beneficial to miss a few passengers in order to substantially reduce passenger delay. When making decisions in real-time, service managers have neither the time nor the analytic resources to predict passenger delays and act accordingly. Adopting service goals can simplify the analysis and response process. Once the proper service goals are determined, they will guide which disruption responses are considered.

Service managers should first determine whether there is enough capacity to serve passengers with the first arriving vehicle (or set of vehicles). If there is not enough capacity, they should then determine whether focusing exclusively on the capacity objective may negatively affect their ability to minimize passenger delay. If not, the capacity problem should be the priority. If so, a solution that serves both goals should be sought. For example, extra capacity could be added at the constrained point and more even headways for high frequency service or improved schedule adherence for low frequency service could be provided elsewhere along the route.

In addition to providing service that satisfies passenger interests, service managers must make sure that no work rules are violated and contain agency costs. Therefore, they must also anticipate work rule conflicts that will need to be addressed, and look for opportunities to move operators around in advance to avoid negative passenger impacts and taking operators out of service, even at the expense of passenger inconvenience.

### 4.2.4. Disruption Response Assessment

Once the type and nature of the disruption and the proper service goals have been determined, the service manager must develop a disruption response plan. Service managers should first determine whether no-cost replacement service, such as a scheduled gap vehicle and operator, is available because it would reduce passenger impacts without increasing agency cost. If this is unavailable, the service manager must evaluate his other feasible options. The evaluation criteria values are calculated or estimated to the extent possible, for each of the feasible service restoration options. The service manager then compares the values and selects the one that best reflects the agency's desired mix of service quality and agency cost.

In general, no-cost replacement options will be rare, unless they are used regularly for different problems each day, because providing these resources costs the agency money, whether or not the service is used. If no-cost service is available, the service manager must decide whether this is likely to be the most disruptive problem to develop while it is committed to addressing the current problem. If he believes it is, the service manager should use it. If no-cost replacement service is unavailable or needed for larger service problems, the service manager should calculate or estimate the values of each of the three primary evaluation criteria for each of the feasible service restoration options.

This process assumes the complete set of feasible service restoration techniques given the nature of the disruption and the service goals is known or that there is a reference manual with this information available when developing a disruption response. Infeasible service restoration techniques and those that do not serve the service goal are not considered. The matching of service restoration techniques and disruption types was discussed in Chapter 3 and reviewed below in Table 3-1.

**Table 4-3: Service Restoration Techniques by Disruption Type**

| Service Restoration Techniques | Disruption Type: Service Gap | Insufficient Capacity | Route/Rail Blockage |
|---|---|---|---|
| Replacement/Extra Service | + | + | + |
| **From Terminal** | | | |
| Respacing[8] | + | 0 | + |
| **In-Route** | | | |
| Limited Stop/Drop-off Express | + | + | 0 |
| Holding | + | 0 | 0 |
| Short Turn | + | + | 0 |
| Re-route[9] | N/A | N/A | + |

+ Improves service.   0 No service effect.

Once the service goals have been determined and feasible restoration techniques identified, the service manager develops a disruption response plan by using service restoration techniques on different vehicles to create the desired vehicle spacing for high frequency service and, thus, passenger loads and passenger wait times.

For example, insufficient capacity on a high frequency route requires that more service be provided. Possible solutions to this are adding more service, holding vehicles in the disrupted section of the route until they are at capacity, and/or expressing vehicles behind the disrupted section to the disrupted section, or some combination of the promising feasible options discovered when evaluating the techniques individually.

The number of missed passengers, maximum passenger wait time, passengers with an excessive wait time, and agency costs should be calculated for each of the feasible options. Additional agency cost can be estimated from the number of additional part time operator hours worked plus the additional full-time operators hours worked on overtime (hours above 8 hours that day or 40 hours that week). When pre-planning, options that are inferior to other available options should be eliminated from the feasible set considered by service managers in real-time.

---

[8] High frequency service only

If the disruption is anticipated to last longer than one trip, a response to deal with the immediate disrupted trip as well as future trips must be developed. Disrupted trip responses need to use in-route service restoration techniques, holding, expressing, and short-turning, while long-term responses should maintain service under a new operations plan. These options include respacing the service from the terminal for high frequency service and adding or replacing service and improving schedule adherence for low frequency service, as well as short-turning or expressing selected runs.

### Immediate Disruption Response

Responding immediately is important because the remainder of the first disrupted trip is where passenger impacts will be the greatest, especially for service gaps and rail blockages. Service gaps will be the largest during this trip because of the additional delay created for the follower.

A brief example is presented to illustrate the disrupted trip service restoration selection process. For service gaps on high frequency service, the supervisor's primary goal is to reduce the size of the service gap to provide enough capacity and to reduce passenger delay. Several in-route service restoration options and combinations of options are available to reduce the gap. If the gap is not expected to create capacity problems, a typical solution is to hold the leader and express its follower to diminish the gap. Other possibilities that should be evaluated before selecting this include adding extra service or, when there is a clear capacity problem, using short turning with holding and expressing. The route profile will determine which techniques are most appropriate.

Likely solutions for the other two disruption types are briefly listed. Insufficient capacity disruptions are likely to be best addressed with a combination of adding service and holding, expressing, or short-turning other service on the route. Past studies have shown that lengthy rail blockages are best served by holding trains in front of the blockage when disruptions will create limited capacity problems or a combination of holding and short-

---

[9] Route/rail blockages only

turning for disruptions producing more serious capacity problems (Ortiz, 2000; Puong, 2001).

### Maintained Disruption Response

Disruptions lasting longer than a trip will create the need for a modified operations plan during the disruption. There are several options available for high frequency service: adding extra service, respacing service from the terminals to provide even headways, routinely expressing every other or every 3rd or 4th bus along specific route segments, or routinely short-turning every other or 3rd or 4th bus from a certain point to provide service. The first option provides extra service to counter the disruption. The second provides a reduced amount of service evenly spaced along the route. The last two concentrate the remaining service where it is most needed.

A brief example of developing and maintaining a new operations plan is given for a service gap. To maintain the scheduled service headways in the primary direction of travel, every third bus is short-turned into the gap at the beginning of the route in the primary direction of travel to keep the gap on the small portion of the route at the end of the secondary direction and beginning of the primary direction because there is relatively little demand on this section. This can be maintained until the extra capacity is no longer needed, leading the service manager to respace service from the terminals, or until the disruption ends, so he can return service to the original operations plan. One caution is that short-turning moves operators out of their scheduled positions and the service manager will need to deal with this at some point in the next couple of trips.

## 4.2.5. Disruption Response Implementation

Once the service manager has identified the best response to the disruption, he must implement the plan. There are four phases of service restoration: initial disruption response, transition to the new plan, maintain the new plan, and transition back to the original plan, which were introduced earlier.

The first stage is the initial response, which includes the initial service restoration steps taken to contain the negative passenger impacts from the time the disruption occurs until the affected vehicles complete their trips. In-route service restoration techniques are used to contain service gaps, find additional or reallocated service quickly to address capacity problems, or provide re-route instructions.

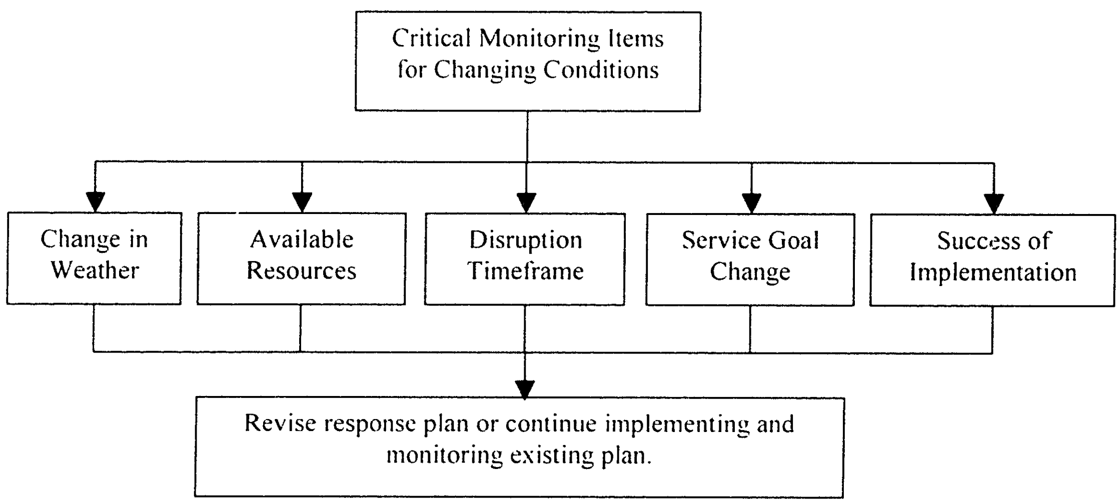
The second stage, if necessary, is moving the vehicles into new positions to provide adequate service on future trips. This involves terminal-based service restoration techniques, such as respacing, short-turning, and deadheading. Once the service is operating under the new plan, the service manager will need to confirm that service has successfully transitioned to the new plan, instruct operators joining the route (those coming off a relief, moving from another route as part of a scheduled interline, and garage pull-outs), and monitor the route for service degradation, as during normal operations. He will also need to monitor the agency cost impacts and work rule compliance, and may need to move operators around or take them out of service due to those constraints. Lastly, once the disruption has finished or permanent replacement service has been found, the service manager must determine the least disruptive way to transition operators with vehicles back to the original operating plan.

One of the challenges during implementation is to develop and communicate specific instructions for each operator in advance. Giving instructions too far in advance may reduce the service manager's flexibility for future changes, make the transition back to the original operations plan more difficult, or confuse operators. Additionally, the process assumes the service restoration plan is implemented in harmony, such that one service plan is decided upon and only instructions that support that plan are given to operators, which is not always the case.

### 4.2.6. Monitor and Reassess

The last steps of the disruption response process are the two feedback loops. The first is real-time feedback to identify changing conditions that will affect the disruption response, and the second is to understand the effectiveness of current service restoration practices by monitoring and evaluating service managers' disruption responses. Throughout the four phases of disruption response, service managers must look for changing conditions that threaten the plan they have devised and implemented, identify and make changes that will reduce the negative passenger impacts, or provide an opportunity to resolve the disruption. As shown in Figure 4-2, the primary concerns are changes in the weather, passenger demand, available resources, anticipated disruption timeframe, service goals, and the success of the implementation.

**Figure 4-2: Real-Time Monitoring and Reassessment Process**




Agencies can also improve the disruption response effectiveness through monitoring and providing feedback to service managers. Through this information and reflection process, they can learn from other service managers' responses, as well as their own. Without proper information and feedback, they will not learn the lessons from the implemented plan, and will likely implement the same plan the next time this type of disruption occurs, as long as it did not result in serious problems. Thus, service managers' responses are likely to be what they perceive worked adequately to resolve the problem.

If assessment information is limited, other mechanisms, such as ridership, complaints, or customer satisfaction surveys could be used as very rough tools to track progress and effectiveness, but are not recommended substitutes for gathering this information directly.

## 4.3. Disruption Response Support Process

While the first process identified the steps service managers should take to develop and implement the best service restoration plan, the second process, disruption response support, considers disruption situations in advance to ensure service managers can implement an effective response despite the constraints on time and information inherent in real-time disruption response. The aim of this process is to enable agencies to identify specific disruption types for which responses can be effectively pre-planned (partially or totally) in order to increase the likelihood that the service manager will be able to implement the best disruption response. Thus, it is designed to identify what information service managers should know in advance, which disruption types would benefit most from pre-planning, and the extent to which different disruption types should be pre-planned. Once responses have been considered for pre-planning, they can be pre-planned to the extent practical by applying the disruption response process in advance. The findings should be readily available to and used by service managers when disruptions occur. The support process steps are listed below.

1. Determine what information should be known in advance to improve real-time disruption response.
2. Determine which disruption situations should be considered for pre-planning and the appropriate mix of real-time and pre-planned decision-making for each scenario.
3. Follow the disruption response process to pre-plan elements of or entire disruption response, as appropriate.

### 4.3.1. Pre-planning Overview

Service managers will need to, partially or fully, pre-plan responses before certain disruptions occur if the response is to be effective. As described in Chapter 3, the purpose of pre-planning disruption responses is to reduce the difficulty of the real-time disruption management problem by identifying effective responses beforehand where possible and improving the implementation process. Pre-planning recognizes that service managers generally do not have the necessary information or time to develop the best possible service restoration plan in real-time. Additionally, as time passes, passengers and agency cost impacts increase, which means there are likely to be better responses available if service managers spend less time developing their response and focus on the response implementation sooner.

While the agency does not have current conditions or available resources information when pre-planning disruption responses, reasonable assumptions can be made about weather, traffic conditions and available resources, depending on what are the important cases to pre-plan.

### 4.3.2. Know Route Characteristics In Advance

The first step in the disruption response support process is providing service managers with route information and expecting them to know it in advance of a disruption. Knowing in advance route profile information, the operations plan, and after what points holding, short-turning and expressing are no longer useful for passengers by route, time of day, and direction, is critical because this information does not change and it eliminates the need to spend time determining them.

Typically, service managers, whether centrally- or field-based, work with the same routes during the same times of day, so they need the same information on a daily basis. While service managers should be expected to know the basics off the top of their head, this information should also be available in a reference, so the specific values are available if service managers need more detailed information or if they are managing unfamiliar routes.

If he does not know this information in advance or have it available in a reference, he will need to estimate it or select a more conservative disruption response that does not depend on having accurate estimates of this information. While he can quickly look up or estimate the route profile and operations plan information, he will need some time to think through and determine the appropriate locations for using each service restoration technique. He may ultimately determine them incorrectly due to making errors in the primary information.

For example, traffic congestion and passenger demand may change slowly over time, but generally stays the same. Thus, service managers should know all route characteristic information in advance of the disruption or have it readily at hand while managing service. In addition, service managers should also anticipate how they would respond to different disruption types. They should have rules of thumb for managing disruptions by disruption type, anticipated disruption length and service gap size (if applicable), which could be simple thresholds, such as knowing the maximum service gap this route can sustain without leaving passengers behind in each direction by time period. Only some basic disruption information would need to be collected or estimated at the time of the disruption.

### 4.3.3. Identify Disruptions for Response Pre-planning

The second step is to determine which disruption scenarios are good candidates for response pre-planning. As discussed in Chapter 3, there are three criteria for selecting disruptions to pre-plan: disruption type, frequency of occurrence and passenger impacts. The mix of these three elements determines whether it is worthwhile for the agency to pre-plan a response because each disruption type has a different set of service restoration options and potential to be addressed in advance, and frequency of occurrence and passenger impacts are an assessment of the extent of problem created for passengers.

While it is important for agencies to prioritize the worst disruptions scenarios, the particular cutoff at which the agency stops pre-planning disruption scenarios is not important because there are so many that can be helpful to have some pre-planning. Additionally, once a scenario is pre-planned, it should not need to be revised if only minor schedule changes occur. A substantial change to the operations plan, such as the number of vehicles operating on the route, would warrant a second look, though.

For those disruptions warranting pre-planning, service managers must determine to what extent the disruption response can be effectively pre-planned. The framework in Chapter 3 describes the gamut of pre-planning options, ranging from full pre-planning with no real-time decision-making to no pre-planning with all real-time decision-making. Disruption response elements that can be pre-determined given a disruption scenario include identifying the appropriate service goals, the feasible service restoration techniques, whether extra service is essential, and how responses should be implemented. Table 3-3 below indicates the level of pre-planning that is practical for each disruption type depending on how often and severe the impacts are.

**Table 4-4: Real-Time / Pre-planned Decision-Making Mix**

| | | **Disruption Type** | | | |
|---|---|---|---|---|---|
| **Frequency of Occurrence** | **Passenger Impact** | Hold-in and Rail Blockage | Insufficient Capacity or Running Time | Bus Route Blockage | Standing Vehicle |
| Common | Severe | F | F | P | P |
| Common | Moderate | F | P | P | G |
| Uncommon | Severe | P | P | P | G |
| Uncommon | Moderate | P | G | G | R-T |

**F** – Fully Pre-plan  **P** – Partial Pre-plan  **G** – Guidelines  **R-T** – Real-Time

These categorizations should be intuitive given the characteristics of the disruption. For example, hold-ins can be controlled to a great extent, especially if anticipated in advance, because garages control vehicle and extraboard operator assignments and can choose to hold-in runs that are not as critical to successful operations. Rail blockages create tremendous passenger impacts because of the high volume of passengers and the inability of trains to pass the route blockage. Thus, blockages on each route segment by time of day and direction must have a pre-planned response for service managers to follow. Standing vehicles, on the other hand, are difficult to anticipate, control, or respond to. Much of this response development must be done in real-time.

### 4.3.4. Pre-plan Disruption Response

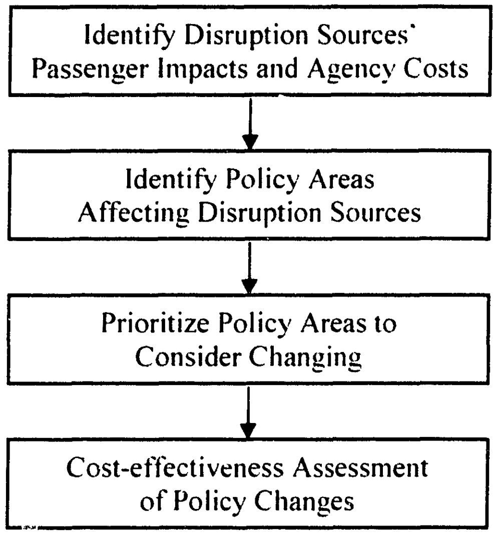
Once the extent of the pre-planning has been determined, the last step is to pre-plan the elements of the disruption response worth pre-determining or the entire response, if warranted. This is done through the disruption response process described in Section 4.2, making assumptions about the disruption, as necessary. Response guidelines and partial or full pre-planning responses should be developed from this pre-planning process and compiled as a reference for service managers and other agency personnel.

## 4.4. Agency Policy Evaluation Process

While there will always be disruptions, modifying agency policies may help to reduce the incidence of disruptions and/or improve disruption response. The agency policy evaluation process is intended to identify which policies should be considered for further analysis. Once policies are identified for further review, the specific cost-benefit analysis depends on the type of policy considered. For example, evaluating the effectiveness of the types and level of training programs requires a totally different analysis than how to set the extraboard levels.

While the framework in Chapter 3 focused on how policy changes could reduce the impacts of disruptions through reducing their incidence and improving the disruption response, the agency policy evaluation process develops a way for agencies to identify which policies should be considered for change. The first subsection develops a process to identify the passenger impacts and costs of each disruption source. The second identifies which policies could contain the disruption sources. The third prioritizes the disruption sources, and the last proposes a cost-effectiveness approach to evaluating policy changes. The process includes four steps, as shown in Figure 4-3.

**Figure 4-3: Policy Evaluation Process**




### 4.4.1. Identify Disruption Sources' Passenger Impacts and Costs

In order to target and improve service reliability problems, the agency must determine its primary disruption sources and their impacts. The agency must understand where and how it is experiencing service problems in order to effectively improve service reliability. Therefore, agencies should collect passenger impact and agency cost information by disruption source. Some agencies keep this information manually, making it difficult to develop any summary information, but others enter and store this information electronically, which allows for analysis and report generation.

Reliability problems may stem from not having enough operating vehicles or operators to operate all scheduled service, in-route equipment failures or incidents, missed reliefs, insufficient headways to carry all passengers, or right-of-way failures or blockages. As shown in these examples, disruptions are created by both internal and external sources. The agency must analyze and address internal sources, as well as manage its externally-generated disruptions through recognizing how it can work with other agencies and the public to reduce the negative impacts and how it can respond internally to cope with the external source disruptions.

### 4.4.2. Identify Policy Areas Affecting Disruption Sources

Once the impacts of each disruption source are known, the next step is to map the disruption sources to the policy areas. This will help agencies bridge the disruptions they are experiencing with the policy areas that may affect the disruption. From drawing general connections, agency policy-makers can narrow their focus when determining which policies are contributing to or mitigating the amount and effects of different disruptions.

For example, the time needed to move vehicles where they need to be to respond to the disruption will always be a problem, but agencies may be able to stage extra vehicles at critical locations near several routes' peak load points. Additionally, while real-time vehicle location information is becoming increasingly available for transit agencies, the availability and accuracy of information from which to make decisions is a problem. These problems may be reduced by changing different policy areas. The first could be improved through resource availability and the second could be improved through communications and information technology applications.

Another area of interest is extraboard productivity. Agencies should collect specific information about extraboard usage by day of week and by season to evaluate the benefits of increasing the number of operators scheduled compared with increasing operator reassignment flexibility through work rule changes to allow supervisors more freedom in what work operators do during disruptions. They may find that there is no benefit because operator use is not a critical constraint, but it is important to know whether this is a problem.

### 4.4.3. Prioritize Policies

Once the disruption sources are identified and their impacts quantified, policy makers can prioritize which policies to consider. This can be done by estimating the cost-effectiveness of expected passenger impact and agency cost improvements. Agencies can then choose which policy evaluations they want to pursue. While most cases will have fairly clear benefits and limited cost, some may have an increase in both or have uncertain cost and benefit estimates.

### 4.4.4. Assess Cost-Effectiveness of Policy Changes

The prior steps have identified which policies the agency should evaluate first. The last step is to estimate the costs and benefits of a policy change using the specific evaluation processes developed in other studies. While this study has focused on how to improve service during disruptions, it is important to note that most agency policies are designed to influence both regular and disrupted operations. When the costs and benefits of changing policies are calculated, it is important to include the impacts during regular operations.

## 4.5. Summary

Based on the frameworks and criteria developed in Chapter 3, this chapter developed three processes to avoid disruptions and improve disruption response. The first process addresses the core disruption response problem. It identifies the information service managers need and the decisions they must make to develop an effective disruption response in advance or in real-time. The second process is designed to support the first. It identifies which disruption situations should be pre-planned and to what extent they should be practically pre-planned. The last, an agency policy evaluation process, is removed from the core disruption response problem and helps agencies prioritize which policies they should consider changing in order to avoid more disruptions and improve disruption response.

# Chapter 5: Chicago Transit Authority Bus Application

This chapter applies the three processes developed in Chapter 4 to the Chicago Transit Authority's (CTA) bus system. A specific case, the Route 77 Belmont, is analyzed to show how service managers and agency policymakers can use the three processes to improve disruption response. Specific scenarios were selected to illustrate different disruption types and response strategies, address service managers' key concerns about disruption response, and highlight the key steps and considerations in this process.

The first three sections provide an overview of CTA policies affecting disruption response, develop the route profile of the route used for this study, and describe the disruption scenario considered. The fourth develops the disruption response options available under ideal conditions. The next two sections show the disruption response and disruption response support processes under CTA constraints. The seventh section describes the agency policy evaluation process and identifies policies for further analysis, and the last section summarizes the chapter's findings.

## 5.1. Agency Policy Overview

This section discusses current CTA policies and practices that directly influence disruption response, as described in Chapter 3. Policies provide service managers both opportunities and constraints. They apply to every service restoration situation, regardless of the characteristics of the specific disruption and the service on which it occurs.

### 5.1.1. Work Rules

CTA's governing work rules pertaining to disruption response are typical of most agencies. There are two primary areas of concern with respect to the agency's work rules: wage rules, such as overtime pay, and operator-use limitations, such as spread time limits. CTA employs part-time, regular full-time (five 8-hour shifts per week), and block full-time (four 10-hour shifts per week, with Monday or Friday and the weekend off) operators. Part-time operators may not work more than 30 hours per week, unless there is an emergency, and are entitled to overtime (150% of their regular wage) when working more than 8 hours in a day or 40 hours in a week. Once part-time operators report to work, they are guaranteed 2 hours of pay. They may operate full-time runs, as long as all full-time extra board operators have already been assigned work and they do not exceed their maximum hours.

Regular full-time operators are guaranteed 8 hours of pay daily and 40 hours weekly, while block full-time operators are guaranteed 40 hours per week. Both earn overtime pay at 150% of their regular wage rate for any time worked above these thresholds. For runs scheduled longer than 10 hours, regular full-time operators earn an additional 50% wage premium, in addition to the regular or overtime wage paid. Block operators do not earn a spread premium, but the total spread time for any run can be no more than 13 hours (CTA and ATU, 1996).

### 5.1.2. Policy/Exclusions

Policy and exclusions include hiring and promotion practices, standard operating procedures, excluded service restoration techniques, and reverse incentives. CTA practices in this area are also representative of the industry. Operators are hired as part-time operators and full-time operators are selected exclusively from the pool of part-time operators. Part-time operators may not be considered for full-time employment until they have one consecutive year of service, without accidents or attendance problems, however, after two years of service, each part-time operator must be hired as a full-time operator, regardless of his accident or attendance record. Operators must apply and compete for line instruction and service management positions and these positions are filled exclusively from the pool of full-time operators.

CTA allows service managers to use any of the service restoration techniques described in this study. Standard operating procedures generally do not allow operators to re-route service or take any other restorative action without instruction from a controller or supervisor. The exception is on designated operator empowerment routes, where operators have some authority to initiate service restoration (CTA, 2001c).

### 5.1.3. Training

CTA training focuses on providing intensive multi-day or multi-week introductory courses for new operators, controllers and supervisors that give an overview of the basics they will need to perform their new job. Short courses on specific elements or rehearsing different scenarios are not mainstays, with the exception of defensive driver training or, infrequently, special training programs that were not covered in the basic training, such as customer service or diversity trainings.

Specifically, CTA provides its operators the initial commercial drivers license test training, defensive driving training, and post-accident retraining. All accepted supervisor applicants take a 3-week supervisor training class before joining the supervisor pool. They are offered a modified version as a refresher course if they have been out on long-term absence. Likewise, controllers have an initial intensive training course and modified refresher course. There is no formal continuing education for supervisors and controllers where they explore different topics on an annual basis or forums in which to share their findings on what works well and what does not.

### 5.1.4. Communications

While being technically advanced and offering both data and voice capabilities, the CTA communications system does not support service management as well as it was designed to. This is primarily because of how information is made available to different parties and which groups are in direct communication. Supervisors and operators do not have a direct communications link; they only converse on the street when operators pass supervisors. This creates service management problems because supervisors can only instruct vehicles as they pass, at which point it may be too late to influence how their current trip is operated. Additionally, supervisors do not currently have any text messaging or automatic vehicle location (AVL) tracking or viewing capabilities. They can only receive information from controllers through the supervisor voice channel, which broadcasts to all supervisors. Lastly, direct supervisor-to-supervisor contact is via the supervisor voice channel, which is at capacity during the peak service hours (Barker, 2002).

Controllers also face communications limitations. There are two types of controllers involved in bus service management at CTA: bus channel and supervisor channel controllers. There are several garage controllers who work with all runs associated with their assigned garages. One supervisor channel controller maintains radio communication with all supervisors in the entire system. Operators and bus channel controllers can contact each other directly, but the text messages convey a limited amount of information, which generally requires a follow-up radio call. Radio contact takes a few minutes to establish and requires the operator to curb the bus while using the radio.

Bus channel controllers can forward text messages to the supervisor channel controller to be announced over the supervisor channel when delay, accident, or standing vehicle reports and other information needs to be relayed to supervisors. Therefore, the process for getting information about a problem that developed in the field to supervisors is cumbersome. Information travels from the operator to the bus channel controller via data messaging and voice channels, from the bus channel controller to the supervisor channel controller via data and direct voice contact, and from the supervisor channel controller to the supervisors via the voice channel (Barker, 2002).

### Service Management Staff and Responsibility Distribution

CTA manages service and responds to disruptions through a combination of control center-based controllers, commonly called dispatchers at other agencies, and field-based supervisors, called inspectors at some agencies. As described above, controllers are often the first point of contact and instruction for operators who have encountered a problem. Thus, controllers gather information and alert supervisors to problems. They also dispatch tow trucks and vehicle maintenance crews, and place bus orders with garages when a vehicle exchange or additional vehicle is needed (Barker, 2002).

Supervisors respond directly to field problems, such as accidents or other disturbances, as well as take the primary role in service restoration for both degraded and disrupted service. Post supervisors are assigned to a location, typically a relief point, terminal, or major intersection with other routes, and monitor the service passing his post. Mobile supervisors have a vehicle and monitor and respond to problems on all routes within a defined district. Mobile supervisors respond to in-route disruptions, such as accidents and vehicle failures, and manage the disruption scene. Thus, post supervisors along the disrupted route have primary responsibilities for disruption response, which are likely to be distributed, such that a terminal-based supervisor deals with all terminal-based service restoration strategies, while supervisors along the route deal with operator reliefs and in-route service management (Barker, 2002).

### 5.1.5. Available Resources

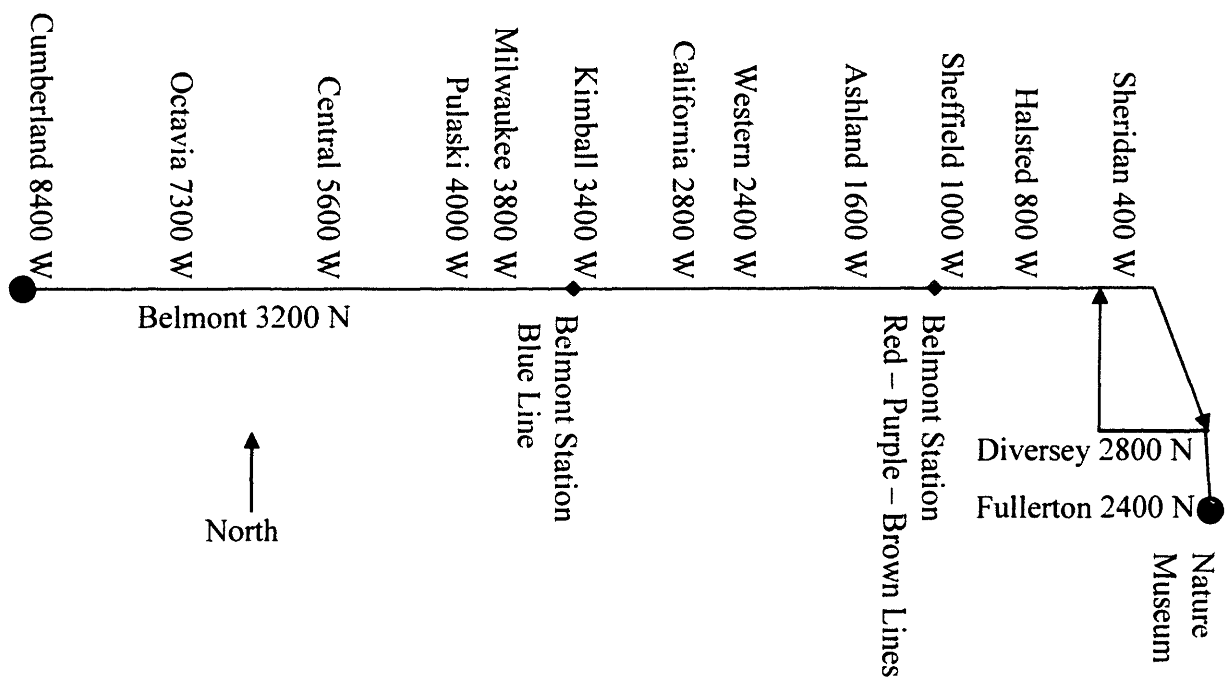
There is no specific CTA policy about the availability of operators and extra vehicles. Historically, a few gap buses and operators have been scheduled during the morning and afternoon peaks and stationed with a post supervisor to use as he sees fit. Extraboard and part-time operators are assigned open work a day in advance and a list of operators interested in working overtime or making up missed work is available. The available operators are pre-assigned work and there are few, if any, operators available and unassigned during pullouts to cover no-show absences as they occur.

## 5.2. Route Characteristics

### 5.2.1. Route Overview

Route 77 Belmont, a high-frequency east-west crosstown route on the north side of Chicago will be used as the case study route. Several junior and senior high schools are located along the central and western portions of the route, and dense mixed use commercial and residential development characterizes the eastern portion. It intersects the Red and Brown elevated rail lines at Sheffield, one block west of Halsted on the eastern part of the route, and intersects the Blue subway line at Kimball, in the middle of the route. A schematic of the route is shown below in Figure 5-1.

**Figure 5-1: 77 Belmont Route**




First, the datasets used to develop the route profile are discussed. Then, a basic route profile is developed. These characteristics are used for the disruption response processes.

### 5.2.2. Datasets

Two datasets were used to estimate passenger demand information: the automatic fare collection (AFC) data and manually recorded point checks. The AFC data was used to calculate boardings per trip and, thus, hourly demand by direction. The point check data provides average loads at the maximum load point and headway variance. While service managers do not have this information in real-time, it provides historical information on how the route usually operates. Lastly, the operations plan provides information about where the vehicle and operators should be, which service managers can use to discern where conflicts may develop in the future and potential opportunities for pull-in vehicles and operators, as well as scheduled headways, relief times and locations and scheduled travel times.

### 5.2.3. Route Profile and Operating Plan

#### Infrastructure Characteristics and Operations Environment

Several miles of the eastern portion of Belmont Avenue has a single lane in each direction with parking along both sides, which contributes to traffic congestion and limits the ability of vehicles to pass each other. There is an additional lane in each direction on the western portion of the route and less traffic congestion, making passing easier. The Blue Line has a terminal on the southeast corner of Belmont and Kimball, which requires all westbound buses to make an uncontrolled left turn from Belmont into the terminal and a right turn into the left turning lane on Kimball. These two moves cost a lot of running time. Stops are generally every 2-3 blocks and there is no signal priority along the route.

#### Demand Profile

There is limited passenger demand information available at CTA. Ideally, service managers would like to know three basic pieces of passenger demand information: the average and maximum bus load under typical operating conditions and how close these are to the vehicle capacity in each direction by hour. In the absence of this information, supervisors use operations-based indicators, such as vehicle headways to ascertain the passenger impacts, knowing the greater the headway, the greater the load and the longer the average wait time.

The demand profile in this study relies on manually recorded point check data, which provides passenger loads for select locations along the route. The point check locations are points near the ends of the route and at major connection points along the route. Because there are no ride checks, these locations may not include the true peak load location, but they are the CTA service planning department's best guess. The point checks also mask minor increases and decreases in the actual load along the route. The

AFC data can be used to calculate the total number of passengers served by each trip, but it does not record where the boarding occurred or where passengers disembark, and, thus, does not show vehicle loads.

Table 5-1 shows the average and maximum loads at the maximum load point on the 77 Belmont. The primary directions of travel are eastbound in the morning and westbound in the afternoon. The highest demand direction overall is westbound in the afternoon, with an average load of 50. Kimball is at the Blue Line station and Pulaski is east of the station.

**Table 5-1: Route 77 Belmont Highest Maximum and Average Load by Time Period and Direction**

| Time of Day | Direction | Average Boardings per Trip[10] | Location | Average Peak Load | Location | Highest Single Peak Load |
|---|---|---|---|---|---|---|
| Morning Peak | Eastbound | 85 | Kimball | 38 | Pulaski | 68 |
| | Westbound | 70 | Pulaski | 24 | Pulaski | 51 |
| Afternoon Peak | Eastbound | 71 | Pulaski | 29 | Pulaski | 54 |
| | Westbound | 99 | Kimball | 50 | Kimball | 72 |

Supervisors also need to know approximately how the total trip demand is distributed along the route, which is shown in Table 5-2 below. This was estimated from the manual load checks and limited knowledge of the route.

[10] The morning average is 6-8am and the afternoon average is 4-5pm.

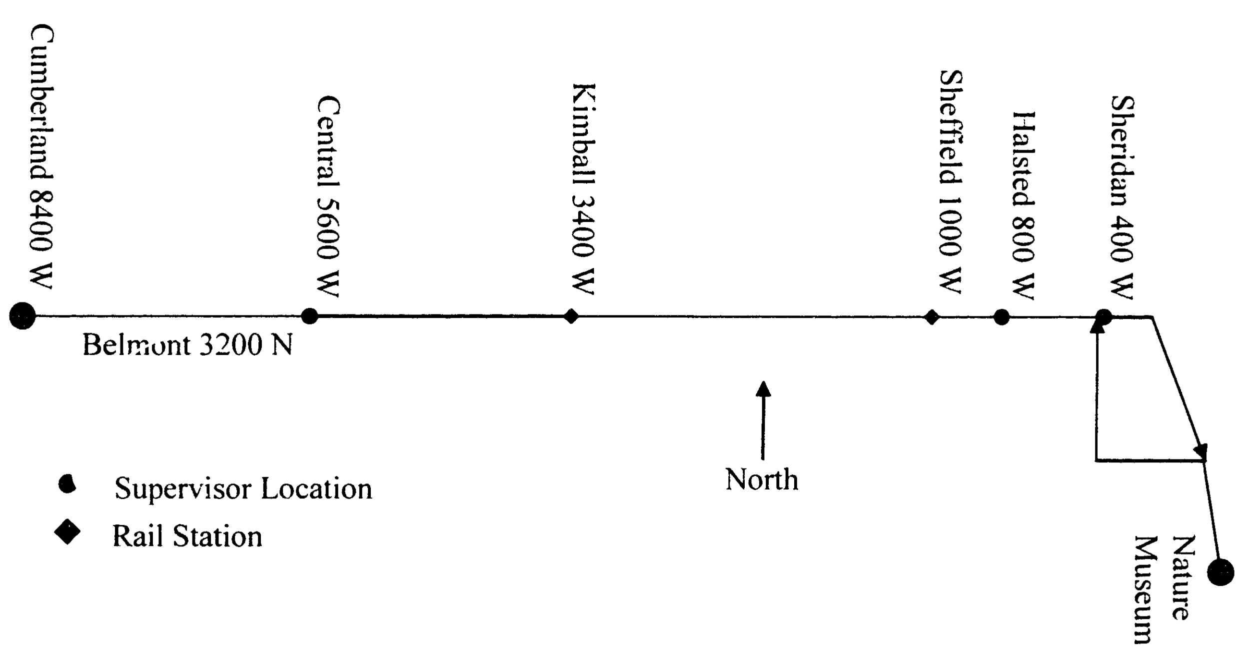
**Table 5-2: Route 77 Belmont Afternoon Estimated Boarding Distribution**

| Westbound | | Eastbound | |
|---|---|---|---|
| **Route Segment** | **Percentage** | **Route Segment** | **Percentage** |
| Nature Museum – Red Line | 5% | Cumberland – Central | 20% |
| Red Line – Western | 45% | Central – Pulaski | 25% |
| Western – Pulaski | 25% | Pulaski – Western | 30% |
| Pulaski – Central | 10% | Western – Red Line | 20% |
| Central – Octavia | 10% | Red Line – Nature Museum | 5% |
| Octavia – Cumberland | 5% | | |

## Service Management

There are 3 post supervisors stationed along the route. From east to west, they are at Sheridan, Halsted, and Central (shown in Figure 5-2 below), but no supervisors are stationed at either terminal. There are, however, four mobile supervisors with districts that include a portion of the 77 Belmont route. Supervisors have printed copies of each route's operating plan, which includes each route's schedule, run numbers, when each run should be at timepoints along the route, and next trip information. They are expected to know the primary direction of travel and major boarding and alighting points.

**Figure 5-2: Supervisor Locations**



### Operations Plan

Buses are scheduled every 5-6 minutes during the morning and afternoon peaks and 8 minutes during the midday. The relief point is at Central with operators being relieved both eastbound and westbound. While reliefs occur throughout the day, most reliefs occur 8-10am, 11am-1pm, and 3-5pm. The scheduled running time is about 80 minutes eastbound and 70 minutes westbound in the morning, 70 minutes in the midday, and 80 minutes during the afternoon peak (see Table 5-3). The same amount of recovery time is scheduled at each end of the route, with 8-9 minutes at each end during the morning peak, about 15 minutes during the midday, and 10-15 minutes during the afternoon peak. There is one interlined trip in the afternoon.

**Table 5-3: Route 77 Belmont Scheduled Afternoon Travel Times**

| Westbound | | Eastbound | |
|---|---|---|---|
| **Route Segment** | **Travel Time (min)** | **Route Segment** | **Travel Time (min)** |
| Nature Museum – Western | 27 | Cumberland – Central | 15 |
| Western – Pulaski | 17 | Central – Pulaski | 13 |
| Pulaski – Central | 10 | Pulaski – Western | 20 |
| Central – Octavia | 12 | Western – Halsted | 15 |
| Octavia – Cumberland | 6 | Halsted – Nature Museum | 10 |

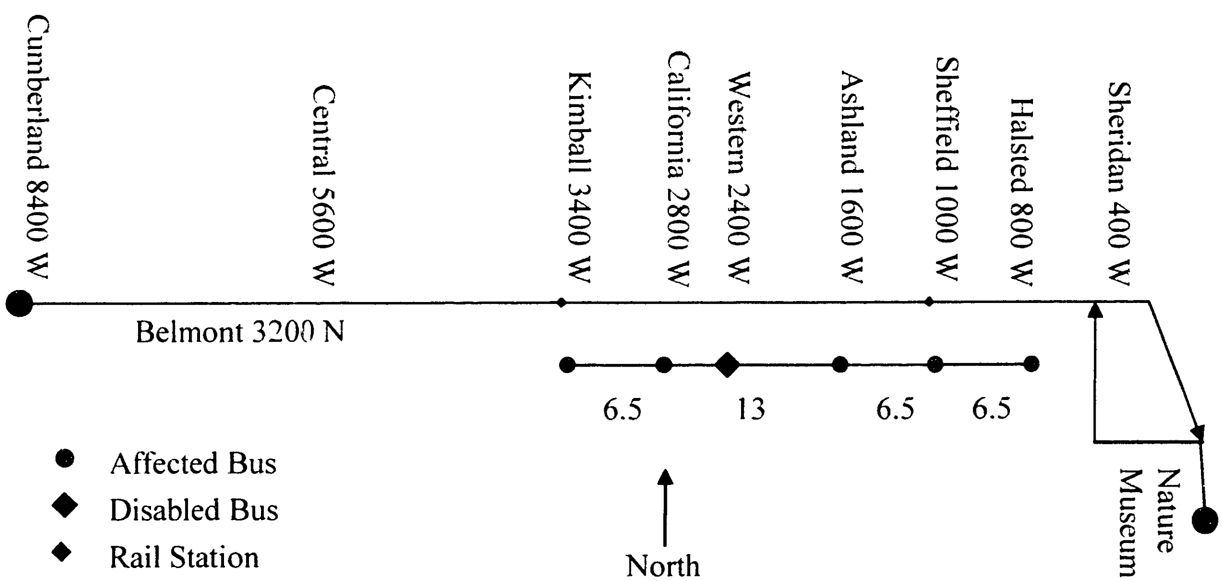
## 5.3. Disruption Overview

The process is applied to a standing bus disruption with a bus having a mechanical failure at Western during a westbound afternoon peak trip. At the moment of the disruption all vehicles are on time and at the locations listed in Table 5-4 and shown in Figure 5-3 below. The full-time afternoon pull-out operator is with the vehicle trying to restart it. According to current practice, the operator will stop the following bus and transfer all passengers to that bus.

**Table 5-4: Vehicle Location and Load Estimates when Disruption Occurs**

| Vehicle | Schedule Adherence | Headway | Load at Time of Disruption[11] | Location |
|---|---|---|---|---|
| Second Leader | On schedule | 6.5 | 50 | Kimball (Blue Line) |
| Leader | On schedule | 6.5 | 42 | California |
| Disabled Bus | Not Operating | N/A | 38 | Western |
| Immediate Follower | On schedule | 13 | 35 | Ashland |
| Second Follower | On schedule | 6.5 | 32 | Sheffield (Red Line) |
| Follower 3 | On schedule | 6.5 | 16 | Broadway (before Halsted) |

**Figure 5-3: Vehicle Location when Disruption Occurs**




Two cases are discussed in the following section: the first is under ideal conditions where the service manager has all pertinent information and can implement a disruption response plan immediately, and the second is under the constraints that currently face CTA bus operations.

## 5.4. Ideal Disruption Response

Given the route and disruption information described above, this section develops the best possible response if there were perfect information on vehicle location, passenger load and passenger demand, and the agency could respond at the moment the disruption occurred. While the best specific response to each disruption can be found using methods developed in prior studies and having accurate passenger demand, passenger load, and vehicle location data, this study uses available CTA data and makes some general assumptions to identify which restoration techniques would be the most appropriate. The ideal case also highlights constraining policies that, if changed, could result in improved disruption response.

### a) Immediate Disruption Response

First, a "No Intervention" case is considered, so the agency can know what to expect if they do nothing. The following cases consider options available to manage the service for the remainder of this trip. Only combinations of extra service, holding, and expressing are considered. Short-turning and borrowing are not considered as a response on the initial trip because the disruption occurred in the middle of the route, making them difficult to justify due to the impacts to passengers on the service from which it would be borrowed and the time needed to get the bus in service.

#### 1. No Intervention

If the trip is allowed to finish without intervention, about 132 passengers are expected to wait longer than 1.5 headways (see Table 5-5). This is developed from the average westbound trip demand for trips starting between 4 and 5pm. An extra bus to pick up the passengers on the disabled bus would be necessary to eliminate the additional headway of wait time created for passengers on the following bus.

From the manual load checks, on average between 4 and 5pm, there are 38 passengers on each westbound bus at Western (see Table 5-1). Passengers on the disabled vehicle have

[11] Estimated load from manual point check data.

the initial wait for the bus, the wait for the following bus and the delay to transfer to the following bus. Therefore, the 38 passengers from the disabled bus, on average, experience up to a one headway wait when waiting initially, plus 2 headways while waiting for the following bus to restart service for a total of 2-3 headways of wait time and stopped delay. Passengers on the following bus have their initial wait for the bus (up to one headway) and the delay to transfer passengers to their bus, so they experience 1-2 headways of wait time and stopped delay. Thus, 57 passengers wait at least 1.5 headways. One passenger would need to wait for the next bus because he would not fit on the following bus if the transfer were instantaneous. If it takes 6 minutes to move passengers from the disabled to the following bus, the second following bus will be pulling up as the first follower completes boarding passengers from the disabled bus. Therefore, the last passenger can board the second follower. This initial impact will remain the same, regardless of what restoration techniques are implemented after this passenger transfer occurs.

From Table 5-2 we assume 50% of passengers have yet to be served on this trip and the gap is 3 headways long once the passengers have moved to the follower, so 149 passengers (99 * .5 * 3) is the total number served by the 2 bunched buses following the gap. This is shown in Table 5-5 below.

**Table 5-5: No Intervention Passengers Impacts**

| Route Segment | Largest Service Gap | Boarding Demand[12] | Missed Passengers | Total Waiting >= 1.5 Headways |
|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 0[13] (76 – 150) | 57 (38 * 1) (38 * .5) |
| Western – Cumberland | 20 minutes (3 headways) | 149 (99 * .5 * 3) | 0[14] (50 * 3) – (75 * 2) | 74 (149 * .5) |
| **Total:** | **20 minutes (3 headways)** | | **0** | **132** |

### 2. Holding

This response uses a simple holding technique. Because the gap is 3 headways long after the disruption occurs and passengers are transferred, two gap leaders are held to spread the gap. The disabled bus' second leader is instructed to wait 3 minutes at the Kimball Blue Line station. Its first leader is told to lose 4 minutes on his schedule between his current position at California and to wait at the Blue Line station at Kimball another 3 minutes. This will create three headways at 10-13 minutes, rather than two at 6.5 and one at 20 minutes. The tradeoff is the delay created for passengers approaching and leaving the Blue Line station. This reduces the number of passengers waiting more than 1.5 headways from 132 to 92.

[12] Boarding Demand = Average Total Boardings per Trip * Proportion Boarding in Route Segment (from Table 5-2) * Headway

[13] The second follower arrives as the first follower is finishing loading up passengers, so the capacity of two buses (150) is available for picking up passengers from the disabled bus.

[14] Missed Passengers at Peak Load Point = Anticipated Load – Provided Capacity

**Table 5-6: Holding Leaders Disruption Intervention Passengers Impacts**

| Route Segment | Largest Service Gap | Boarding Demand | Missed Passengers | Total Waiting >= 1.5 Headways |
|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 0 (76 – 150) | 57 (38 * 1) (38 * .5) |
| Western – California | 20 minutes (3 headways) | 15 (99 * .05 * 3) | 0 (76 + 15 – 150) | 8 (15 * .5) |
| California – Kimball[15] | 18 minutes (2.8 headways) | 42 (99 * .15 * 2.8) | 0[16] (91 + 42 – 150) | 18 (42 * .44) |
| Kimball – Cumberland | 13 minutes (1.3 headways) | 39 (99 * .3 * 1.3) | 0[17] | 9 (39 * .23) |
| **Total:** | **20 minutes (3 headways)** | | **0** | **92** |

### 3. Expressing + Holding

This response uses expressing and holding to reduce the gap size, which is 3 headways long when the following buses leave Western. Ideally, the first following bus, which is at capacity after picking up passengers from the disabled bus, would operate express making stops only to drop passengers off and serve major stops until it makes up 8 minutes[18] or reaches Pulaski, after which it should serve all stops, eventually creating an 8 minute headway. However, there is only 20 minutes in the schedule between Western and Pulaski, so the expressed bus cannot make up 8 minutes by the time it reaches the peak load point.

Assuming buses travelling express can make up 10% of the scheduled running time, it can only make up 2 minutes by Pulaski. With this information, the second gap leader is instructed to wait 2.5 minutes at the Blue Line station, creating a 9 minute headway for him. The first gap leader should lose 3 minutes between California and the Blue Line station and wait 3 minutes at the Blue Line station, creating a 10 minute headway for him

---

[15] Passengers on following vehicles.

[16] Estimated Load – Capacity (Alighting rates are unknown. As shown, nobody alights, so this estimated demand is greater than what the real demand would be.)

[17] Enough capacity is provided after the peak load point.

[18] Putting it 2-3 minutes up on its printed schedule.

and a 12-14 minute headway for his follower, the expressing bus that picked up all of the passengers on the disabled bus. This would create 3 headways under the threshold of 10 minutes (1.5 scheduled headways) and one at 12-14 minutes (2 scheduled headways), rather than two at 6.5 (one scheduled headway) and one at 20 minutes (3 scheduled headways), followed by two buses. This creates a bigger window for passengers connecting from the Blue Line to catch a bus and substantially reduces the service gap, but the tradeoff is the delay created for the passengers held at the Blue Line station travelling through the station and the 9 missed passengers as the bus expresses (which are served within 2 minutes by the following bus). It reduces the number of passengers waiting more than 1.5 headways from 132 with no intervention to 85. Clearly, when compared to the holding only option shown above, nearly all of the benefit in this case (95%) comes from holding the preceding vehicles, rather than from the expressing.

**Table 5-7: Holding Leaders and Expressing Follower Passengers Impacts**

| **Route Segment** | **Largest Service Gap** | **Boarding Demand** | **Missed Passengers** | **Total Waiting >= 1.5 Headways** |
|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 0 (76 – 150) | 57 (38 * 1) (38 * .5) |
| Western – California | 20 minutes (3 headways) | 15 (99 * .05 * 2.95[19]) | 0[20] | 7 (15 * .47) |
| California – Kimball | 18 minutes (2.8 headways) | 37 (99 * .15 * 2.5) | 9[21] (37 * .25)[22] | 12 (28 * .25) + (9 * .5) |
| Kimball – Cumberland | 12 minutes (1.8 headways) | 53 (99 * .3 * 1.8) | 0[23] | 9 (53 * .17) |
| **Total:** | **20 minutes (3 headways)** | | **9** | **85** |

[19] Average headway for this segment
[20] The buses are still bunched for this segment, so passengers are not considered missed even if the following bus serves them, instead of the expressing bus.
[21] These passengers wait another 3.5 minutes on average.
[22] Assume 25% of passengers were missed by the expressing bus.

### 4. Extra Vehicle

Another group of passengers to target are those waiting to board between Western and Pulaski. An extra vehicle could be used to split the 20 minute gap into two 10 minute headways, eliminating the need for any passengers to wait for more than 1.5 headways. An available vehicle that could be inserted in the middle of the gap would eliminate missed passengers and reduce the number of passengers waiting more than 1.5 headways to just the passengers on the disabled and following bus.

**Table 5-8: Extra Vehicle Inserted Mid-Gap Passengers Impacts**

| Route Segment | Largest Service Gap | Boarding Demand | Missed Passengers | Total Waiting >= 1.5 Headways |
|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 0 (76 – 150) | 57 (38 * 1) (38 * .5) |
| Western – Cumberland | 10 minutes (2 headways)[24] | 99 (99 * .5 * 2) | 0 | 0 |
| **Total:** | **20 minutes (3 headways)** | | **0** | **57** |

### 5. Extra Vehicle + Holding

This response adds an extra vehicle at the disruption location to serve passengers one headway after the disruption occurs. While this does not reduce the wait for passengers on the disabled bus, it provides more capacity, eliminates the impacts to passengers on the following bus. By allowing the following bus to continue in service rather than stop to pickup a load of passengers, the headway gap only gets as large as two headways for passengers, except those on the disabled bus. The gap leading bus is instructed to lose 3 minutes on his schedule between California and Kimball. This reduces the number of passengers waiting more than 1.5 headways from 57 to 42. Overtime for the extra vehicle to complete this trip and return to the supervisor or pull-in could be estimated at 1.25

[23] Passengers are not skipped after the peak load point.

[24] The service gap for the rest of the route is two headways because the following bus is not delayed to pick up passengers on the disabled bus, so it continues on its schedule.

hours to complete the trip to Cumberland and return to the pull-in location with the supervisor at Central. Table 5-9 shows the passenger impacts of this response.

**Table 5-9: Extra Vehicle at Disruption Location + Leader Holding Passenger Impacts**

| Route Segment | Largest Service Gap | Boarding Demand | Missed Passengers | Total Waiting >= 1.5 Headways |
|---|---|---|---|---|
| Disabled Bus + Extra Vehicle | 20 minutes (3 headways) | 38 (Load) (38 * 1) | 0 (38 – 75) | 38 (38 * 1) |
| Western – California | 13 minutes (2 headways) | 10 (99 * .05 * 2) | 0 (48 – 75) | 2 (10 * .23) |
| California – Kimball | 13 minutes (2 headways) | 25 (99 * .15 * 1.7[25]) | 0 (73 – 75) | 2 (25 * .09) |
| Kimball – Cumberland | 10 minutes (1.5 headways) | 45 (99 * .3 * 1.5) | 0 | 0 |
| **Total:** | **20 minutes (3 headways)** | | **0** | **42** |

## Ideal Response Summary

This exercise of considering the best possible solution has developed a baseline against which responses developed within the agency constraints can be evaluated. It highlights the potential improvement available through changing agency policy, which policies are constraining, and which disruptions have only poor disruption responses available indicating they should be targeted for avoidance.

Of all responses considered, passengers stand to gain the most if extra service is provided immediately in the middle of the service gap, coupled with effective holding of the gap leaders. If extra service is unavailable, expressing and holding vehicles into the gap is the most effective way to reduce the number of passengers with excessive wait times. The impacts of the disruption response cases considered are shown in Table 5-10 below.

[25] Average headway for this segment

**Table 5-10: Ideal Responses Summary of Initial Trip Passenger Impacts and Agency Cost**

| Disruption Response | Missed Passengers | Passengers Waiting >= 1.5 Headways | Maximum Wait Time[26] | Agency Cost[27] |
|---|---|---|---|---|
| No Intervention | 0 | 132 | 3 headways (20 minutes) | 0 |
| Holding | 0 | 92 (30% less)[28] | 3 headways (20 minutes) | 0 |
| Holding + Express | 9 | 85 (36% less) | 3 headways (20 minutes) | 0 |
| Extra Service | 0 | 57 (57% less) | 3 headways (20 minutes) | 1.25 hours overtime |
| Extra Service + Holding | 0 | 42 (68% less) | 3 headways (20 minutes) | 1.25 hours overtime |

## b) Long-term Response

If the disruption is anticipated to also affect future trips, missing service must be replaced to maintain the existing operating plan or a new operations plan developed, given the prior headway is unachievable with a missing vehicle. Because this disruption is a service gap and adding a vehicle would allow the scheduled operating plan to work, first it should be determined whether there is a no-cost service option, a gap bus and operator or extra vehicle and extra board operator from the garage, available. If there is an available gap bus and the passenger impacts created by this disruption are anticipated to be the greatest during the estimated disruption duration, he should use it.

If no-cost service is unavailable, there are several options for improving delays on high frequency service: having operators make extra trips (typically on overtime), re-spacing service from the terminal, or using short-turning or deadheading to concentrate service where it is most needed by moving service from the lower demand portions of the route to the main trunk section. Capacity problems are not anticipated when the service is

---

[26] Although this depends on the route segment, the longest wait or service gap is 3 headways, which is what passengers on the disabled bus experience.
[27] Available gap buses would not cost the agency overtime because the extra service is scheduled.
[28] When compared to the "no intervention" case.

evenly spaced with one vehicle missing, so the primary service goal becomes reducing passenger delay.

Most of the 77 Belmont demand occurs between the Red Line and Pulaski, making the Red Line the most promising short-turning location. This could be done in two ways. The first is terminating an eastbound trip at the Red Line and starting its westbound trip from there. The second is deadheading a bus from the eastern terminal and having it start westbound from the Red Line. Short-turning would get the vehicle in place faster and the head sign could indicate that it would terminate at the Red Line, so passengers would know before boarding that the service would terminate early. Operators would also need specific turnaround instructions to keep the bus on safe streets.

To re-space service from the terminal, two adjustments are needed. The disabled bus leader should leave two minutes after his scheduled departure time, making his headway 8-9 minutes and the gap follower should leave 2 minutes early, 8-9 minutes behind the gap leader. Leaving two minutes early also creates an 8-9 minute headway for its follower. This headway does not create capacity problems for the reverse direction or following trips, as shown in Table 5-11 below. If buses can keep their intervals spaced and passengers arrived evenly, passengers would not be missed or wait longer than 1.5 headways. However, the westbound trips will be very full and could encounter capacity problems due to passenger demand variability.

**Table 5-11: Long-Term Disruption Response Options**

| Service Restoration Options | Max Peak Load | Missed Passengers | Passengers Waiting >= 1.5 Headways[29] | Additional Agency Cost |
|---|---|---|---|---|
| Re-space | 65[30] (50 * 1.3) | 0 | 0 | 0 |
| Extra Service | 50 | 0 | 0 | 1.56 overtime hours |
| Short-turning | 50 | 0 | 2[31] (71 * .05 * .23) + (99 * .05 * .23) | 0 |
| Dead-head | 50 | 0 | 1[32] (99 * .05 * .23) | 0 |

## 5.5. Disruption Response Process

This section presents a case study applying the disruption response process as it would be followed in the CTA policy context, to develop a response for a specific disruption. It follows the steps developed in Chapter 4: information gathering, identifying feasible service restoration options and developing a disruption response plan within their agency policy constraints. The first part describes the information service managers would like to have to determine the best response and then follows the steps in the disruption response process. Then, the process of identifying the appropriate service goals and evaluating the feasible disruption responses is shown.

There are several critical differences between the ideal and CTA-context cases, which limit the responses available to CTA: available information, communications and response time. The differences and their impacts are highlighted throughout the case.

[29] Only occurs if reverse direction is not spaced.
[30] Boarding spikes are likely at the rail stations and could put some vehicles over capacity.
[31] Short turning eliminates the short-turned operator's recovery time and the operator will need to be put back in place at some point.
[32] Dead head time comes out of recovery time, so it may not be reliable.

### 5.5.1. Route Characteristics

In order to select among service restoration options, the agency must understand the route characteristics, such as the operations plan and route profile. These frame the available options. The route characteristics for this case-study are described above in Section 5.2.

Because of the availability and stability of this information and lack of time when disruptions occur, this should be determined during the pre-planning support phase. Ideally, service managers should know this information in advance for each of the routes he manages and should have it available for other routes.

### 5.5.2. Disruption Information Gathering

Ideally, the supervisor knows the current weather, transit system status, passenger load patterns, and traffic conditions, as well as the current availability of extra service, such as no-cost gap buses, extra vehicles or operators at the garage, pull-ins, runs without a full 8 hour shift, and the availability of part-time operators before the disruption occurs.

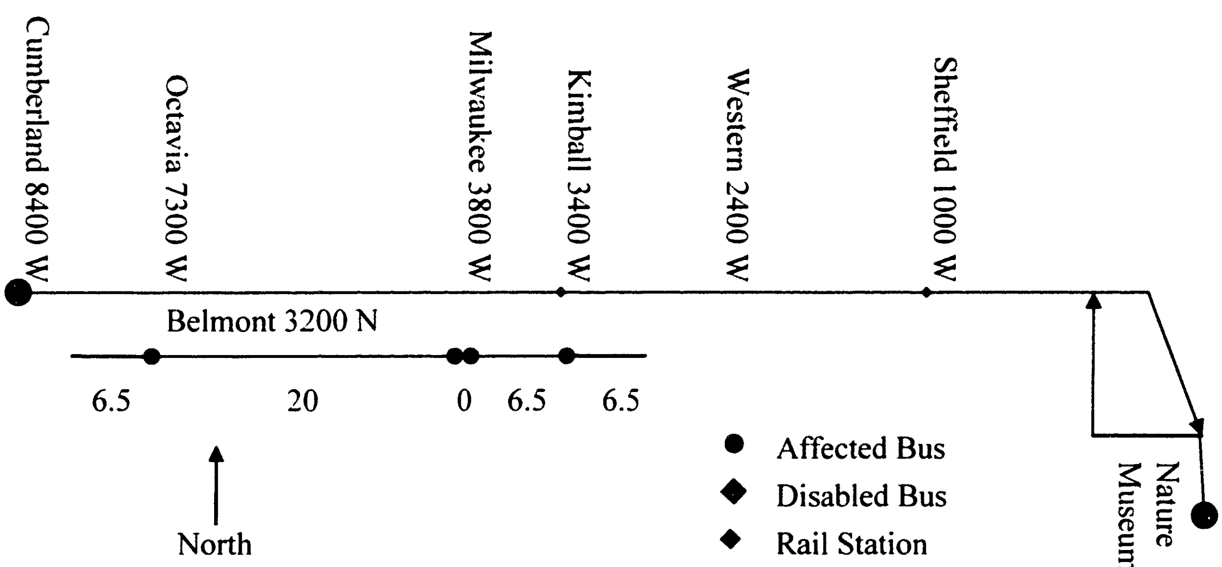
As seen in the ideal responses developed in Section 5.4, the effectiveness of the response selected depends heavily on the vehicle locations, passenger loads, and passenger demand at the time of the disruption. Without this information, it is difficult for the agency to pre-plan an intelligent response, much less develop one in real-time. CTA supervisors record headways and passenger loads as buses pass them and buses that do not pass them are suspected problems, but they do not have comprehensive real-time passenger or vehicle location information. Table 5-12 identifies the information available to the supervisor at Central on bus location and passenger load estimates 20 minutes after the disruption started, which is an average of how long it takes for supervisors to be notified of the problem via the controller managing the supervisor radio channel. Bus location and headway estimates are also shown in Figure 5-4.

**Table 5-12: Vehicle Location and Peak Load Estimates when Supervisors are Notified[33]**

| Vehicle | Schedule Adherence | Headway | Estimated Load | Location |
|---|---|---|---|---|
| Leader[34] | On schedule | 6.5 | 13 | Octavia |
| Disabled Bus | Not operating | N/A | 0 | Western |
| *Immediate Follower[35]* | *One headway late* | *20* | *68* | *Milwaukee* |
| *Follower 2* | *On schedule* | *0* | *45* | *Milwaukee* |
| *Follower 3* | *On schedule* | *6.5* | *50* | *Kimball (Blue Line Station)* |

*Assumed location, but supervisor is not certain.*

**Figure 5-4: Vehicle Locations when Supervisors are Notified**




Supervisors would also like to know what resources are available when developing a response. Table 5-13 below also identifies the resources available to respond to this disruption. as well as their costs and the additional passenger impacts of using them.

---

[33] Twenty minutes after the disruption occurred.

[34] Estimated location from his service check information.

**Table 5-13: Available Resources**

| Source | Available | Impact to other Service | Agency Cost |
|---|---|---|---|
| Gap vehicle and operator | Yes | Yes, if needed for larger disruption | No-cost |
| Extra Trip(s) | Yes | No | Over-time |
| Additional vehicle and operator | No | No | Over-time |
| Borrow from other Route | Yes | Yes | No-cost |

The only piece missing is the specific disruption information. From the moment he is notified about the disruption, the supervisor begins collecting information about it. During the initial controller announcement the supervisor learns the information shown in Table 5-14 about the disruption.

**Table 5-14: Disruption Information from Controller Announcement**

| Category | Disposition | Specifics | Fact/Inferred |
|---|---|---|---|
| Disruption Type | Standing Vehicle | Mechanical Failure – Dead Motor | Fact |
| Impact Location | Western | | Fact |
| Anticipated Timeframe | Longer than one trip | Possibly entire peak period | Inferred from personal experience |
| Disruption Magnitude | Limited to service gap. | | Inferred from personal experience |
| Operator Experience Levels | Mostly experienced | Only 2 part-time runs on the street | Inferred from knowledge of current pick |

## Disruption / Supervisor Assumptions

Supervisors do not have real time vehicle location or passenger count information to track the buses and thus they can only make reasonable guesses about buses that have passed them. This case has been developed with average loads and assuming that prior to the disruption buses were on schedule and each bus was carrying an average passenger load. In order to illustrate a typical disruption clearly, the disruption and the supervisor's

[35] Unknown, but estimated from past experience.

assumptions are the same. While this will not always be true, this case study is designed to show the disruption response process, not the difference in response given the supervisor's perceptions and the actual situation.

### 5.5.3. Identify Relevant Service Goals

Once the supervisor has information about the disruption, he can identify the proper service goals. While it is important to identify if passengers will be skipped because of insufficient capacity and respond to that problem, the agency must balance this with responses that would reduce the number of passengers with excessive wait times, particularly if the additional wait for skipped passengers is less than half a headway.

There are two critical capacity points: at the disruption location and at the peak load point. The supervisor should determine whether there will be enough capacity provided by the first set of buses at each of these locations. To do this, he estimates the size of the service gap, the passenger arrival rate, and the average capacity of the vehicles assigned to the route, from which he can estimate expected loads on the buses following the gap.

The capacity of buses assigned to the 77 Belmont is 75. At the time and location of the disruption, there was enough capacity to pick up almost all passengers from the disabled bus with the following bus. When using the average load for the disruption location, one passenger must transfer to the second following bus, which pulls up as the first follower is finishing loading passengers. The service gap is now 3 headways, which means there will be 3 times as many passengers wanting to board the first set of vehicles as usual and there is only 50% of one bus' capacity available.

The two followers will almost always be bunched because of the time needed to move passengers from the disabled bus to its follower at the disruption site and the additional travel time needed once moving because of increased passenger boardings. This is true because of the difficulty of getting passengers on and off the bus, as well as the need to make more stops overall at minor bus stops because of the increased demand. With the two vehicles arriving together providing a combined capacity of 150 passengers, in this case there would be enough capacity to serve all passengers with the first set of arriving vehicles.

The peak load point should also be evaluated because it is the most constrained point during normal and disrupted operations. In this case, given the average passenger load at the peak load point is 50 and the unmanaged service gap is 3 headways, there will be 150 passengers who would like to be on a vehicle at that point and two vehicles passing through at the same time, if unmanaged. From this quick analysis, the supervisor can conclude there will be just enough capacity for this disruption. Passengers will be stuffed onto the buses, making providing more capacity ideal, but the primary goal becomes minimizing passenger delay.

### 5.5.4. Develop and Evaluate Service Restoration Options

This section shows how an agency would estimate the benefits of different responses and describes how supervisors should evaluate and select among service restoration options to develop a disruption response, subject to the limitations of available information. First, the key route and disruption characteristics are briefly reviewed, the feasible options to achieve the service goal are outlined and the evaluation criteria are calculated for each option.

The disruption type, anticipated disruption timeframe, current conditions, and the service goals determine the service restoration options to consider, which are summarized in Table 5-15 below.

**Table 5-15: Primary Attributes Determining Feasible Service Restoration Options**

| Attribute | Disposition |
|---|---|
| Disruption Type | Service Gap: Standing vehicle |
| Anticipated Disruption Timeframe | Several Trips – End of Time Period |
| Operation Goal | Minimize Passenger Delay |
| Operations Conditions | Normal<br>Disruption Impacts Limited to the Route |
| Feasible Service Restoration Techniques | Adding Service<br>Respacing from the Terminal<br>Holding<br>Expressing<br>Short-turning |

All of the service restoration techniques described in Chapter 3 are feasible for addressing service gaps. They are adding service, respacing from the terminal to provide an even headway, and all three in-route service restoration technique, holding, expressing and short turning. Because the supervisor anticipates the disruption will exist longer than one trip, he needs to evaluate the effectiveness of developing a disruption response for the immediate service gap problem as well as for future trips.

In this case, the primary impediment to improving service during this trip is that the vehicles following the gap are approaching the peak load point just as the supervisors learn about the disruption, making it difficult for any service restoration actions to improve most passengers' experiences. Other impediments to diminishing the service gap on this trip include the inability of the supervisor to contact and instruct operators in time for them to take the service restoration action and an inability to hold the gap leader because he is too far along his trip for holding to reduce passenger delay. Specifically, if the vehicle leading the gap is close enough to the end of the route, where his passenger load is declining, holding the leader to diminish the gap size would delay passengers on-board more than it would reduce delay time for passengers arriving during the gap.

### a) Immediate Disruption Response

First, the supervisor should consider whether he can manage the gap in the first trip. To do this, he must evaluate the costs and benefits of each service restoration option after he has eliminated techniques that will not help passengers given the specific disruption situation. Vehicle location and passenger load information, as well as current conditions are the primary indicators of which service restoration options are available.

In this case, with his estimate of vehicle locations and knowledge of the passenger boarding distribution, he can eliminate holding the leader and short-turning responses. Even with his limited information, he knows the leader is already somewhere around Octavia, beyond which there are few boarding passengers and holding would result in worse rather than better service. He can also eliminate short turning for the initial trip because the following buses are also likely to have passed the peak load point, which makes the extra capacity unnecessary and the impacts to eastbound service very disruptive. Borrowing service would have similar benefits and impacts and would do more harm than good, considering the negative passenger impacts on the parallel neighboring routes, and the time needed to move the bus into place. Expressing is not useful at the supervisor location and can also be eliminated. Extra service would be an effective headway shortening option, but the passenger benefits would be limited if started at Central at the supervisor location. Feasible options are summarized in Table 5-16 below.

**Table 5-16: Feasibility of Disruption Responses**

| Initial Disruption Response | Feasible |
|---|---|
| Holding | No |
| Short-turning | No |
| Borrow Service | No |
| Expressing | No[36] |
| Extra Service | Yes |
| Extra Service + Expressing | No[37] |

While supervisors could never calculate all this information in real-time, it would be ideal for supervisors to have this information accurately in real-time. Given that this is not the

[36] Expressing is only effective if implemented from the disruption site in this case, which is not currently CTA policy. Expressing after the peak load point is worse for passengers because of the impacts to missed passengers and the small number of passengers who will benefit.

case, supervisors, instead, develop general or historically-based passenger impacts in order to select among the options. Additionally, when evaluating options, the supervisor can concentrate on what the specific improvements of the option, knowing that the initial disruption impacts have already taken place, making this a smaller problem to estimate. One final note is that service restoration option evaluation is only as good as the information upon which it is based. If the information is incorrect, it may lead supervisors to make decisions that do not reduce negative passenger impacts.

### 1. No Intervention

The "No Intervention" option was considered in the ideal case. While the results are the same, the key difference in the CTA context is that instead of having specific numbers, the supervisor can estimate that no or few passengers will be missed and all passengers on the disabled bus, half of the passengers on the following bus, and half the remaining passengers for 3 times the typical trip (132 passengers) will have been delayed by 1.5 or more headways. See Table 5-5 for the specific passenger impacts.

### 2. Extra Service

The first service restoration option considered is using extra service to fill the gap. Generally, finding service in time to complete a trip is unlikely given the limited amount of extra service available, the amount of time needed to instruct the operator and put the bus in place, and the location of the gap with respect to the peak load section when the supervisors are notified. However, in this case, it is worth reviewing because there is a supervisor at Central and the middle of the gap does not pass his location until 30 minutes after the disruption occurred, making this a feasible option if the service is available.

Extra service options include a gap bus or having a pull-in finish the trip. A pull-in is likely to incur overtime for the additional trip and pull-in time from the terminal, whereas a gap bus is already scheduled. The benefits of providing additional service between Central and the end of the route are limited because it is past the peak load section.

---

[37] See the previous footnote.

Ideally, providing extra service in the middle of the service gap at the peak load point would provide the greatest reduction in the number of passengers with excessive wait times, but the supervisors will not receive the disruption notification in time to provide service even at the peak load point. The buses following the gap arrive at the peak load point just as the supervisors are notified of the disruption.

Again, the supervisor will not have the actual passenger counts and vehicle locations from which to consider specific results, but he can estimate that there will be few, if any, missed passengers, but the gap follower will be absolutely packed. The negative passenger impacts to those on the disabled and following bus remain the same, regardless of the option selected, because most of the impacts have already occurred by the time supervisors are notified. Even with extra service added at Central, the impacts for passengers boarding before Central remain unchanged, limiting the number of passengers who can benefit because passengers boarding after Central are the primary beneficiaries of this option. None of them wait longer than 1.5 headways and passengers on the buses following the gap feel less crowding after Central. As shown in Table 5-17, while it is a very effective technique, there are few passengers to benefit from the extra service.

**Table 5-17: Extra Service Passengers Impacts and Additional Cost**

| **Location** | **Largest Service Gap** | **Boarding Demand**[38] | **Passengers Waiting >= 1.5 Headways** | **Additional Agency Cost** |
|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 57 (38 * 1) + (38 * .5) | |
| Western – Central | 20 minutes (3 headways) | 104 (99 * .35 * 3) | 52 (104 * .5) | |
| Central – Cumberland | 10 minutes (1.5 headways) | 22 (99 * .15 * 1.5) | 0 | 35 minutes overtime |
| **Total:** | **20 minutes (3 headways)** | | **109** | **35 minutes overtime**[39] |

[38] This information is estimated from historical averages. If buses are not evenly spaced, the passenger loads will vary substantially.

### 3. Expressing

Although CTA cannot effectively express under current operations procedures, this section is included to show what CTA could achieve through changing procedures under their current communications system. For expressing, the following bus that picked up the passengers from the disabled bus could drop off only, allowing passengers to board where it stopped as well as serving all major stops, until he has made up 10 minutes, or half the service gap. This serves the primary goal of reducing the gap, while providing faster service to passengers on the disabled bus and serving major boarding locations. However, assuming the bus can only increase its speed by 10%, he can move up 4.5 minutes in the 45 minutes remaining in this trip. However, from the disabled bus' location, there is only 10 minutes to the peak load point at Pulaski. Therefore, even if the controller could instruct the follower while the passengers from the disabled bus boarded his bus, the service gap at the peak load point would be about 18 minutes with no other interventions made.

Once past Pulaski, the express bus should still try to make up time, but should not miss passengers, even if it slows his progress in regaining the schedule because it will irritate passengers who have been waiting for the bus and the number of passengers benefiting from the shorter headway is small. This is because there will be relatively few boarding passengers once past the peak load point and the first arriving bus will have pulled away from his follower by this point, making the additional wait extremely onerous for passengers who have already waited more than one or two headways.

In the no intervention and extra service cases, the two following buses are bunched and, thus, provide one unit of twice the normal capacity, which together provides enough capacity to serve all passengers. For the first 2-3 minutes, the buses are still bunched and these passengers are not considered missed. However, once the expressing bus has had a chance to catch up some time, they spread out by 2 minutes at the peak load point and the

---

[39] Pull-ins are the most likely source of additional service and they will incur overtime for the extra trip. Using a gap bus would not incur additional cost.

first vehicle arriving by itself does not have enough capacity to serve all passengers in this case. Nineteen passengers are missed between Western and Pulaski and, thus, wait another 1-2 minutes for the second bus, half of whom are assumed to have waited 10 or more minutes for service.

Because of the capacity constraints when serving the peak load point, expressing has a limited ability to reduce the number of passengers waiting at least 1.5 headways because so many passengers could not benefit from the reduced headway provided. This also assumes a limited ability to both make up time and serve all passengers once past the peak load point and has provided a range for the number of passengers waiting at least 1.5 headways because of this uncertainty. This is more difficult to estimate because of the changing headway as the bus expresses. The case results are summarized in Table 5-18 below.

**Table 5-18: Expressing from Disruption Location (Western) Passenger Impacts**

| Location | Largest Service Gap | Boarding Demand | Second Bus Headway | Missed Passengers | Total Waiting >= 1.5 Headways |
|---|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 6.5 | 0 (76 – 150) | 57 (38 * 1) + (38 * .5) |
| Western – Pulaski | 20 minutes (3 headways) | 71 (99 * .25 * 2.9) | 0-2 | 13 (71 * .25 * .75[40]) | 34 (58 * .47) (13 * .5) |
| Pulaski – Cumberland | 18 minutes (2.8 headways) | 69 (99 * .25 * 2.8) | 2 | 0 | 30 (69* .44) |
| **Total:** | **20 minutes (3 headways)** | | | **13** | **121** |

[40] The buses are bunched for the first 25% of this segment, so these passengers are not considered missed.

### 4. Extra Service + Expressing[41]

The last case considered is adding extra service and expressing service from Central. This case uses the expressing technique and adds the extra service in the same ways as the previous expressing and extra service cases. The expressed bus expresses from the disruption point at Western through Pulaski once it has reached capacity and the extra bus leaves Central one headway behind the gap leader. The service gap at Pulaski (the peak load point) would be about 18 minutes (2.8 headways), at Central would be 8.5 minutes (1.3 headways), and at Octavia would be 7.5 minutes (1.1 headways), but again the bus may not be able to continue reducing its headway after Pulaski because of the need to stop and serve all passengers after the peak load point. As shown in Table 5-19, the improvement in this case is also limited because most of the passenger impacts occur at the peak load point and the extra service can only be used at the end of the route.

**Table 5-19: Extra Service + Expressing Case Passengers Impacts and Additional Cost**

| Location | Largest Service Gap | Boarding Demand | Missed Passengers | Passengers Waiting >= 1.5 Headways | Additional Agency Cost |
|---|---|---|---|---|---|
| Initial Load (Disabled + Following Bus) | 20 minutes (3 headways)/ 13 minutes (2 headways) | 75 (Maximum Load) (38 * 2) | 0 (76 – 150) | 57 (38 * 1) + (38 * .5) | |
| Western – Pulaski | 20 minutes (3 headways) | 72 (99 * .25 * 2.9) | 13 (71 * .25 * .75) | 34 (58 * .47) (13 * .5) | |
| Pulaski – Central | 18 minutes (2.8 headways) | 28 (99 * .1 * 2.8) | 0 | 12 (28 * .44) | |
| Central – Cumberland | 9 minutes (1.4 headways) | 21 (99 * .15 * 1.4) | 0 | 0 | 35 minutes overtime[42] |
| **Total:** | **20 minutes (3 headways)** | | **13** | **103** | **35 minutes overtime** |

---

[41] This response can also only be used if pre-planned and is shown as a potential option if operators were instructed by agency policy before the disruption occurs to express from the disruption location to the peak load point.

[42] Assumes a pull-in was used to make the additional trip on overtime.

### Comparing the Service Restoration Options

The supervisor must now compare his options and select the one that best satisfies agency values. The primary agency value judgment that must be made, whether set formally or left to the case-by-case discretion of service managers, is how much weight to place on each of the evaluation measures. Deciding between agency cost and reducing passenger impacts is a clear case of competing objectives, but deciding between missing passengers and reducing long passenger delays may be a more difficult question. Values will vary across agencies and within agencies over time, but value judgments will be made, whether or not agencies formally develop a policy.

The options in this case illustrate some of these interesting trade-offs and constraints. Expressing would be effective with a full bus before the peak load point, but once it reaches the supervisor at Central, it hurts more passengers than it benefits. It is not perfect, though, because it creates a short service gap between the two following buses before the peak load point and some passengers must wait for the second bus. The only way to provide enough capacity with the first set of vehicles is to keep the buses bunched through the peak load segment, which ultimately creates a longer wait for more passengers. The benefits of extra service are also limited because it cannot be put in service until after the peak load point. The in-route service restoration options are summarized in Table 5-20 below.

**Table 5-20: CTA Context Summary of Initial Trip Passenger Impacts and Agency Costs**

| Service Restoration Option | Missed Passengers | Passengers Waiting >= 1.5 Headways | Maximum Wait Time | Additional Agency Cost | Considered [43] |
|---|---|---|---|---|---|
| No Intervention | 0 | 132 | 3 headways (20 minutes) | 0 | Yes |
| Extra Service | 0 | 109 (17% less)[44] | 3 headways (20 minutes) | 35 minutes overtime | Yes |
| Express[45] | 13 | 121 (8% less) | 3 headways (20 minutes) | 0 | No |
| Express + Extra Service[46] | 13 | 103 (22% less) | 3 headways (20 minutes) | 35 minutes overtime | No |

First, the manager should quickly review the options to see if there are any dominant options. In this case, there is no such option. CTA has only two likely options: add extra service at the supervisor's location at Central or do nothing. The supervisor would probably add a bus if he can find a pull-in. Assuming an extra vehicle is available, which is unlikely between 4 and 5 pm, and if the agency is not concerned about spending 35 minutes of overtime on passengers facing a wait of up to 3 headways, 29 passengers can be saved from waiting more than 1.5 headways. This is effectively at a cost of 1.2 minutes of overtime per "saved" passenger. If an extra vehicle cannot be found, there is nothing the point supervisor at Central should do within the CTA context to restore service on the initial trip. It is best for him to do nothing.

While adding the Express and Express + Extra Service options provides CTA with more options, the primary source of service improvement is adding extra service. Expressing from the disruption site saves 11 passengers, or 8% of the passengers, affected in the "do nothing" response case from large impacts, while the Express + Extra Service saves 29 passengers, or 22% compared with the "no intervention" response.

---

[43] Express options are not considered under current CTA policy because the instructions would have to be a standing policy, such that the operator would know immediately that he should express from the disruption location to the peak load point.

[44] When compared to the "no intervention" case.

[45] From disruption location at Western.

Adding extra service is the greatest single intervention that can be made, but if expressing can be done immediately, it does not create additional agency cost and can improve a few passengers' experiences. This can be underway while the supervisor reviews his extra service options and tries to identify a pull-in or other extra service.

### b) Long-term Response

As discussed in Section 5.4, a more permanent response is needed if the disruption is anticipated to affect future trips. The supervisor can either replace the missing service to maintain the existing operating plan or develop a new operating plan, given that the prior headway is unachievable under the disruption conditions. Because this disruption is a service gap and adding a vehicle would allow the scheduled operating plan to work, he should first determine whether there is a no-cost service option available.

The options available in the CTA context are slightly limited compared to the ideal options. CTA rarely has extra operators or vehicles during the morning and afternoon peak hours. Thus, this case study assumes service replacement options are unavailable. The other ideal options: re-spacing service from the terminal, or using short-turning or deadheading to concentrate service to the main trunk section, are available in the CTA disruption response context. They depend on getting supervisors into position in time to instruct operators due to the limited communication ability with operators. These options were discussed in detail for the ideal case in Section 5.4 and are shown in Table 5-21 below.

---

[46] See previous footnote.

**Table 5-21: Long-Term Disruption Response Options**

| Service Restoration Options | Max Peak Load | Missed Passengers | Passengers Waiting >= 1.5 Headways[47] | Additional Agency Cost |
|---|---|---|---|---|
| Re-space | 65[48]<br>(50 * 1.3) | 0 | 0 | 0 |
| Short-turning | 50 | 0 | 2[49]<br>(71 * .05 * .23) +<br>(99 * .05 * .23) | 0 |
| Dead-head | 50 | 0 | 1[50]<br>(99 * .05 * .23) | 0 |

## c) Maintaining New Operations Plan and Returning to Original Plan

Maintaining the original operations plan by adding extra service or a new operations plan by re-spacing the service is straightforward. For re-spacing, the two operators should leave the terminal on future trips 2 minutes early or late. Once the disruption concludes, returning to the original plan can easily be done through dispatching future trips according to the original schedule or by attempting to get the vehicles back on schedule while in-route.

Short-turning and deadheading require more interaction. The supervisor must identify which runs should be short-turned into the vehicle gap and whether there is enough time to deadhead vehicles from the Nature Museum terminal to the Red Line. Once the disruption is over, the supervisor must return the operators to where they should be with respect to their run paddles.

---

[47] Only occurs if reverse direction is not spaced.
[48] Boarding spikes are likely at the rail stations and could put some vehicles over capacity.
[49] Short turning eliminates the short-turned operator's recovery time and the operator will need to be put back in place at some point.
[50] Dead head time comes out of recovery time, so it may not be reliable.

### 5.5.5. Service Restoration Implementation

The disruption response implementation is done in four parts. They are the initial response (short-term management of the remainder of the disrupted trip), the transition into the long-term management (terminal-based management of re-spacing the service), maintaining the new operations plan, and returning to the original plan once the disruption concludes. However, while planning should be done in order due to time constraints, instructions to operators need not be given in the same order; especially if communication between the service manager and operator is by direct contact, as it is between supervisors and operators at CTA.

#### Communication and Service Managers Needed

Service restoration implementation depends on the agency's communications system. Available service options and the timeline of disruption response depend on the ability to communicate relevant information to the parties involved as soon as possible. While the communications system ultimately allows this, it also serves to constrain disruption response because no system can provide all information to all parties instantaneously.

Specifically, it is important to know how effectively the radio channels can be used to coordinate disruption response among service managers and whether or not service managers can communicate directly with operators. There are currently several communications limitations for CTA bus operations. CTA supervisors cannot communicate directly with operators. Voice contact between operators and controllers require the operator to stop the bus and take the call (thus, delaying service), controllers must make a phone call to garage personnel to order replacement buses. Finally, the radio channels are too full to allow a coordinated disruption response for each incident (Barker, 2002). These limit the service restoration options available, as highlighted in the differences in options available in the ideal and CTA context cases.

In this case, up to six service managers are needed (3 mobile supervisors, 2 point supervisor, and 1 controller). One mobile supervisor responds to the disruption scene and works with the disabled bus to try to fix the particular problem it is facing[51] to return it to service as quickly as possible. The point supervisor at Central is needed to instruct operators to put extra service in place. The point supervisors at Halsted would put short-turned service in place for additional westbound trips or instruct eastbound operators to deadhead back. if that option were selected instead. Mobile supervisors at the terminals would be used to ensure proper dispatching for the next eastbound and westbound trips. Lastly, the controller would be needed to instruct the gap follower to express while he is loading the disabled bus' passengers.

### 5.5.6. Monitor and Reassess

There are two feedback loops. The first is in real-time. The supervisor evaluates the effectiveness of his plan and the ability of operators to carry out his instructions. as well as monitors changing conditions and available resources. This is important to identify new opportunities and impediments to his plan. With the limited vehicle location and passenger load information available at CTA. supervisors will not have this information readily available in real-time or even after the disruption. He would need to speak with the instructed operators and supervisors to get an idea of how his response worked or look at the AFC and AVL data collected by different departments in a different CTA office.

The second feedback loop is used for the agency to set its service management and disruption response standards. Policy makers should look at how disruption response is handled in different situations and recommend best practices as the preferred response.

### 5.5.7. Disruption Response Process Findings

This case study illustrates several key disruption response management points. The first is that accurate route profile information known in advance of the disruption is essential to disruption response. The second is drawn from the first observation: service restoration

[51] For example. for equipment failures. he tries to repair the vehicle and for accidents or other disturbances.

options are developed with the specific route, disruption, and situation at the time of disruption in mind. Therefore, while the rules of thumb are the same, disruption response is different in each case. The third is the importance of focusing the disruption response on the peak load section of the route. The fourth is the importance of receiving disruption reports and starting the disruption response process as soon as possible because of the need to instruct operators before they reach the peak load point. The last is that service managers need a process to guide them through the response development, selection and implementation stages, rather than a series of specific things not to do.

As the case study shows, service restoration options are developed and selected based on the passenger flow profile, making accurate information essential to disruption response. Service managers cannot make intelligent disruption response decisions if they do not have a reasonable understanding of where passengers are, where they are going, how long they have likely waited for service, when the next vehicle will be arriving and how full the vehicles are. Grossly inaccurate information can be more detrimental than a lack of information because of the risk of service managers implementing misguided service restoration strategies, rather than choosing a conservative option because he knows the limitations of the available information.

Secondly, different options are available depending on the location and timing of the disruption, as well as the route characteristics. Passenger load profiles (both boarding / alighting distribution along the route and peak load segment information), average V/C ratios, the operations plan, service manager locations (particularly for agencies with a limited communications system), vehicle loads and locations when service managers are notified, and the location of the service gap and peak load point are important factors in determining the options available. For example, generally, adding service in the middle of the gap will be the best option to reduce passenger delay. However, as seen in the scenario analyzed, extra service does little to improve the experience of most passengers if added after the peak load point and most passengers have boarded, making the

---

he secures the scene.

expressing options better because they can be implemented in time to improve most passengers' experience.

This leads to the third point: the disruption response must focus on improving the experience of passengers through the peak load segment because this is where capacity problems are most likely to arise and where even small adjustments can lead to significant reductions in passenger impacts. Likewise, large improvements provided to passengers, such as providing extra service, after this point do very little good if there are few passengers to benefit.

The fourth point is the need to receive information about, and respond to, the disruption as quickly as possible. As time passes, the gap travels along the route, creating more delays for passengers, and making it increasingly likely that the gap leader has already passed the peak load point once the service manager is ready to act, thus eliminating the holding service restoration technique.

The last key point is that service managers need a process to follow to develop and select disruption responses that reduce passenger impacts and are consistent with agency values. This gives them a framework from which to approach the important decisions they must make affecting how passengers experience transit by outlining the information needed, how to use the information, the options available, and how to select which options to implement. Ultimately, this provides an environment in which their objectives are clear and action is encouraged, rather than an environment that encourages them to do nothing for fear of making service worse because they do not really understand how these elements interact.

Without a framework to guide their decision-making process, rules of thumb are likely to exist to protect passengers from the most misguided service restoration actions, such as short turning from the primary demand direction to fill a gap in the reverse direction. However, they do not provide service managers with guidance as to what they should do, resulting in the "do nothing" option be adopted all too often. Thus, it creates an environment where service managers cannot do a, b, and c, but does not indicate they should do x, y, or z, or even help them select among competing reasonable service restoration options.

## 5.6. Disruption Response Support Process

As described in Chapter 4, the pre-planning process has two primary objectives. The first is to use pre-planning to improve the chances that an effective disruption response will be selected when a disruption occurs despite the limited time available for analysis and decision-making during a disruption. The second is to protect critical service from disruptions to the extent possible. Both a hold-in and standing vehicle case are considered to illustrate what elements of the response should be pre-planned. The standing vehicle case from Section 5.5 is also used as an example of what should be pre-planned and how it can improve the agency's response.

### 5.6.1. Standing Bus Case

The standing bus case developed in the previous section illustrates why and how pre-planning would be useful for disruption response improvement. There are elements that can be pre-planned in each of the disruption response process. Generally, enough information can be collected in advance that agencies should try to develop guidelines for each route by time of day, direction, and location of disruption that eliminates infeasible restoration options from consideration and, ideally, guides them to the most beneficial one or two options. This section steps through this process and highlights the elements that can be pre-planned for each step.

#### Information Gathering

The elements of the information gathering step, discussed in Chapter 4, are reviewed in Figure 5-5 below. Service managers need to implement service restoration as quickly as possible to minimize passenger impacts. Therefore, information gathering during the disruption should be limited as much as possible so only information not available before the disruption, such as the specifics of the disruption, must be collected once the disruption occurs.

**Figure 5-5: Information Gathering Elements**

| *Route Characteristics and Disruption Information Gathering* | | | | |
|---|---|---|---|---|
| Nature of Disruption | Route Profile | Operations Plan | Current Conditions | Available Resources |

The route profile and operations plan are clear examples of elements that should be known in advance. Service managers should have a general idea of these two elements and should carry references to find specific information once the disruption occurs, such as trip departure and relief times. As the service managers are out in the field monitoring service, they should also be aware of current conditions and available resources. Lastly, the nature of the disruption can also be considered in advance, in the sense that the service manager should have a plan for a variety of disruption scenarios given the conditions of the first four elements.

## Service Goals

Pre-planning the service goals is really identifying in which situations there will not be enough service provided to serve all passengers with the first vehicle or where runs will develop problems with respect to the work rules. Insufficient capacity depends on the disruption gap size, capacity provided, and passenger demand.

As shown in Sections 5.4 and 5.5, the impact of a standing bus depends on where along the route the disruption occurs by time of day and direction and how quickly service restoration can occur. The agency should have all the relevant information in advance to advise the service manager of the conditions under which capacity problems will occur. The agency could also indicate runs within one trip's worth of time of their use limits in the operations plan[52], so the service manager would not have to determine whether he can use each operator he may like to.

### Develop Service Restoration Response

To use his time effectively during disruptions, service managers should have potential leads for likely or severe disruption situations and then pursue each one to see if they are still available, in order of estimated benefits, if there is a disruption. The agency should limit the service manager's list of considered options for the initial and maintained disruption responses, as much as possible, so that he only spends time considering the best options given the disruption situation. This means eliminating both infeasible and ineffective disruption responses and identifying the most promising one or two options, as highlighted in the ideal and CTA-context case studies above. This information would reduce wasted service manager time.

Pre-planning returning to the original plan once the disruption ends is difficult to anticipate. Rather, a few helpful tools may be developed. The first is developing a list of runs that would encounter work rule constraints before the affected vehicle was scheduled to pull-in or the end of the peak period when other vehicles would be available. The second is identifying the transition back to the original operations plan at the likely conclusion of the event, such as those listed previously. However, this has limited use; it can only be used when one option is almost always going to be implemented.

Thus, standing bus disruption responses should be pre-planned generally for the location of failure by route segment and time period (morning peak, midday, afternoon peak) for each route, but not for a specific run or situation. Ideally, a reference that indicates the gap size needed before capacity problems occur, the achievable headway with a missing vehicle, the immediate and long-term agency-feasible response options available, their impacts and costs, and implementation steps should also be available to service managers.

---

[52] This is the Supervisor Guide at CTA.

### 5.6.2. Hold-in Bus Case

While the case study considers a hold-in on a high frequency route, as mentioned previously, all runs on low frequency routes are essential to providing reasonable service. Therefore, hold-ins on low frequency routes should be avoided and pre-planning should focus on replacing the service.

The hold-in disruption considered is a common last-minute hold-in that could not have been foreseen. A full-time operator calls in sick 30 minutes before his report time. He was scheduled to operate a "swing block run," which drives 4-5 hours through the morning peak into late morning that ends with a relief, takes a break then makes a relief and works for another 4-5 hours through the afternoon peak into early evening. This is a critical run because it is in service during both of the peak periods and is involved in two reliefs.

Hold-in disruptions are pre-planned differently than standing vehicle disruptions. As discussed in Chapter 3, the primary difference is the ability to channel the missing service where it is least disruptive to passengers. The agency can select which run or vehicle block is affected and prepare for the disruption with a few minutes advance notice. This allows agencies to minimize passenger impacts by targeting where the service is removed and managing it from the terminal, in advance. The agency should develop a list of hold-in candidates and entirely pre-plan each of their disruption responses. Because the same limited set of hold-ins make it into the field when there are hold-ins, the disruption responses should be well understood by service managers and, hopefully, the involved operators. Thus, hold-in pre-planning consists of identifying which runs should be held-in and pre-planning each of those specific disruptions and should not focus on managing all hold-ins that could happen.

Swing block runs are too critical to hold-in. Therefore, an operator should be assigned to the entire run (or one operator assigned to the morning piece and another to the afternoon piece). For the morning, a less critical run from the established hold-in list should be held-in, but, ideally, with several hours advanced notice, the afternoon portion of the run should be covered by an extraboard or another part-time operator called in to cover it and no service missed.

## 5.7. Agency Policy Assessment Process

Problems will always exist to some extent, but agencies can set their policies to reduce the number of disruptions facing operations and minimize passenger impacts when they do occur. This section steps through the policy assessment process described in Chapter 4. It first identifies policies that commonly affect standing vehicles and hold-in disruptions, evaluates their impacts in the CTA case, and recommends policy settings for CTA. Then, it evaluates the impacts of one policy area, available resources, on passenger impacts in the examples given in Sections 5.5 and 5.6.

### 5.7.1. Identify Agency Delay Sources and Impacts

In order to properly prioritize policy areas to assess, the agency must understand the source of its delays and disruptions and the impacts of these disruptions on passengers. Therefore, agencies should collect this information because disruption sources are specific to the agency and will change over time as different sources are addressed. Tracking this information over time will also allow agencies to see the effectiveness of different initiatives on reducing disruptions.

Disruption source information for the CTA specifically is discussed below. Other agencies will likely have different sources of disruptions, depending the age of their fleet, the average operators experience level, weather conditions, the general driving conditions and environment, and the incidence of passenger disturbances.

Another note of caution for agencies is to be careful about applying system-wide or garage-level disruption source information to individual routes. Ideally, this information should be available at the garage- or route-level, depending on the source. For example, vehicle and staffing problems are likely to be at the garage-level, assuming the pool of vehicles, maintenance staff, and operating staff are assigned to a garage, while capacity or insufficient running time problems would be at the route-level.

### CTA Hold-in Sources

Of all runs held-in at CTA in 2001, 63% were due to equipment and 36% were due to personnel[53]. Trippers are the hold-in of choice; they comprise 25% of runs and 93% of runs held-in, indicating full runs are protected whenever possible. They were also more than twice as likely (63% vs. 36%) to be held-in than full runs. Virtually all (99.5%) equipment hold-ins are trippers, while 96% of regular runs were held-in due to missing personnel, highlighting the flexibility of equipment assignment and the rigidity of operator assignment.

### CTA Standing Bus Sources

The source of standing vehicles is important to understand and address in order to reduce the negative impacts of standing vehicles. CTA's field-based disruption sources and delayed or missed service hours spent on each are summarized in Table 5-22 below.

**Table 5-22: Sources and Delay for Standing Bus Disruptions[54]**

| Disruption Source | Disruption Events | Disruption Delay Time |
|---|---|---|
| Equipment Failures | 74% | 77% |
| Disturbances | 10% | 5% |
| Accidents | 5% | 13% |
| Miscellaneous | 7% | 2% |

Clearly, once service is underway, the greatest disruption threat is equipment failures. They account for about ¾ of both the number of disruptions and the missed or delayed service time. CTA's database further summarizes the events by type of equipment failures, so Maintenance can see trends in which components are failing. Accidents are also an important contributor to standing bus disruptions because of the

[53] CTA bus garages record hold-in-information.
[54] This information is recorded through a Power Control/Communications Center event database. Controllers complete an electronic event report for each event they manage.

disproportionately high (5% of disruption events vs. 13% of missed service hours) amount of missed service time. This is because the vehicles and operators involved in accidents do not return to service. Disturbances, however, are not as important because they do return to service quickly. Two disruption types considered in this study that are not important delay sources for CTA are problems with reliefs and route blockages, which created 2% and 1% of missing service time, respectively.

### 5.7.2. Identify Policy Areas to Review

Once the agency knows the source of each disruption type, it should identify the primary reasons for its disruptions, such as not filling the schedule for pullout, having vehicles fail while in-service, or having insufficient running time or capacity. The standing vehicle case developed in Section 5.5 is discussed because it is more difficult to respond to once in the field than a pre-planned hold-in.

#### Initial Disruption Response

There are many potential barriers between the ideal and achievable response situations, as described in the case above: the time needed to get vehicles in place, communication among service managers and operators, and the availability of extra service. The first problem is really a problem of response time, which can be reduced, but not eliminated. One way of addressing this, when the vehicles behind the disruption location can do something to reduce passenger impacts, is to pre-instruct them, which would ultimately be included in operator training. The training could be developed to teach operators to identify specific disruption scenarios, so they can respond immediately to some disruptions. Agencies would need to identify situations in which they always want the operator to respond in the same way and can train, or pre-instruct, operators. One example of this is that CTA bus operators are already instructed to stop their follower when they are experiencing a disabling mechanical problem and have their passengers board the following bus.

The second problem highlighted in this disruption scenario is communications. CTA supervisors cannot communicate directly with operators, making it practically impossible to instruct them unless they are at the supervisor's location (Barker, 2002). This will always be a problem during disruptions, since responses require vehicles to begin disruption responses at various points along the route and supervisors can only be in one location at a time. Last, the limited availability of extra resources also constrains response options, and vehicle and operator availability is limited during peak service hours.

**Long-term Disruption Response**

The greatest impediments to successfully implementing the long-term disruption responses are virtually the same: training, communication, and extra resources. Operator and supervisor training are critical to implementing the ideal disruption response because it relies on the supervisor's ability to anticipate the future state of the route, such as where operators will end up after respacing or short-turning with respect to where they need to be, and the operator's ability to understand what the supervisor is instructing him to do. The communications system also plays a large role because of the need to instruct operators remotely while maintaining the new operations and when returning to the original plan. Lastly, having extra resources available allows the supervisor to replace the service without disturbing other runs.

### 5.7.3. CTA Policy Area Assessment

This section discusses policy areas CTA could review to improve disruption response and develops possible policy changes to be evaluated for resource availability and training. Communication is clearly one policy area that limits both the initial and long-term disruption response options. Barker (2002) discussed the state of CTA's communication system, outlined its limitations, and recommended improvements, so this is not included in this study.

## Resource Availability

The availability of extra resources will always be a limiting factor because, if available, extra resources help in all disruptions. It can restore missing service, add needed capacity, or provide another vehicle to meet increased running time requirements. Extra service should be made available when disruptions that are individually unpredictable occur frequently enough to justify making it available. For example, a number of vehicle failures and accidents will occur each day, but the specific vehicles are unpredictable. Agencies also need to establish the number of extra operators and vehicles needed to make the scheduled pullouts based on their anticipated absenteeism and spare ratio needs, which could include anticipated in-service failures. Extra resources provided to cover in-service failures could be determined from a threshold of passenger impacts for different failure events and how likely qualifying disruptions are to occur. Another important question is the acceptable rate of in-service vehicle failure and the maintenance practice modifications required to improve vehicle reliability.

One approach to providing extra service on high frequency routes is to develop a base schedule that is filled every day (thus, virtually eliminating hold-ins), where passengers are can be served by the first bus, but that allows some equipment and operators to be available for use on different routes as needed for single-trip capacity problems, delays, and disruptions. This flexible service structure is different than CTA's current approach of scheduling every available bus and operator and then scrambling to make pull out and juggling which runs should be held-in. To move to this type of schedule, CTA would need to increase its number of available buses[55] or reduce the amount of scheduled service. Reducing the schedule to increase vehicle availability or making vehicles available to supervisors could create concern about having equipment that may not be used on a daily basis. This should lead to a comparison of the costs of having more scheduled service that is less reliable with the costs of a thinner schedule (fewer vehicles and longer headways) at greater reliability.

---

[55] Most (63%) CTA hold-ins are due to vehicle shortages, so this is the primary constraint, not operator availability.

An important question is how extra resources can be made available for disrupted service, but are used each day to supplement the base schedule if not needed to respond to disruptions. One potential way of improving the flexible service option is to move away from having vehicles sitting with supervisors, who choose whether to save it for a disruption or provide extra service along busy routes, and to formalize how service is used when not needed for disruptions. Instead of viewing it as flexible service that gets deployed differently or directly by supervisors each day, pulling service for disruption problems could be handled as an exception. Thus, the flexible service would have a regularly scheduled run and would only be called to fill in as needed. As with hold-in disruptions, there could be a limited set of movable trips from which to take extra service. Their schedules would be designed to have little impact on passengers when taken off their regularly scheduled run and should have entirely pre-planned disruption responses for the route from which they are taken.

## Training

The last area constraining disruption options is operator and supervisor training. The specific concerns for operators are whether they have the driving skills to make up time, the route knowledge to identify which are the major stops while expressing, and the knowledge and experience to understand supervisor instructions to put themselves at the time and location requested.

### Operator Training

The first training area is an issue of driving skills and how to improve junior operators' skills to be closer to those of senior operators who know how to move in and out of traffic safely and quickly and can make up time when asked to. Junior operators' skill levels are currently a concern at CTA, where over 50% of operators have less than 5 years experience with CTA. Besides its constraints on disruption response, varying skill levels is also a concern for bunching problems during normal service. The second, route knowledge, is also likely tied to the operator's years of experience at CTA. To effectively serve only drop-offs and major boarding locations in express mode, operators must know which stops are essential to serve.

The last training issue concerning disruption response is whether operators understand and can properly follow instructions. CTA did not report this as a problem, but most instructions are currently simple, individual instructions. Operators may find it more difficult to follow instructions if several are given at once[56], instructions are sent via the data messaging system, or sustained complex operations, such as short-turning, are used more often.

### Supervisor Training

For supervisors, the training issues are whether they have the skills to understand the available options, develop plans in advance, select from pre-developed plans they have used in the past, think through disruption scenarios and disruption responses to identify future problems, and give operators advance instructions when possible. The first three concerns stem from the need for pre-planning at CTA and highlight its importance. Pre-planning would involve supervisors and other agency policymakers and this involvement in thinking about how to respond to problems would improve their knowledge of the basics, such as passenger boarding patterns along routes, as well as their disruption response development skills.

The last two concerns are also addressed in the informal training that occurs when pre-planning disruption responses, but they focus on a different problem. Thinking through disruptions and potential responses to anticipate future operator work rule or service problems requires some thought and may intimidate supervisors who have dealt primarily with single trip service restoration strategies. CTA service restoration generally relies on instructions that relate to the operator's schedule because of the easy reference for the operator. Giving instructions in advance is not common, but, again, this may stem from most service restoration focusing on improving service for the remainder of the trip.

---

[56] With such a limited communications system, supervisors might need to give immediate and future instructions.

## 5.8. Summary and Findings

This chapter illustrated the disruption response and agency policy assessment processes to disruption scenarios by applying them to the CTA 77 Belmont bus route. The applications found a gap between the disruption responses available in an ideal situation and those available within the current CTA policy context. Ways of closing this gap were discussed and improved pre-planning was suggested as a primary method. Lastly, policy areas constraining CTA's disruption response were identified and policy changes were suggested.

### 5.8.1. Disruption Response Findings

There were several key findings from the disruption response process for the ideal and CTA context cases. Service restoration options are limited by the available information and may be harmful with inaccurate information. Secondly, different disruption responses are feasible under different circumstances and the developed response should reflect those differences. Thirdly, disruption response should try to save the greatest number of passengers from negative impacts, which means the focus should be on the peak boarding segments. Fourth is the need for service managers to begin implementing a disruption response as quickly as possible; the response should be developed and implemented before options become infeasible. Last, these considerations highlight the need for agencies to institute a framework through which to approach service management during disruptions.

The specific passenger impacts estimates support these findings. The ideal disruption response available with extra service would reduce the passenger impacts by 68% from 132 to 42 passengers with an excessive wait time, at a cost of 1.25 hours of overtime. Without extra service available, it could be reduced by 36% to 85 passengers. However, within existing CTA constraints, there are only two options: adding service or doing nothing. Extra service reduces the passenger impacts by 17% to 109 passengers with excessive wait times. If extra service is unavailable, 132 passengers will wait more than 1.5 headways.

### 5.8.2. Indicators of Pre-planning

These cases also illustrated the value of pre-planning. A hold-in and standing vehicle case were considered. Fully developing responses would benefit CTA's disruption response because it would minimize the disruption before even making it into service. In the standing vehicle case, the lack of accurate information in real-time delayed the supervisor's ability to respond quickly, which eliminated all feasible options, except adding extra service. Guidelines developed in advance to guide disruption response for standing vehicles by route segment, time of day and direction were recommended, which would include both the primary service goals and the most promising service restoration options for the initial and long-term responses. A specific recommendation was to pre-instruct operators to drop-off and serve major stops only from the disruption location through the peak load point. Lastly, moving back to the original schedule will need real-time planning, since it would be difficult to estimate in advance all the conditions in which the service manager would need to do this transition. The recommended response to hold-in disruptions was to pre-determine which trippers could be sufficiently managed if held-in and to move operators and vehicles from the manageable tripper hold-ins to the critical blocks and runs.

### 5.8.3. Agency Policy

Lastly, the agency policy assessment application highlighted several findings. The first is the need to track delay and disruption source information to recognize trends and the try to establish the benefits or problems with changes. Most constraints can be reduced, but few can be completely eliminated, such as the lead time needed to respond to disruptions and the effectiveness of the communications system. The agency's constraints affect both the initial and long-term disruption responses. The primary policy areas constraining CTA's disruption response are the communications system, resource availability, and training.

# Chapter 6: Summary, Conclusions, and Future Research

Disruptions are one of the greatest threats to passengers' transit experiences. How the agency responds to disruptions will either allow the worst impacts to occur and result in alienated passengers or save them from long waits and the irritation of being passed by full vehicles. Unanticipated, lengthy service delays are a major deterrence for passengers and could drive passengers to find other ways of making their trips. Procedures to ensure service managers develop effective service restoration plans and have the information and tools necessary to implement the desired response plan are critical to the success of service management during disruptions.

This research focused on developing and applying a process for improving transit service management during disruptions, reducing the incidence of disruption occurrence, and evaluating policy areas affecting service restoration. The processes were applied to the CTA bus system on a high frequency crosstown route, the 77 Belmont. The application developed an unconstrained ideal and a CTA-context constrained disruption response. Pre-planned disruption responses were discussed for hold-in and standing vehicle cases. Lastly, based on the constraints found when applying the first two processes, the policy areas contributing to the constraints were discussed and policy changes to explore were suggested.

## 6.1. Summary and Conclusions

There were three primary objectives: develop a framework for agencies to make use of, develop criteria to evaluate passenger impacts during, and evaluate agency policies affecting transit service management during disruptions. Chapters 1 and 2 provided a background of service management concepts and prior research. Chapter 3 developed a framework for agencies and service managers to think about service management's disruption response and agency policy evaluation, as well as the evaluation criteria they should use. Chapter 4 developed formal processes for developing a disruption response and evaluating agency policies. Chapter 5 applied the processes to a standing vehicle and hold-in disruption on the CTA bus system. The following subsections summarize the approach and findings of each objective.

### 6.1.1. Structuring the Disruption Response Approach

The first objective was to structure an approach to the problem of transit service management during disruptions. The framework defined the primary management problem as a disruption response problem with the secondary issue of how to best support that response and identified the important information service managers need to determine the feasible options to respond to disruptions and select among them. The disruption response process suggested steps agencies could take to respond to disruptions, how to use information to develop a response, and the steps and potential constraints on implementation and the feedback and monitoring that needs to occur to improve disruption response. The pre-planning framework identified the range of pre-planning options, the criteria for determining which disruption situations warrant pre-planning and to what extent, and the concept of disruption reallocation.

#### Findings

There were several key findings. Pre-planning and policy assessments were identified as the primary means of improving disruption response and most of the recommendations pertain to them.

1. Timely and accurate information is critical to the success of disruption response. The boarding profile and peak load point with respect to the disruption location determine the service restoration options available to service managers. Service managers should have readily available all information that can be known in advance. This includes static information, such as route profile and operations plan information, and dynamic information, such as current conditions and the availability and locations of extra resources. Only the specifics of the disruption should need to be gathered at the time of the disruption. To the extent possible, service managers should also identify potential disruptions and promising responses, which may be available from the set of pre-planned

disruption responses. How they collect information could also help lead service managers to focus on managing the whole route, rather than individually instructing operators.

The case study developed in Chapter 5 highlights this. The ideal disruption response reduced severe passengers impacts by 68%, compared with the best possible CTA response, which could only reduce impacts by 17%.

2. Because most of the passenger impacts occur at the onset of the disruption, a quick response is required to eliminate most of the initial trip's passenger impacts.

Most of the impacts affect passengers on the disrupted and immediately following vehicles. Responses will be limited if service managers must develop a response after the disruption has occurred because as time passes, more passengers are impacted and fewer restoration options are available. Depending on the agency's communication system, disruption notification could be delayed, making immediate action even more important. Pre-planning and policy changes can help address this by identifying and removing constraints at the policy level or responding better through a pre-planned response.

3. Most disruptions can and should be pre-planned to some extent, but few can be entirely pre-planned.

Agencies have a limited ability to pre-plan unpredictable disruptions, such as standing vehicles and unanticipated bus route blockages, which places them on the real-time response portion of the spectrum. These disruptions should have guidelines developed for how to respond by route segment, direction, and time of day, but the agency could not reasonably try to pre-plan all instances of them. At the other end of the pre-planning spectrum are hold-ins, scheduled bus route blockages, and rail blockages, which should be entirely pre-planned. The first two can be managed in advance such that the agency has control over how the disruption manifests itself. For example, a list of hold-in candidates should be developed and each of those situations pre-planned. Rail blockages are so disruptive to passengers and their different manifestations so limited, that they should be pre-planned entirely by track segment, direction, and time of day.

4. Low frequency service must be protected.

Agencies cannot effectively respond to disruptions on low frequency service. There are few options available to them because passengers arrive just before the vehicle is scheduled to arrive. Thus, providing even headways off schedule does not substantially decrease passenger impacts. Therefore, disruptions on low frequency service cannot be managed well and the scheduled service must be replaced as soon as possible.

5. CTA disruption response could be improved with increased pre-planning.

CTA has developed hold-in guidelines, but the bus garages do not necessarily have a current list of specific runs or vehicle blocks to hold in. They also reroute service and develop new schedules for planned bus route blockages. However, the agency has not established pre-planned responses to implement when hold-ins do occur or partially pre-planned elements for other disruptions. They informally assign a mix of vehicles to some routes and concentrate their more reliable vehicles on others, but they could explore this more and formalize a program of assigning more reliable vehicles to essential runs.

Another useful rule of thumb for high frequency routes facing a running time problem is pre-determining a new headway that accommodates most disruptions and developing a response for that case. For example, moving from a seven-minute headway to an eight- or nine-minute headway. Another pre-identified element for routes that develop capacity problems quickly should be parallel routes from which to borrow service.

### 6.1.2. Developing Evaluation Criteria

The second objective was to develop appropriate evaluation criteria, or passenger impact and agency cost indicators, that properly value the primary passenger impacts and agency cost during a service disruption, without overwhelming service managers with too many indicators. Many indicators were considered and four selected based on how well they identified the primary passenger impacts and are designed to measure severe impacts. They were also considered for their ease of developing in advance from historic data and ease of estimating in real-time. They are the:

1. Number of Missed Passengers
2. Number of Passengers with an Excessive Wait Time
3. Longest Headway[57]
4. Agency Cost

## Findings

The evaluation criteria do not measure gradual service degradation or improvement and, therefore, do not reflect if the agency has marginally improved passengers' experiences through a reduced average wait time. Rather the agency only has an improved disruption response if fewer passengers are skipped by the first arriving vehicle or suffer a long wait, if the longest service gap is reduced, or agency costs are reduced with all other passenger impact reductions being equal.

### 6.1.3. Agency Policy Assessment

The last objective was to define and evaluate policy level issues that can either assist the service manager in mitigating negative passenger impacts during disruptions or to reduce the number of disruptions occurring. Agency policies can serve to reduce the number of disruptions that occur or reduce the negative passenger impacts when disruptions occur. Agencies can use the policy evaluation process to assess the cost-effectiveness of changing different agency policies that exclude otherwise feasible service restoration options or affect the frequency of disruptions.

## Findings

1. Changing agency policies affecting disruption response is a very powerful tool because of the wide-reaching effects of policies, so it is important to determine whether specific policy changes are worthwhile.

This process helps agencies understand how their policies affect the agency's ability to respond to service disruptions and what policy changes may be cost effective.

---

[57] This approximates the longest wait time for high frequency service.

Additionally, service managers can improve the agency's response to these disruptions by understanding which disruptions are amenable to mitigation or avoidance and how they should be managed.

2. The greatest constraints on CTA's disruption response are time and information. These problems stem primarily from constraints in the communications, training, and extra resources policy areas. There are several constraints on disruption response from the communications policy area. Supervisors and operators cannot directly communicate with each other. Thus, there are significant delays before supervisors, who are responsible for service management, learn about disruptions and also in instructing operators once a plan is developed. Contact between the central bus controllers and operators is also limited because of the need to curb the bus while having a conversation on the radio.

Training is always an important policy area when the agency needs its workforce, operators or service managers, to change their behavior. There are several training avenues that could benefit operators. In 2001, fifty percent of CTA bus operators had less than 5 years experience with the CTA. Defensive driving training reduces accident rates and general driving skills training can improve operators' abilities to be able to implement the service restoration techniques service managers ask them to do. While most agencies, including CTA, have accident reduction and driving skills improvement programs, an area that has not been explored is identifying situations in which operators could respond immediately to disruptions. CTA could benefit substantially in these situations, especially given the long timeframe for supervisors to be notified and respond. CTA supervisor training should focus on developing them to be part of the pre-planning process. This would teach them more about service restoration in general and in the specific cases in which they may be involved in the future.

The other problem is a lack of information, either historic or available in real-time. Supervisors and controllers do not have vehicle location information, passenger load, or passenger demand information, the primary determinants of negative passenger impacts and framers of feasible disruption responses.

3. Increasing extra service availability and its cost-effectiveness.

Agencies would like to have extra service available to respond to disruptions, but having an operator or vehicle sitting idle in case an unpredicted disruption occurs is usually not the best use of resources. Therefore, a system making service available to disruptions, but being productive when not supporting a disruption response would be ideal. As in the case of hold-ins, there could be a list of candidate runs to be borrowed and used during disruptions which were added into a base schedule that provides enough service even when they are missing. Entirely pre-planned responses should then be developed for when each of these runs were taken from the schedule, in order to minimize the disruption to the route from which the service is being taken.

## 6.2. Future Research

This research has provided an overview and strategy for of service management during disruptions, but there are still many issues to explore. The ones suggested here are in-depth analyses of the policy areas and service restoration technique usage for those that have not already been studied, and exploring methods in the feedback loop, such as how to monitor and determine the effectiveness of service managers' disruption response.

### 6.2.1. Evaluation Framework and Assessment

Further research could be done on specific elements of the process and framework. One area is the specifics of how and when to use different service restoration techniques and the future impacts of service restoration techniques. Past studies have evaluated the best use of specific operations control strategies, typically for a one way trip, and this study outlines some primary characteristics that frame the service restoration options available, but comprehensive guidelines could be developed for high frequency routes by loading profile, the availability of extra resources, and initial response time. For example, they would list the likely service goals and most promising one or two disruption responses.

They could also include recommendations for whether operators would be able to identify the situation and do something to respond right away.

Further research is also needed to guide agencies on the specifics of setting their policy areas. While methods for setting the extraboard, determining the proper spare ratio, and the effects of using part-time operators under different work rule environments have been developed, training and work rules are policy areas that have not been studied in depth.

### 6.2.2. Information Availability

Two interesting questions are the benefits of accurate real-time information and how passengers perceive different impacts. This research illustrates the service restoration limitations when service managers do not have general vehicle location and passenger load and demand information, and other studies have shown the improved disruption response when having real-time vehicle location information and vehicles available to respond (Strathman et al, 2001). Showing the benefits of the service managers also knowing the passenger impacts would help allow agencies to be more responsive to passengers, rather than just to operators. This could also include identify or quantify the specific service restoration limitations when missing either of these components.

### 6.2.3. Assessing Passenger Impacts

Another topic that has not been fully explored is how passengers truly value different impacts and how likely are different impacts to deter them from using transit for future trips. One example is that past studies have looked at the relative values of in-vehicle and at stop/station wait times and have spawned the use of weighting factors in subsequent studies. These weighting factors remain flat, while the irritation level of passengers who have waited a relatively long time is probably higher and the irritation of passengers who have just arrived is probably lower than these simple models suggest. Research could be done on how this factor changes over time for passengers who have been delayed for several minutes on a vehicle or who have waited longer than one headway at a stop/station. Thus, their irritation level probably increases geometrically, rather than linearly over time. Knowing the key thresholds of when their irritation level changes in absolute minutes or with respect to the scheduled headway could be used to set the concept of the excessive wait time.

Another area of concern that is poorly understood is which particular disruption events and disruption responses and the frequency of their occurrence drive passengers away. How important is it to passengers to be served by the first set of vehicles compared with waiting a few minutes less and to what extent does real-time information change the effects of these events? Additionally, is there an important threshold of incident frequency that turns passengers away from transit? For example, expressing vehicles creates a second irritation event. The first is the delay and the second is a vehicle passing them. Therefore, expressing would be a poor option if passengers are more irritated with vehicles passing them than waiting a few more minutes and being served by the first set of vehicles.

### 6.2.4. Feedback Evaluation Loop

Lastly, while this study identifies the need to understand the successes and shortcomings of service managers' service restoration, it does not explore potential performance indicators or an effective way to evaluate their performance. Future studies could look at the process of gathering feedback and using it to improve service management in general and disruption response specifically. This may also help service managers identify and mitigate problems before they become disruptions.

## References


Abkowitz, Mark, Eiger, Amir, and Engelstein, Israel. 1986. Optimal Control of Headway Variation on Transit Routes. *Journal of Advanced Transportation*, 20(1).

Abkowitz, Mark and Lepofsky, Mark. 1990. Implementing Headway-Based Reliability Control on Transit Routes. *Journal of Transportation Engineering*, 116(1).

Barnett, Arnold I. 1978. Control Strategies for Transport Systems with Nonlinear Waiting Costs. *Transportation Science*, 12(2): 119-135.

Barker, David P. 2002. Communication, Information, and Responsibility Distribution Strategies for Effective Real-Time Transit Service Management. MS Thesis, Massachusetts Institute of Technology.

Barnett, Arnold I. 1974. On Controlling Randomness in Transit Operations. *Transportation Science*, 8(2): 102-116.

Chicago Transit Authority. 2001a. Bus Garage Manpower Reports. Bus Operations Department.

——. 2001b. Bus Garage Performance Reports. Bus Operations Department.

——. 2000a. Bus Service Supervisor Training Program. Training and Instruction Department.

——. 1999a. Bus Service Workshop Instructor Guide – Refresher Course. Training and Instruction Department.

——. 1999b. Customer Satisfaction Survey. Market Research Department.

——. 2001c. Operator Empowerment Program. Bus Operations Department.

——. 2000b. Traveler Behavior and Attitude Survey. Market Research Department.

Chicago Transit Authority and Amalgamated Transit Union, Locals 241 and 305. 1996. Wage and Working Conditions Agreement.

Deckoff, Anthony. The Short-Turn as a Real Time Transit Operating Strategy. MS Thesis, Massachusetts Institute of Technology.

Eberlein, Xu Jun. 1995. Real Time Control Strategies in Transit Operations: Models and Analysis. PhD Dissertation, Massachusetts Institute of Technology.

Eberlein, Xu-Jun, Wilson, Nigel H.M., Barnhart, Cynthia, and Bernstein, David. 1998. The Real-Time Deadheading Problem in Transit Operations Control. *Transportation Research B* 32(2): 77-100.

Froloff, Edith, Rizzi, Michel and Saporito, Antoine. 1989. Theory and Practice of Bus Service Management. RATP, English Edition, Revision 1, 1994.

Fu, Liping, and Yang, Xuhui. 2001. On Design and Implementation of Bus Holding Control Strategies under Real-Time Information. Transportation Research Board, 81st Annual Meeting.

Holroyd and Scraggs. 1966. Waiting Times for buses in Central London. *Traffic Engineering and Control*. 8(3):158-160.

Koffman, David. 1978. A Simulation Study of Alternative Real-Time Bus Headway Control Strategies. Transportation Research Board, 57th Annual Meeting.

Levinson, Herbert S. 1991. Supervision Strategies for Improved Reliability of Bus Routes. NCTRP Synthesis Report 15. Transportation Research Board.

MacDorman, Littleton. 1986. Use of Part-Time Operators. NCTRP. Synthesis Report 9. Transportation Research Board.

O'Dell, Susan. 1997. Optimal Control Strategies for a Rail Transit Line. MS Thesis, Massachusetts Institute of Technology.

Ortiz, Iris. 2000. Analysis of Real-Time Operations Control Strategies for Tren Urbano. MS Thesis, Massachusetts Institute of Technology.

Osuna, E.E. and Newell, G.F. 1972. Control Strategies for an Idealized Public Transportation System. *Transportation Science*, 6: 52-72.

Puong, André. 2001. A Real-Time Train Holding Model for Rail Transit Systems. MS Thesis, Massachusetts Institute of Technology.

Rahbee, Adam. 2001. Rail Transit Operations Analysis: Framework and Applications. MS Thesis, Massachusetts Institute of Technology.

Shen, Su. 2000. Integrated Real-Time Dispatching Recovery Strategies: A Model for Rail Transit Systems. MS Thesis, Massachusetts Institute of Technology.

Song, Wei. 1998. Real Time Dispatching Control in Transit Systems. MS Thesis, Massachusetts Institute of Technology.

Strathman, James G., Kimpel, Thomas J., Dueker, Kenneth J., Gerhart, Richard L., Turner, Kenneth, Griffin, David, and Callas, Steve. 2001. Bus Transit Operations

Control: Review and Experience Involving Tri-Met's Automated Bus Dispatching System. Transportation Research Board, $80^{th}$ Annual Meeting.

Turnquist, Mark. 1978. A Model for Investigating the Effects of Service Frequency and Reliability on bus Passenger Waiting Times. *Transportation Research Record* 663, 70-73.

Turnquist, Mark and Blume Steven. 1980. Evaluating the Potential Effectiveness of Headway Control Strategies for Transit Systems. *Transportation Research Record* 746, 25-29.

Welding, P. 1957. The Instability of Close Interval Service. *Operations Research Quarterly*, 8(3): 133-148.

Wilson, Nigel H.M., Macchi, Richard, Fellows, R., and Deckoff, Anthony. 1992. Improving Service on the MBTA Green Line through Better Operations Control. *Transportation Research Record* 1361, 10-15.