Accessible Transportation Technologies Research Initiative (ATTRI) Performance Metrics and Evaluation

Evaluation Framework Report

www.its.dot.gov/index.htm

Task 2 Technical Report—July 2020
FHWA-JPO-20-784

U.S. Department of Transportation

Produced by Cambridge Systematics, Inc.
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Federal Highway Administration
Federal Transit Administration
Intelligent Transportation Systems Joint Programs Office

Notice

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-JPO-20-784		

4. Title and Subtitle	5. Report Date
Accessible Transportation Technologies Research Initiative (ATTRI) Performance Metrics and Evaluation, Final Evaluation Framework Report	July 2020
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Anat Caspi, Mark Hallenbeck, Shannon Tyman	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Cambridge Systematics, Inc. 101 Station Landing, Suite 400 Medford, MA 02155 Washington State Transportation Center Taskar Center for Accessible Design University of Washington 1107 NE 45th St, Suite #535 Seattle, WA 98195-4802	11. Contract or Grant No. DTFH61-16-D-00051
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
U.S. Department of Transportation ITS Joint Programs Office 1200 New Jersey Avenue, SE Washington, DC 20590	Final
	14. Sponsoring Agency Code HOP

15. Supplementary Notes

16. Abstract

This report provides a framework that can be used to develop an evaluation of technology applications designed to remove barriers to transportation for people with visual, hearing, cognitive, and mobility disabilities. The framework helps evaluation teams follow a consistent process that starts with an understanding of 1) the goals of the technology; 2) the needs of the intended user population; and 3) the intended technical performance of the system, to create a logic model that identifies the evaluation hypotheses to be tested, the appropriate performance measures for evaluating the those hypotheses, the data sources from which the required data can be obtained, and the areas where collaboration and agreement are needed between the evaluation team, the technology developers, and the evaluation sponsors to finalize and perform the evaluation.

The framework specifically accounts for—and provides guidance for—evaluations of a wide variety of accessibility development projects (ADP) and is designed to help project sponsors, participants, and independent evaluation (IE) teams focus their evaluation efforts on the key outcomes of importance for each ATTRI-funded ADPs being studied, while also keeping in mind the need to understand the effect the ADP has on overall trip making capabilities of users. Thus, the framework is designed for use with all types of projects sponsored by U.S. Department of Transportation's (USDOT) ATTRI Program.

17. Keywords	18. Distribution Statement
Accessible Transportation Technologies Research Initiative, wayfinding, pre-trip, scenario, logic model, Accessibility Development Projects	No restrictions

19. Security Classification. (of this report)	20. Security Classification. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	128	N/A

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	Pound force	4.45	newtons	N
lbf/in^2	Pound force per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Table of Contents

List of Tables

Chapter 1. Introduction

Report Introduction

This report outlines a technology project evaluative framework that takes a holistic approach to mobility and transit and that is specifically focused on whether technologies being developed and deployed change the accessibility of travel modes, are usable by their target populations, and improve that populations' ability to travel. The report describes the process, logic models, and performance metrics that can be used to evaluate the wide range of Accessibility Development Projects (ADP) whether supported by Accessible Transportation Technologies Research Initiative (ATTRI) or via other mechanisms.

This evaluation framework can be used to develop and then perform independent evaluations (IE) of the ATTRI's funded development projects. The framework is focused on evaluating the performance of tools or technologies that have previously been selected by ATTRI and its partners to meet identified needs or help travelers surmount identified barriers to travel. Evaluating the needs of users and identifying the nature of barriers were the subjects of previous ATTRI efforts and are not addressed in this framework other than applying the results of those previous ATTRI projects.

The report is structured into four chapters:

- Chapter 1 contains introductory material.
- Chapter 2 introduces a number of very important concepts that are key to understanding the evaluation process.
- Chapter 3 presents the actual evaluation process.
- Chapter 4 presents a detailed example of how the framework might be applied.
- Finally, the report includes appendices that provide assistance to IE teams in the development of logic models and the selection of performance metrics.

Project Background

The U.S. Department of Transportation's (USDOT) ATTRI is a joint USDOT initiative, co-led by the Federal Highway Administration (FHWA), Federal Transit Administration (FTA), and Intelligent Transportation Systems Joint Program Office, with support from the National Institute on Disability, Independent Living, and Rehabilitation Research and other Federal partners.

The ATTRI Program is leading efforts to develop and implement transformative applications to improve mobility options for all travelers, particularly those with disabilities. ATTRI research focuses on removing barriers to transportation for people with visual, aural, cognitive, and mobility disabilities through every step of the trip-making process.

ATTRI seeks to remove barriers to transportation across the “complete trip” chain, leveraging advanced technology to enable people to travel independently at any time, to any place, regardless of their individual abilities. ATTRI intends to improve the ability of all people to travel in an efficient and affordable manner, emphasizing transportation system improvements that allow individuals with disabilities (and all travelers) to reliably, safely, and independently plan and execute seamless, complete trips, from origin to destination.

This report defines complete trips in terms of an individual’s ability to plan for and complete a trip from origin to destination without gaps (disruptions) in the travel chain. The links of this chain include trip planning, travel to a station, station/stop use, boarding a vehicle, using a vehicle, leaving a vehicle, using the stop or transferring, and traveling to a destination after leaving the station/stop. If the traveler is not able to complete one step in this chain of activities, then the trip cannot be completed, decreasing overall accessibility for the individual unable to make the trip.

Through extensive research and outreach, the ATTRI program has identified four key areas for technology development with the potential to address gaps in the mobility of people with disabilities:

- Smart wayfinding and navigation.
- Pre-trip concierge and virtualization.
- Safe intersection crossing.
- Robotics and automation.

ATTRI-funded development projects across the four technology areas should work together to enable more individuals to complete more trips, providing the basis for a more inclusive and effective transportation network that is far more economical, expansive, and welcoming than exists today. This report provides a framework for evaluating ADP across all four technology areas regardless of where funding for those ADPs comes from, although the guidance does focus on outcomes of particular importance to ATTRI. Guidance on evaluating the spatial, temporal, economic, physiological and social impacts of the ATTRI efforts can be found in the reports Shared Mobility and Equity Primer and Mobility Performance Metrics (MPM) for Integrated Mobility and Beyond.[1,2]

Introduction to the Evaluation Framework

This evaluation framework specifically accounts for—and provides guidance for—evaluations of a wide variety of accessible transportation technology related projects, including helping project sponsors, managers, participants, and IE teams focus their evaluation efforts on the key outcomes of importance for each individual ADP being studied, while also keeping in mind the need to understand the effect the ADP

[1] Travel Behavior: Shared Mobility and Transportation Equity, for the Federal Highway Administration, by Susan Shaheen, et. al., August 2017, Report #PL-18-007.

[2] Mobility Performance Metrics for Integrated Mobility and Beyond, for the Federal Transit Administration, by TransitCenter, Applied Predictive Technologies, and Texas A&M Transportation Institute, February 2020, FTA Report # 0152.

has on overall trip-making capabilities of users. That is, how has the ADP changed the ability of its users to travel more spontaneously and flexibly? Are they better able to make complete trips, and if not, why not?

The recommended ATTRI evaluation should be set up and performed using the following steps:

- Identify the key details needed for the evaluation, which includes the IE (with support or input from the ADP team and sponsors) performing the following tasks.
  - Review the goals and expected performance/outcomes of the ADP being evaluated.
  - Characterize the primary target population for the ADP.
  - Document the travel requirements, perspectives, and needs of that target population.
  - Understand the ADP technology, including:
    - The travel outcomes it intends to achieve or the travel barriers it is intended to remove.
    - The technological steps it will perform.
    - The user interactions with the ADP technology needed to achieve those outcomes.
  - Determine the key objectives (results) that the evaluation sponsors wish to learn from the evaluation effort.
- Develop a threat model for the ADP functions and required user interactions for each of the affected travel activities.
- Establish the specific travel activities to be affected/improved by the ADP.
- Develop an initial logic model based on the above information that serves as a guide for the evaluation and estimate a budget for the activities described in that logic model.
- Work cooperatively with the project sponsors and ADP team to refine the logic model, data collection plan, and resulting evaluation plan to match the scope of work with the available budget, given the key objectives developed above.
- Perform the evaluation activities.
- Perform a qualitative complete trip evaluation and continuing needs assessment based on the outcomes of the evaluation activities and the target population user needs.

Each of these activities is described in detail in chapter 3 of this report. While the list above provides a useful order in which to approach these topics, the need to work interactively with both the ADP team and the evaluation sponsors may result in the IE team approaching these tasks in a different order as the evaluation proceeds.

Chapter 2. Key Concepts Necessary for Evaluating Accessibility Projects

This chapter briefly introduces key sets of concepts that are important throughout the evaluation process. Due to the complexity of the different types of evaluation concepts, this chapter only introduces this topic. More details on each concept can be found in the appendices. These details become more important the more technically detailed the independent evaluations (IE) is that is, the more the IE is focused on the technical performance of the Accessibility Development Projects (ADP) and its use by the subject population.

These concepts include:

1. **Understand Target Population, User Needs and Barriers—**This first of the key concepts are the user needs and barriers to travel for the target population. The goal of the Accessible Transportation Technologies Research Initiative (ATTRI) program, and for any accessible transportation technology, is to remove barriers (whether structural, systemic or circumstantial) that impact travelers negatively, thereby instrumenting individuals to achieve their travel goals more effectively. Therefore, it is important to understand baseline traveler requirements, needs and barriers in order for the evaluation to measure whether the ADP intervention can effectively facilitate improvements in achieving travel goals. A starting point for this is the *ATTRI User Needs Assessment: Stakeholder Engagement Report*, published in May 2016. These needs were identified through stakeholder coordination. Any specific project may identify requirements beyond these, and the analyst/project manager would need or want to know the unique needs for their project.
   - **Disability Types—**The U.S. Census report, *Americans with Disabilities: 2010* categorizes types of disabilities into communicative, physical, and mental domains.[3] It is important to note that people can have multiple disabilities. The ATTRI team has adopted these definitions with certain modifications. To facilitate development of technological solutions designed to address a specific functional requirement, the ATTRI team divided the “communicative domain” into visual disability and hearing disability. In this document, the term cognitive disability is used in place of “mental domain,” and mobility disability refers to conditions in the “physical domain.” Thus, ATTRI focuses on technological solutions to remove barriers to transportation according to four functional disabilities: Visual, hearing, cognitive and mobility, defined below.
     - **Visual—**People who have a visual disability report they are blind or have difficulty seeing.
     - **Hearing—**People who have a hearing disability report they are deaf or have difficulty hearing.

---

[3] Americans with Disabilities: 2010, by Matthew W. Brault, for the United States Census, July 2012, Report Number P70-131.

- **Cognitive—**People who have a cognitive disability report one or more of the following:
  - Have a learning disability, an intellectual disability, developmental disability or Alzheimer's disease, senility, or dementia.
  - Have some other mental or emotional condition that seriously interfered with everyday activities.
- **Physical—**People who have a physical disability report one or more of the following:
  - Use a wheelchair, cane, crutches, or walker.
  - Have difficulty walking a quarter of a mile, climbing a flight of stairs, lifting something as heavy as a 10-pound bag of groceries, grasping objects, or getting in or out of bed.
  - List arthritis, rheumatism, broken bone, cancer, or other condition that limits activity or movement.

- **User needs** are based on that person's capabilities, expectations, personal schedule, etc., and are classified into the four categories:
  - **Information** for Travelers with Disabilities, the most frequently identified category of user need, is a critical component for mobility. Existing and emerging technologies in the areas of Wayfinding and Navigation and Assistive Technologies present strong opportunities to meet the information needs of travelers with disabilities.
  - Providing travel **Options** to travelers with disabilities before and during their travel enhances the trip experience and increases the probability of an uninterrupted trip. The Enhanced Human Services Transportation focus area is well-suited to facilitate enhanced traveler options through coordination between agencies, jurisdictions, and nonprofit organizations.
  - More travel **Assistance** could be given to travelers with disabilities during their travel, particularly in the forms of Assistive Technologies, Automation and Robotics and Data Integration.
  - **Access** to transportation assets could be enhanced through technology solutions, but most traveler needs related to access pertain to information about access-related amenities.
- **Barriers** to completing trips may be encountered during each segment of a trip (pre-trip planning, departure, en route, arrival, and return). The potential cause of barriers is a function of internal, external, and natural factors that fall within the realm of the transportation and transit agencies. The categories include:
  - Adverse perception of travel.
  - Cost.
  - Inadequate Infrastructure. signage or wayfinding tools.
  - Inadequate transportation options and amenities.
  - Lack of technology access.
  - Lack of travel support/customer service.
  - Driving barriers.

2. **Scenarios and Travel Activity Links (TAL)—**The second concept involves scenarios and the travel activities (steps or tasks) that are required to complete trips and that may present circumstances that limit specific user groups from traveling easily. These travel activities include

tasks such as using trip planning services or navigating through a specific travel environment. Most ADP technologies are designed to address barriers in the built environment, transportation vehicles, services or travel environments that are not designed to accommodate one or more groups. For example, an ADP technology might be designed to ease the travel planning process for people using wheelchairs since typical routing applications do not have information that is critical knowledge for wheelchair travel. Another technology may be designed to provide indoor navigation localization to orient blind or visually individuals to their location in multi-level underground transit stations.

- **Scenarios—**Scenarios describe actual or hypothetical trips being made by individuals with specific mobility profiles. As a result, those individuals have characteristics that make some travel activities challenging within current transportation environments. The specificity of these scenarios allows for a detailed analysis of the potential challenges of different groups of travelers. To help analysts create accessible technology specific evaluation, the scenario process relies on a structured view of trip-making behavior
- **Travel Activity Links—**Acknowledging that the accessibility of a complete trip depends on an individual's ability to complete every link within the travel chain, TAL provide structured, manageable pieces by which to deconstruct a travel scenario and separate the specific activities during which travelers may confront barriers. Every trip (from Origin A to Destination B) requires that travelers perform one or more of these activities. The number of TALs performed and the order in which they are performed vary from trip to trip. The 11 TALs are listed below. Appendix D provides detailed descriptions of all 11 TALs. TAL 5 is described below to illustrate how TALs are can be used to better understand the detailed actions travelers must perform.

- TAL 1: Trip planning (both pre-trip and midtrip).
- TAL 2: Accessing trip itineraries midtrip and assessing trip progress.
- TAL 3: Identifying entry/egress:
  - 3a to/from a transit vehicle.
  - 3b to/from a travel environment.
- TAL 4: Entry/egress:
  - 4a to/from a transit vehicle.
  - 4b to/from a travel environment.
- TAL 5: Pedestrian-only environments.
- TAL 6: Street crossings and intersections.
- TAL 7: Mixed environments with moving vehicles and pedestrians.
- TAL 8: Indoor and underground transit facilities.
- TAL 9: Outdoor transit facilities (e.g., transit transfer centers).
- TAL 10: Riding a vehicle.
- TAL 11: Transit payment (includes identifying payment location).

Breakdown of Travel Activity Link 5: Pedestrian-Only Environments

This TAL includes the tasks associated with travelers moving through above-ground, pedestrian-only environments. (It does not include intersections, indoor environments, or underground environments.) Specific activities include travelers doing the following:

- Gaining situational awareness (includes orienting oneself in space, identifying nearby objects and safely avoiding collisions, and assessing whether objects are moving or stationary).
- Predicting the path of moving objects and forecasting the speed at which those objects will move along that path.
- Identifying paths that are suitable for them within the pedestrian environment (e.g., can travelers actually walk on that surface, or should a different path with a better surface be taken?).
- Selecting the specific travel path they wish to take through that environment to their next travel waypoint (tactical navigation).
- Physically moving through that environment, including obstacle avoidance and mobility assistance (e.g., balance support).

- **Combining TALs to Create Trips:** Different trips require different combinations of TALs. For example, a simple walking trip from an office in downtown to a restaurant across the street, made by an individual every day, would include only three TALs, although some of the TALs would occur more than once during the trip. The TALs would occur in the following order: Link 4b Entry/egress, Link 5 Pedestrian-only environments, Link 6 Street crossings and intersections, Link 5 again, and finally Link 4b for a second time. These links are used to identify specific issues or user needs associated with a specific population's difficulties, such as 1) exiting the building and entering into the sidewalk environment; 2) navigating to the intersection; 3) getting across the street; 4) navigating to the proper door of the restaurant; and finally 5) navigating into the restaurant itself.

  An example of a more complex trip might be if that same individual needed to take a transit bus across town to meet a client at a restaurant. That trip might include Link 1, Link 4b, Link 5, Link 6, Link 5, Link 3a, Link 4a, Link 11, Link 10, Link 3a, Link 4a, Link 5, Link 3b, and Link 4b. In this case, the traveler would need to plan her trip—determining which bus to take, where to catch that bus, how to strategically navigate to that bus stop, and when she needed to arrive at that stop in order to catch that bus. Then she would leave her location, travel to the stop, identify that the correct bus was arriving and where the door to the bus was located. Next, she would need to board the bus, pay for the transit trip, and find her seat on the bus. Next, the traveler would need to identify when it was time to disembark, signal the driver, and successfully exit the bus onto the sidewalk when the bus stopped. Finally, she would need to navigate to the restaurant, identify how to enter the restaurant, and complete the trip.

- **Converting Trips to Scenarios:** After trips have been created, scenarios can be created by defining the user characteristics of the individuals traveling. Different users (populations of individuals) experience a given trip (set of TALs) differently as a result of their different abilities. TALs that are easy to perform for some populations can be very difficult for others to perform. Linked together, they allow the IE to explore the combined impacts of a travel population's disabilities on their travel mobility, and the effects that a new ADP technology will have on that overall mobility.

  A variety of scenarios—both different types of trips and potentially different user populations—are needed for each ADP evaluation, and the trips and scenarios must be specifically crafted to examine the performance of the ADP and its use by the user population it is intended to help, given the types of trips that population takes.

  The scenarios should be developed on the basis of the travel barriers associated with the intended users of the ADP technology or service, and they should reflect the types of trips those individuals need to make in their daily lives. At least some of the scenarios should include both the need to travel serendipitously and the need to navigate on-the-fly strategically (making origin-to-destination trip decisions) and tactically (making decisions about issues encountered en route). (For example, the scenarios should include the ability of the traveler to respond to unexpected travel conditions to change route midtrip or to add a stop midtrip.) These scenarios are key to both evaluating how the ADP affects the activities of individuals in traveling and to understanding how the ADP technology or policy affects the user population's overall mobility (i.e., the effect on the "complete trip").

3. **Evaluation Contexts—**The next set of concepts discussed in this chapter are the "evaluation contexts." This set of six topic areas helps the IE team understand the full breadth of attributes that contribute to the success or failure of an accessible transportation technology and therefore need to be considered for inclusion in a project evaluation. By explicitly describing specific contexts that need to be considered in the evaluation, the framework helps ensure that the IE team considers all the important attributes of the ADP technology as it develops the scope of the evaluation. The evaluation framework divides evaluation subjects into six specific contexts. These six contexts are defined as follows:

   1. Technical Function: This context evaluates whether the product functions according to design specifications. It also explores whether the product improves user's travel efficiency (e.g., decreases travel time) and increases the efficiency of particular TALs.
   2. Technological Robustness: This context asks whether the technology is high quality, reliable, safe, and durable through user testing. This context provides structure to analyze safety measures, including the protection and privacy of user information.
   3. Usability: This context examines the customization available to meet the needs of particular traveler subpopulations.
   4. Communication and Closing Information Gaps: This context examines whether the ADP technology can effectively communicate with the user population. Successful communication includes both allowing the user to request specific information when and where it is needed and to receive, perceive, and comprehend a response to those requests or receive, perceive, and comprehend other necessary information.
   5. System and Service Integration: This context is concerned with the steps in the travel chain, which agencies are potentially affected by the ADP technology, and how agency stakeholders are affected by the technology (other riders, operators, management, etc.).
   6. User Empowerment and Social Acceptance: This context examines whether the ATTRI ADP technology increases user empowerment (the ability of the individual to control their own life), and whether the ADP technology can be used in public without drawing unwanted attention to the user. That is, the user is able to travel more freely, and is comfortable using the ADP technology in public.

4. **Performance Measurement—**An evaluation of an accessible transportation technology can be focused on many different aspects, such as the interface, the technical performance, or the travel outcomes resulting from the technology's deployment. Most ATTRI ADP evaluations are expected to focus on four basic types of outcomes listed below:
   1. The functional performance of the technology viz-a-viz the defined primary target population.
   2. The ability of the target population to successfully interact with (use) the technology (including successful user interaction and task completion).
   3. The impact the technology has on the ability of the target population to perform specific travel activities.
   4. The degree to which the technology facilitates greater mobility and travel opportunities for the target population (i.e., whether they gain the ability to perform more complete trips, or whether experiences in a particular travel link are improved.)
5. **Threat Model—**Defined broadly, a "Threat Model" is an understanding of the things that could go wrong with the operation and use of the ADP technology, as well as an understanding of the ways the technology is designed to address those failure points if and when they occur. It covers topics such as the following:
   - Mechanisms that could cause the technology to fail and that are critical for understanding the overall robustness of the ADP technology's performance (e.g., what happens if a loss of communications occurs, or if a user encounters an unexpected travel outcome and requires assistance?).

- Safety concerns that could result from use of the technology (e.g., are there ways in which use of the ADP technology could put users in harm's way, such as texting while driving with a smartphone, or leading an individual into an environment that s/he cannot safely traverse?).
- Indirect threats to the technology user, such as invasion of user privacy or cyber security concerns associated with use of the technology (e.g., how does the technology prevent a stranger from obtaining a secret access code if it is part of wayfinding instructions for a user?).

6. **Logic Models—**Logic models present the traceable connections across a project's goals to intended outcomes and impact. The logic models formalize the traceable mapping from evaluation contexts, to hypotheses that could be evaluated and the details by which those hypotheses are tested given the goals and objective of the technology, including performance metrics and data requirements.

   For the Evaluation Framework, three logic models mapped to performance outcomes are:

   - Performance of Travel Activities—Examining the **effectiveness** of the ADP in achieving those desired outcomes and the **efficiency,** which measures an ADP's performance by comparing how many resources the user spends in order to complete their trip before and after the introduction of the ADP.
   - Ability to Mitigate Threats.
   - Ability to Address Target Populations Needs.

7. **Gap Analysis—**Once the basic analysis described in the Logic Model is completed, if resources allow, a "complete trip gap analysis" may be performed. The complete trip gap analysis is designed to examine in a qualitative manner, the expected impacts of the ADP on the variety of trips made by the target population and assess the project's contribution within the context of the larger transportation network.
8. **Complete Trip Analysis—**In contrast to the gap analysis, the complete trip analysis is intended to give the sponsoring agency insight into the larger travel outcomes that the ADP can provide, as well as insight into the remaining issues that may still prevent the target population from traveling as easily and freely as desired.

Chapter 3. Evaluation Framework

In defining a methodology for performing a useful independent evaluation, the independent evaluations (IE) concentrates its analysis on 1) the population experiencing the travel gap which the Accessibility Development Projects (ADP) seeks to address; 2) the specific travel deficits or gaps that the technology aims to improve or close; 3) the travel experience it seeks to improve; and 4) the travel (and life) outcomes that result from the introducing the ADP intervention to the target population.

This framework is focused on evaluating the performance of tools or technologies to meet identified travel needs or barriers. Identifying and categorizing the needs of users and detailing the nature of these barriers were the subjects of previous Accessible Transportation Technologies Research Initiative (ATTRI) efforts and are not addressed in this framework other than applying the results of those prior ATTRI projects.

The ultimate goal of a developed independent evaluation is to measure the impact of introducing the ADP technology to on travel and participation in travel in the target population. In addition, the evaluation results will embrace users' perspective in order to provide useful recommendations for the ADP team on making its tool/technology more usable and useful for its target users. The combination of these results describe the level of success achieved by the ADP and provide the information needed to further improve the target users' access to transportation and mobility.

Conducting an independent evaluation for an ADP requires the following:

- Understanding travel use and perspectives of the technology's desired, likely, and/or actual (or "target") users.
- Understanding the target users' needs in the context of the complete trip.
  - The problems, risks, and threats users are likely to face.
  - The problems, risks, and threats the technology seeks to mitigate.
  - The problems, risks, and threats the technology cannot mitigate but that users might expect it to.
- Measuring the performance, impact and user satisfaction with the technology in the field.
- Analyzing the technology's use heuristically (i.e., through an expert review).
- Analyzing users' experience of the technology empirically (e.g., through user studies both in the lab and in the field).
- Determining the overall change, perceived change, or likely changes in travel behavior and perception in the target population.
- In response to review and evaluation results, making actionable recommendations to the ADP team, as appropriate, to help them improve the technology's design.

While the framework is intended to measure the changes in travel that result from deployment of the ADP, it is also important to obtain, through the evaluation, an understanding of how well the components of the ADP function, especially how easily and effectively users can interact with it and how those interactions affect the perceptions travelers have of their ability to complete trips when and where they need to make them.

Accessible transportation technologies should be viewed not as a product, but as a process. Successful execution of this evaluation framework will not be transactional but cooperative. This means that the ADP team will not simply receive a report of problems and successes from the IE team as a result of use of the evaluation framework, but that an iterative exchange will lead to the definition of, agreement about, and execution of an independent evaluation process. The IE will then lead to actionable insights and improvements over time for the developers of the ADP, the agencies that need to help implement and support the technology, and the agencies and organizations that will fund these tasks.

The evaluation framework provides project sponsors and IE teams with the ability to focus their evaluation efforts on four different aspects of ATTRI project evaluation and measurement:

1. The performance of the new technology as it will be used by the population (target users) for which it is intended (e.g., Does the technology function as intended? Can users effectively interact with it? Does it successfully fill information gaps?).
2. Improvements that the ADP achieves in meeting ATTRI's adopted user needs or surmounting identified travel barriers.
3. Changes in targeted users' behaviors and ability to perform the travel activities that the ADP was designed to address.
4. Changes in targeted users' ability to travel more freely and spontaneously, that is, in their ability to make more "complete trips."

The first three aspects of the evaluation will be carried out by examining how effectively members of the target population can use the ADP and how effectively they can perform the specific travel activities the ADP is intended to improve. The last aspect of the evaluation will be performed through a gap analysis. Complete trips are defined in terms of an individual's ability to plan for and complete a trip from origin to destination without disruptions (gaps) in the travel chain. The framework recommends examining the targeted users' ability to perform complete trips by creating representative travel scenarios for those users, evaluating how incorporation of the ADP technology helps them perform the travel activities necessary to complete those travel scenarios, determining what gaps in making complete trips the ADP will close and which will remain, and critically considering what new barriers may arise through the introduction of the new technology.

The recommended process for setting up and performing an evaluation of an ADP includes the following steps:

1. Set up the IE by identifying the key details required for the evaluation:
    - Review the goals and expected performance outcomes of the ADP.
    - Characterize the target population for the ADP.
    - Document the travel requirements, perspectives and needs of that target population.
    - Understand the ADP technology, including:
        - The specific travel outcomes it is intended to achieve.
        - The technological steps it will perform.

        - The user interactions with the ADP technology necessary to achieve those outcomes.
    - Determine the key objectives for the evaluation effort. (i.e., the findings that are most important to determine.)
    - Establish which specific travel activities will be affected/improved by the ADP.
2. Develop a threat model for the ADP functions and required user interactions for each of the affected travel activities.
3. Develop an initial logic model for the evaluation project that serves as a guide to the evaluation activities, including the required data collection, and estimate a budget for the activities described in that logic model. Then work cooperatively with the project sponsors and ADP team to refine the logic model, data collection plan, and resulting evaluation plan to match the scope of work with the available budget, given the key objectives determined above.
4. Perform the evaluation.
5. Perform a qualitative gap analysis addressing the complete trip evaluation and continuing needs assessment based on the outcomes of the evaluation activities and the target population's user needs.

Each of these activities is described in detail in the remainder of this chapter. Note that while the list above provides a useful order in which to approach these topics, the need to work interactively with the full range of stakeholders may result in the IE team approaching these tasks in slightly different order as the evaluation proceeds. Also note that setting up the evaluation involves some iteration, as the initially intended evaluation tasks may require data that is unavailable or require tests which cannot be collected within the available budget. Thus, the IE team needs to interact closely with the ADP team while setting up the evaluation, in order to determine what data can be collected, what tests can be performed, and how those activities affect the design and performance of the evaluation.

Step 1: Set Up an Independent Evaluation

To set up the evaluation, the team must first gain a thorough understanding of the technology being developed, the tasks it is intended to perform, and the population expected to use it. The outcome of this review should be a basic understanding of the following:

- The target population(s) for the ADP.
- The travel needs intended to be more effectively met or travel barriers meant to be overcome.
- The basic technical actions to be taken by the ADP (e.g., data to be gathered, information to be delivered, and the technical tasks to be performed to gather those data, convert them into information, and deliver that information).
- The overall outcomes expected from the project, both in terms of travel outcomes (the travel activities expected to change and how) and system implementation (e.g., Is this a proof of concept? A model deployment test? A full-scale deployment?).

These insights can be gained from the proposal that led to the ADP's funding; the final scope of work associated with the project; and discussions with the ADP team, the project sponsors, and the agency or organization funding the evaluation effort. The outcome of the set-up task is a complete understanding of who the ADP is intended to help, how that help will be delivered, and the goals and expected outcomes of the ADP.

To present the evaluation approach within this framework document, we will use an example of an ADP to guide the review of evaluation steps. The example is shown in highlighted call out boxes that accompany the framework guidance.

Working example—Robotics and Automation Technology at an Outdoor Transit Center (Robotics and Automation)

This example evaluation is for a wheeled robot designed to assist blind individuals who need to navigate above-ground transit centers. The robot is designed to assist travelers who have a variety of mobility needs in addition to those with low vision. The robot is called to individuals arriving at the transit center with their smartphone application. Once the robot has arrived, it announces its arrival via voice and vibration through the smartphone application. The smartphone must then be used to acknowledge that the robot has found the correct user via a tap on the phone screen or voice command. At that point, users can either give voice commands to the robot or send text commands to the robot. The ADP is designed to work with smartphones connected to braille data entry and output devices. Once the traveler and the robot have connected, the robot then guides the traveler to their desired destination within the transit center.

Blind individuals may use the robot in one of two ways: 1) they may place a hand on the robot, which allows the robot to lead them; or 2) the robot can physically lead users to the next transit stop while giving them auditory, turn-by-turn directions to follow. The robot follows the path with the best pavement surfaces (fewest tripping opportunities caused by disruptions in the surface profile) to the desired destination within the transit station. Because the robot is wheeled, it does not take stairs or escalators, even if they provide a more direct path. Once users have arrived at the desired vehicle boarding location, the robot waits with them, identifies the vehicle to board when it arrives, takes them to the correct boarding location, and provides voice guidance for boarding that vehicle.

The robot does not currently supply expected real-time arrival information about transit vehicles. It does use short-range object detection to identify local objects in its path so that it can avoid obstructions or potential collisions (e.g., people, suitcases, and other objects that may or may not be part of the fixed infrastructure). The robot also does not physically assist users in boarding or exiting the transit vehicle. (Although the robot does not physically lift an individual, it does attempt to align users with the door opening.)

Step 1a: Understand the Target Population

An appropriate starting point for the IE set-up is to gain a more complete understanding of the population that the ADP is targeting, including the need(s) of that population and the barrier(s) it needs help overcoming. These needs and barriers help in defining the specific travel problem or set of travel issues within the Travel Activity Links (TAL) that the ADP is intended to mitigate. The target users of the ADP technology may include more than one population, with multiple different user needs and abilities. Variations in user needs and capabilities among target user populations are important to recognize because those differences typically require that the evaluation does the following:

- Test the use of the technology among each target population.
- Refine those tests to examine the impacts of specific population needs and capabilities on the population's ability to successfully interact with and use the ADP technology.
- Understand both the improvements in mobility (the complete trip) that result from the ADP for different population groups and understand remaining limitations with the ADP technology.

One aspect of the IE that the ATTRI program is very interested in evaluating is the impact of the ADP on the “complete trip” (e.g., the overall ability of the target population to travel conveniently and flexibly). Therefore, the IE needs to determine not just the effectiveness of the ADP at improving a specific travel activity or activities but the impact of those improvements on the overall ability of the target population to travel. To answer this larger question requires that the IE team understand the target population’s other travel needs/barriers, as those remaining needs/barriers may limit the overall mobility benefits that the target population experiences because of the ADP deployment. The qualitative complete trip (gap) analysis performed at the end of the IE focuses on these broader travel impacts.

Both the initial set-up of the IE and the performance of the gap analysis require an understanding of the full range of the target population’s mobility needs and barriers. The starting point to understanding the target population to a review the first key concept—user needs and barriers as presented in chapter 2 and the ATTRI User Needs Assessment: Stakeholder Engagement Report. Many people may experience more than one of the disabilities described. Therefore, a key task for the IE team is to understand not just the specific TAL (or TALs) addressed by the ADP, but the overall travel needs of the population being targeted, and thus the complete set of issues and concerns that affect that population’s travel ability. Table 22 in appendix A presents a list of the types of travel disabilities along with a robust list of the types concerns these disabilities raise that ADPs are designed to address. Table 1 provides a simplified version of this table.

Understanding the travel activity concerns described in table 1 and table 22 helps the IE team identify the full set of user needs the ADP is designed to meet and the barriers the ADP is designed to surmount.

Table 1. Disabilities affecting travel mobility.

Disability Category	Description of Issue/ Concern	Travel Activity Concerns
Motor	Requires Assistive Device: • Cane, crutches. • Manual or power wheelchair.	• Inability or limited ability to use stairs. • Limited ability to carry devices. • Limited distance that can be traveled. • Need to be locked down on transit vehicles and may require driver assistance to perform this task.
Vision	• Blind. • Partial vision.	• Need for screen readers or other technology. • Inability to see approaching vehicles, creating safety risk. • Difficulty navigating.
Hearing	• Deaf. • Partial hearing acuity.	• Inability to hear approaching vehicles, creating safety risk. • Limited information transfer from auditory clues.
Learning/ Cognition	• Dementia. • Processing disability. • Intellectual disability. • Mental or emotional disability. • Autism.	• Limited short-term memory, making it difficult or impossible to remember directions. • Easily confused by changes in environment. • Difficulty reading signage. • Social anxiety. • Limited problem-solving capacity.
Speech	• Communication disorder.	• Difficulty being understood when asking for assistance (vehicle boarding or attempting to identify exit stop).

Source: Federal Highway Administration.

Example Robotics and Automation—Understand the Target Population

While the focus of this evaluation is on users with low or no vision, the ADP is designed to be used by individuals with a wide variety of mobility challenges, including those that use a variety of mobility devices such as manual wheelchairs, powered chairs, and knee scooters. Users may have multiple disabilities.

To limit the size of the example, the example evaluation focuses on a population of users whose primary disability is visual. However, even with a concentration on users with low vision, subpopulations of this group can be expected to have a variety of other mobility challenges. Many of these users will also use another assistive device, such as a service dog, a cane, or a wheelchair. The need to use these devices may change the paths these individuals can follow. The personalized navigation directions provided by the ADP technology need to enable travelers to overcome all barriers that are presented along the way. However, because the robot is wheeled, only paths that can be traversed by the robot will be used, even if users could traverse a shorter path.

A specific evaluation issue with this target population is the ability for travelers to successfully communicate with the ADP technology. The ADP technology can be used by individuals with good vision, but the sponsors for this effort are focused on those with limited vision. Therefore, for this population, all communication between travelers and the ADP technology must be auditory, although the smartphone application also allows both tactile cues (vibrations) and braille outputs (for smartphones that have braille capabilities). Blind individuals may use the robot in one of two ways: 1) they may place a hand on the robot, which allows the robot to lead them; or 2) the robot can physically lead users to the next transit stop while giving them auditory, turn-by-turn directions to follow. The robot follows the path with the best pavement surfaces (fewest tripping opportunities caused by disruptions in the surface profile) to the desired destination within the transit station. Because the robot is wheeled, it does not take stairs or escalators, even if they provide a more direct path. Once users have arrived at the desired vehicle boarding location, the robot waits with them, identifies the vehicle to board when it arrives, takes them to the correct boarding location, and provides voice guidance for boarding that vehicle.

To help the IE team document these intended outcomes, table 2 presents a simple form that can be used to highlight which user needs are being addressed. The form was developed by combining material from the ATTRI User Needs Assessment: Stakeholder Engagement Report with an understanding of how these needs are distributed across different disability types.

This form can be used in concert with an understanding of the populations whose needs are being addressed and the travel activities for which the ADP is being deployed to help the IE team document the intended outcomes of the ADP and prioritize those that are most important for review within the evaluation effort. The second column of the table lists specific User Needs that may or may not be addressed by the ADP. The first column on the left indicates the disability types that might to which this need is applicable. The third column of the form allows the IE team to indicate whether each need is applicable to this ADP, and the last column indicates the relative importance (Low / Medium / High) of that User Need to the particular evaluation being performed. Table 2 has been filled out in support of the Robotics example. A blank copy of the form can be found in appendix C.

Table 2. User needs checklist for the robotics and automation example.

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable?	Significance to ADP Objectives (L, M, H)
All	1. Does ADP provide information in a variety of accessible formats?	Y	M
All	2. Is information from ADP interface accessible in a variety of environments (i.e., amid heavy crowds and noise, underground)?	Y	H
All	3. Does ADP perform a task that improves safety and security or that provides emergency information?	N	
All	4. Does ADP provide en route assistance and information?	Y	M
All	5. Does ADP provide connection information (where, who, when)?	Y	M
All	6. Does ADP provide estimated trip length and distance?	N	
All	7. Does ADP provide comprehensive travel information?	N	
All	8. Does ADP require access to equipment (phones, computers, charging, training)?	Y	L
All	9. Does ADP allow the user to create a personalized profile?	Y	H
All	10. Does ADP require coordination of information (between agencies, modes)?	N	
Blind and Visually Impaired (BVI) (Visual) Motor Impairment (MI) (Motor)	11. Does ADP provide real-time transportation information?	Y	M (BVI) M (MI)
BVI (Visual)	12. Does ADP provide audible mapping/directions?	Y	H
BVI (Visual)	13. Does ADP provide destination information (hours, addresses, entrances, layout)?	Y	M
BVI (Visual) MI (Motor	14. Does ADP provide transit schedule and other transit information (e.g., stop location)?	Y	M
BVI (Visual) MI (Motor)	15. Does ADP provide information about pathway infrastructure?	Y	H (BVI) H (MI)
BVI (Visual) MI (Motor	16. Does ADP include provision for outside assistance or attendants?	N	

Table 2. User needs checklist for the robotics and automation example (continuation).

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable?	Significance to ADP Objectives (L, M, H)
BVI (Visual) MI (Motor	17. Does ADP provide amenity information (e.g., restroom, shelter, benches, food, drinks)?	Y	L (BVI) L (MI)
BVI (Visual) Cognitive	18. Does ADP provide information about, and interpretation of, signage?	N	
MI (Motor)	19. Does ADP provide transportation facility information (e.g., maps)?	Y	M
MI (Motor)	20. Does ADP provide information about weather conditions?	N	
Hearing Cognitive	21. Does ADP include information about and/or interpretation of announcements?	N	
Hearing	22. Does ADP incorporate speech-to-text or text-to-speech that enables the user to communicate more easily?	N	
Cognitive	23. Does the ADP provide information in a concise and straightforward manner?	Y	L

Source: Federal Highway Administration.

Any given ADP technology may address multiple user needs. However, not all of the needs will be of equal importance in the evaluation. Part of the evaluation development process is to determine which of these needs, or what set of these needs, should be the focus of the evaluation. Table 2 is designed to start that prioritization by indicating which user needs that are applicable to the ADP are of most significance to the project. Making that determination is best accomplished with an understanding of how the ADP is designed to improve the target population’s performance of specific travel activities.

Example Robotics and Automation—Understand the Target Population—Elaborating User Needs

This example provides the details behind those User Needs in table 2 that were indicated "Yes." To save space, it does not discuss the "No" answers.

1. **Does ADP provide information in a variety of accessible formats?** The ADP's smartphone application is designed to allow users to obtain information both aurally and visually. The focus of this test will be on the auditory communication.
2. **Is information from ADP interface accessible in a variety of environments (i.e., amid heavy crowds and noise, underground)?** The ADP's smartphone application and the robot itself are designed to allow users to obtain information both aurally and via a braille reader attachment to the smartphone. The robot is also programmed to repeat spoken directions if users are unclear request the instructions to be repeated.
4. **Does ADP provide en route assistance and information?** The ADP provides information on where the transit stop for the departing bus is located, when it is scheduled to leave, and the directions for navigating to that stop.
5. **Does ADP provide connection information (where, who, when)?** The robot is able to identify the bus/route/stop that is the destination of the user, can successfully travel to that location, and can announce the arrival of transit vehicles at that location.
8. **Does ADP require access to equipment (phones, computers, charging, training)?** The ADP technology requires smartphone technology. While the robot can use an external speaker to give directions and it can respond to voice instructions, the robot must first be "paired" to a user through the smartphone app, so that it understands which directions to follow. Therefore, users must have access to a smartphone with the ADP software loaded on it.
9. **Does ADP allow the user to create a personalized profile?** The ADP's smartphone application has features that allow users to set their preferred method of communication (e.g., audio versus visual, etc.), as well as define their form of locomotion or need for assistive devices (e.g., manual wheelchair, powered wheelchair, cane, guide dog, etc.). These are then used to both determine the types of surfaces that must be used or avoided in selecting the traveler's path and how the robot should communicate with each user.
11. **Does ADP provide real-time transportation information, including 1) real-time vehicle status; or 2) real-time travel condition/obstruction information?** If the agency has a General Transit Feed Specification (GTFS)-Real Time data feed, the ADP technology includes real-time transit arrival information. The robot also includes real-time object detection and avoidance capabilities. This latter capability allows the robot to detect and avoid collisions with moving objects and unexpected obstacles (e.g., luggage in the path).
12. **Does ADP provide audible mapping/directions?** The ADP technology is designed to compute navigation paths and to guide different users within the context of the transit center. It provides those directions via spoken words when that option is selected in the personal profile.
13. **Does ADP provide destination information (hours, addresses, entrances, layout)?** The robot can navigate to all destinations within the transit center, and that it can apply multiple names to those destinations (e.g., the "northbound Route 47 stop" versus "stop number 56102" or "the northern exit to the transit center" versus "to North 53rd street" or "I'm going to an office building at 1107 North 53rd St."

Example Robotics and Automation—Understand the Target Population—Elaborating User Needs (continuation)

This example provides the details behind those User Needs in table 2 that were indicated “Yes.” To save space, it does not discuss the “No” answers.

14. **Does ADP provide transit schedule and other transit information (e.g., stop location)?** The infrastructure database and path finding algorithm used by the robot include transit stop location information (both stop locations and the route numbers and directions serving each stop). If the transit agency has a GTFS real-time feed, the robot has the ability to pass along expected arrival times from that feed to users.)
15. **Does ADP provide information about pathway infrastructure?** The robot must have a very detailed infrastructure pathway map within the transit center for it to successfully lead travelers to their desired destination within the transit station.
17. **Does ADP provide amenity information (e.g., restroom, shelter, benches, food, drinks)?** If there are amenities at the transit center, the robot can identify the locations of those amenities and lead the traveler to them.
19. **Does ADP provide transportation facility information (e.g., maps)?** The robot’s map database includes all transit stop locations within the station area, as well as the attributes (e.g., routes and directions served) associated with those stops. It also knows where all ramps are located for moving between locations at the transit center.
23. **Does the ADP provide information in a concise and straightforward manner?** The robot is designed to provide audio directions in a very simple manner, in order to make its directions very easy to follow.

Step 1b: Understand the Travel Activities

When identifying user needs and barriers, it is best to keep in mind the travel activities the ADP is designed to improve for the target population. The 11 TALs, presented in chapter 2, are:

- TAL 1: Trip planning (both pre-trip and midtrip).
- TAL 2: Accessing trip itineraries midtrip and assessing trip progress.
- TAL 3: Identifying entry/egress:
    - 3a to/from a transit vehicle.
    - 3b to/from a travel environment.
- TAL 4: Entry/egress:
    - 4a to/from a transit vehicle.
    - 4b to/from a travel environment.
- TAL 5: Pedestrian-only environments.
- TAL 6: Street crossings and intersections.
- TAL 7: Mixed environments with moving vehicles and pedestrians.
- TAL 8: Indoor and underground transit facilities.

- TAL 9: Outdoor transit facilities (e.g., transit transfer centers).
- TAL 10: Riding a vehicle.
- TAL 11: Transit payment (includes identifying payment location).

Different trips require different combinations of these TALs. When the evaluation plan is set up, TALs are used to identify specific issues or user needs associated with a specific population's travel needs and barriers, To develop the IE plan, mutual understanding between the IE and ADP teams is needed about the following topics:

- The travel activities that the ADP technology is intended to improve, and if more than one, which TALs are the highest priority in understanding the performance of the ADP technology.
- The reasons that the target population experiences difficulties performing those activities (see table 5 later in this report.)
- The abilities of the target population.
- How the ADP intends to assist the target population in more easily performing the travel activities.

This information allows the IE team to develop an evaluation plan that examines the following:

- Whether the target population can effectively interact with the ADP technology while performing each TAL.
- Whether the technology delivers the required information needed to improve users' performance of each TAL.
- Whether the target population is comfortable using the ADP when undertaking each TAL.
- Whether the target population actually performs each TAL more effectively when using the ADP.

It is important to note that different hypotheses and tests may be needed to examine the use of the ADP technology by different targeted subpopulations performing a particular TAL, as well as for different TALs performed by one target population. For example, an ADP might be specifically designed to help people with low or no vision identify the best entry/exit point to buildings. Such an ADP would meet a need and help remove a barrier to travel. However, an ADP that served a blind or low vision population might not be useful to individuals with those disabilities who also used a wheelchair. Therefore, if the ADP also was intended to serve users of wheelchairs, additional evaluation activities would be needed to test the effectiveness of the ADP on these two very different target populations.

Example Robotics and Automation—Travel Activity Links

The robot in this example does not provide physical assistance to travelers as they board or alight from the transit vehicle, therefore the logic model will focus on TAL 9 (outdoor transit facilities).

This ADP starts with the need for communication between the robot and travelers. Communication is required to provide the robot with travelers' arrival time and location and their desired destination within the transit center. Communication is also required to physically connect the robot to users when they "meet" at the arrival transit stop (or transit center entry point), and for the robot to transmit the necessary navigation instructions to users. The robot is designed to work with several different communication techniques.

In the version of the ADP being tested, the robot is called to individuals arriving at the transit center with a smartphone application. Using a smartphone and the ADP smartphone application, travelers can either speak into their phone's microphone or enter via text that they will be arriving soon at the transit center on a specific route traveling in a specific direction. Travelers then indicate the route and direction to which they wish to transfer. (For individuals with some vision, these options can be selected via a pick list on the smartphone screen.) Travelers also have the option of indicating the intersection next to the transit center and when they will be arriving, the street and direction in which they wish to exit the station, or the amenity located at the transit center to which they wish to travel.

The ADP transmits this information to the robot, which then computes the origin and destination points for the trip within the transit center for each traveler and trip. Based on the traveler's mobility characteristics, the robot then identifies the path it will follow to help each traveler move through the transit center from the origin point to the center (transit stop or entry point to the transit center) to the exit point from the center (transit stop or exit point), as well as the path the robot must take from its current location to the location where it will meet the traveler. Once a request for assistance has been made, the traveler's smartphone also starts to broadcast its global positioning system (GPS) location. This helps predict when the traveler will arrive and confirms the traveler's location within the transit center once they have arrived.

The robot then travels to the expected arrival point at the station and waits for the traveler.

Once both the robot and the user have arrived at that location, the robot announces its arrival via voice and vibration through the smartphone application. The smartphone must then be used to acknowledge that the robot has found the correct user via a tap on the phone screen or voice command. Once the robot has connected with the traveler, the user can either give voice commands to the robot or send text commands to the robot.

To guide users through the transit center, several communication options are available. An individual with sufficient vision can simply follow the robot, communicating by voice or text to slow the robot if it moves too quickly. A blind individual may use the robot in one of two ways: 1) by placing a hand on the robot, which allows the robot to lead; or 2) by following the robot to the next transit stop while it gives audio, turn-by-turn directions. Audio directions can be provided either via the smartphone app and smartphone speakers (i.e., earpieces), spoken out loud by the robot, or both. Travelers may select between these options via their personal profile.

Some ADP users with low or no vision also need to use mobility devices, including manual wheelchairs, powered chairs, knee scooters, and service animals. Thus, an important subset of the target population does not have a free hand to manipulate a smartphone when they are traveling. Consequently, the evaluation needs to specifically test the "hands free" functionality of the ADP.

Step 1c: Understand Stakeholder Objectives

The last major subtask within the project set-up is to understand the evaluation project's priorities. An "evaluation" of an ADP can be focused on many different aspects of the ADP. What the evaluation actually focuses on will be a function of who is paying for the evaluation, and what the project itself is intended to accomplish. This step examines what most evaluations examines. The IE team needs to work with their project stakeholders (i.e., the project funding agency, the ADP team, other groups and agencies involved in the project) to prioritize the outcomes the IE team evaluates.

The vast majority of ATTRI funded ADP evaluations are expected to focus on four basic types of outcomes:

1. The functional performance of the technology, including the ability of the target population to successfully interact with (use) the technology.
2. The impact the technology has on the ability of the target population to perform specific travel activities.
3. The ability of the technology to mitigate threats to the safety and wellbeing of the target population.
4. The degree to which the technology facilitates greater mobility and travel opportunities for the target population (i.e., whether they gain the ability to perform more complete trips).

Most evaluations will include at least some aspects of all four of these evaluation categories, as failure in any one of these areas will result in mobility outcomes that do not meet the goals of the ATTRI program. In addition, they are interrelated, as failure in the first category will be a direct cause of failure in the second and third categories, which in turn may be one of several causes for failure in the fourth category. Unfortunately, studying all four of these outcomes in detail can be expensive. Therefore, many IEs will focus on a limited subset of these four evaluation topic areas.

The relative importance of these four IE outcomes will change depending on the audience for the evaluation. The first category is focused on the development and use of the technology itself. They produce the most relevant evaluation outcomes for stakeholders interested in building, testing, and refining the performance of the ATTRI technologies.

The results in this category of tests are used by the development community to understand whether the base technology is working as intended. Are data collected, transformed, and transmitted accurately and reliably? If the technology does not work reliably, it will never be used.

Human-centered design, which is of key importance for all ATTRI technology projects is an important aspect of this category. The technology may work, but its interface may not be effective for specific populations, or its physical specifications may create issues for the target user population. The lack of an effective human-centered interface is a core problem for many travelers with disabilities. Their abilities do not match well with the features of many current technology solutions. If target users find the technology unreliable or difficult to use or are socially uncomfortable using it in public, then they will not use it frequently, and it will not have the travel impact desired. This evaluation category also typically provides the answer to “why” the outcomes from the other three evaluation categories are occurring. That is, a poor interface or unreliable technology performance explain why the target population does not measurably improve their performance of key TALs or travel more freely, safely, and opportunistically.

The second category (improving performance of travel activities) focuses the evaluation on the travel outcomes—the transportation problem (need or barrier)—that the ADP technology has been designed to address. This evaluation focus concentrates on the degree to which the ADP technology achieves desired travel improvements among the target population. While it may include examination of why the technology performs as well (or as poorly) as observed, the focus of the evaluation effort is not on the details of the technology's functional performance but on the target population's ability to perform specific travel activities. This evaluation focus is the core of determining whether an ATTRI project actually meets identified user needs and/or measurably helps users surmount travel barriers.

Example Robotics and Automation—Understanding Stakeholder Objectives (Technical Functionality, Usability, Travel Outcomes, Safety, and Empowerment)

The independent evaluation might focus on the **Technical Functionality, Reliability, and Usability** of the system, across the wide range of users. The focus of such an evaluation could test topics such as:

- Can all types of users easily call and communicate their destination to the robot?
- Does the robot correctly meet and identify users? If not, why not?
- Does the robot correctly compute the appropriate path for that user?
- Can all types of users identify and follow the robot's navigation instructions?
- Can all users communicate with the robot in all crowd conditions?

Alternatively, the evaluation focus could be on **travel activity outcomes,** including **User Empowerment**. Such an evaluation would examine topics such as:

- The number and percentage of successful transfers made.
- The user population's comfort with their ability to travel, and their level of satisfaction when using the technology in public.
- The interest of the test population has in continuing to use the technology.
- The expected degree to which travel behavior will change as a result of the deployment of the technology.

The third potential evaluation focus area examines **safety,** and specifically **the ability of the technology to mitigate threats.** Tests in this area would include:

- Can users safely follow the robot through the facility to their destination?
- How often do users not understand or follow the robot's navigation directions?
- Whether the robot, and the users following the robot, cause safety concerns for other users of the transit center, or the transit vehicles operating within the center.

The final potential focus of the evaluation would emphasize **Complete Trips and Gaps,** examining:

- The degree to which easier transfers increase the target population's overall mobility.
- The ability of the target population to obtain the smartphone technology needed to use the system.
- The need for other improvements (e.g., better transit facility infrastructure databases) that are required to allow deployment of the ADP.

The third category (reducing threats to safety) of evaluation focuses on the safety outcomes of the ADP. It focuses on understanding the things that could go wrong with the operation and use of the ADP technology, and the ways the technology is designed to identify and respond to those failure points if and when they occur. The evaluation outcome is a determination of whether the ADP can be safely deployed and used by a vulnerable population, or whether specific aspects of the ADP place users at risk, and need to be addressed before deployment can continue.

The last evaluation category examines the overall effect that the ADP has on actual trip-making behavior for the target population. The evaluation concept here is the degree to which solving the specific transportation problem targeted by the ADP actually improves the mobility of the subject population, and to what degree that target population sees a significant improvement in their ability to travel and quality of life, or whether other needs or barriers continue to limit their mobility.

While these outcomes are intertwined, understanding which focus is most important to the stakeholders will allow the IE plan to allocate its limited resources where they will provide the information that is most important to that sponsor.

Outcome of the Project Set-Up

The outcome from the above activities should be a written summary of the following:

- The prioritized goals of the ADP.
- A description of the target population(s) and which populations will be the focus of the IE.
- A short description of the TALs to be affected by the ADP and the travel outcomes the ADP is intended to achieve for each of those TALs.
- The prioritized goals of the evaluation.

The project stakeholders should review and comment on this summary document. It should be refined on the basis of those comments to ensure that the stakeholders all agree on the intent of the ADP technology development effort and the priorities of the evaluation. Finally, the document should identify how the ADP being evaluated contributes to the overall policy goals of the ATTRI program.[4]

Step 2: Develop a Threat Model

ADP technologies focus on providing greater independence and opportunities to individuals within the context of the complete trip. As such, they are designed to overcome problems, risks, or threats that users typically face in day-to-day travel scenarios. Consequently, the technology must provide some level of safety and security guarantee, or risk mitigation, to users. Most ADP teams have thought explicitly about what protections their technology does and does not offer, and a discussion between the IE and ADP teams should occur to enumerate them.

Knowledge of the target population(s) gives the IE team insight into the vulnerabilities of that population, and thus the “threats” that need particular attention as part of the evaluation. For example, if the ADP will be used by people with low vision, does the failure of the device place users in danger, and how does the ADP technology account for that possibility by identifying when such a risk might occur and providing cues to users so they can mitigate that risk?

---

[4] Corhahl, Gustave, A. Auer, A. Cohen, and J. Broader, *Accessible Transportation Technologies Research Initiative (ATTRI) Policy and Impacts Assessment—Policy Assessment, Gaps & Needs—Final Paper*, July 2019.

Given these factors, the next step in creating the IE plan is to develop a Threat Model.

As presented in chapter 2, a Thread Model is an understanding of the things that could go wrong with the operation and use of the ADP technology, as well as an understanding of the ways the technology is designed to address those failure points if and when they occur. It covers topics such as the following:

- Mechanisms that could cause the technology to fail and that are critical for understanding the overall robustness of the ADP technology's performance (e.g., What happens if a loss of communications occurs, or if a user encounters an unexpected travel outcome and needs help?).
- Safety concerns that could result from use of the technology (e.g., Are there ways in which use of the ADP technology could put users in harm's way, such as texting while driving with a smartphone, or leading an individual into an environment that s/he cannot safely traverse?).
- Indirect threats to the technology user, such as invasion of user privacy or cyber security concerns associated with use of the technology (e.g., How does the technology prevent a stranger from obtaining a secret access code if it is part of wayfinding instructions for a user?).

Developing a useful threat model requires both an understanding of the target population and the technical performance of the ADP, information that should have been learned as part of the project set-up. Most ADP teams have thought explicitly about what protections their technology does and does not offer, and a discussion between the IE and ADP teams should occur to enumerate them. In general, each IE should look for threats in three different areas:

- Technical Stability or Reliability: The ADP technology must be as stable and error-free as possible, so travelers are not left without access to personalized information and resources when they need them in the middle of a trip.
- Robust Failure Modes: Should a technology fail (or crash), reasonable messaging should be available to let users know this has occurred and how they can connect to assistance, emergency, or other services as needed.
- Contextualized Assistance and Decision Support: Should users identify a safety or security threat; they should be able to connect to context-relevant resources or people through the technology.

Within each of these basic categories, the IE needs to reflect the capabilities and needs of the target population to identify which types of failures generate the greatest risk, identify how those risks can be mitigated, and then determine ways to test for the presence and performance of those capabilities. The threats determined by examining the intended operation of the ADP, given the needs and capabilities of the target population, then need to be discussed with the ADP team and the evaluation sponsor to determine how to prioritize them within the evaluation effort.

Example Robotics and Automation—Threat Model

Traveler safety is essential in any setting where people are riding public transit, especially if they are receiving transportation services/assistance. Traveler safety is not only contingent on the infrastructure and services provided, but also requires that the system provide proper cues for travelers so that they can respond to the environment effectively.

There are several major areas in which threats to safety need to be highlighted:

- Travelers and the robot are unable to connect, leaving the traveler stranded without assistance.
- Travelers are unable to correctly inform the robot of their desired destination, resulting in the robot taking travelers to the wrong location.
- Travelers are unable to follow the robot, thus stranding travelers within the transit center
- The robot is unable to follow the path selected by the ADP algorithm from the origin to the destination, leaving both the robot and travelers stranded in the transit center.
- Travelers fall as a result of attempting to follow the robot.
- A hacker takes over the robot (or ADP) and intentionally leads travelers into an unsafe environment.

To address the above threats, the following issues need to be included in the evaluation, either through a review of the results of ADP team's tests or by the IE team performing tests as part of the independent evaluation. Other important user safety topics include the following:

- The degree to which collisions and near collisions occur or are avoided by both the robot and the traveler.
- The ability of each user to safely follow the path selected, given each user's capabilities (the safety risk is that the identified path is not safe for that specific individual because it contains features that are beyond the capabilities of that specific user).
- The ability of users to effectively obtain and understand the navigation instructions when those instructions are provided aurally.
- The timing of those instructions, to ensure that travelers following those instructions do not make turns or other movements earlier or later than intended.
- The reliability of the robot itself (What happens if the system fails? How often does it fail?).

Finally, because the robot needs to navigate through what can be a very crowded environment, the robot contains features that allow it to "see" its surroundings to avoid both fixed and moving objects. It is an important safety outcome to also determine the reliability of this function, and to show the degree to which the robot is able to help travelers it is assisting avoid those same collisions, especially in crowded conditions.

Step 3: Develop Evaluation Logic Models

At this point in developing the evaluation plan, the IE team should have a strong understanding of the following:

- The population(s) being served by the ADP and that population's needs and vulnerabilities.
- The travel activities the ADP will be used to support and how it intends to change how individuals will accomplish those activities.
- The technical functions the ADP needs to perform.
- The potential points of failure for that technology.

- The strategies the ADP team has implemented to guard against those failures and mitigate hazards if those failures occur.
- The priorities of the evaluation sponsors.

Given this information, the next step in creating the evaluation plan is to create a document that formalizes what questions the evaluation needs to answer, identifies the metrics that will be used to quantify the answers to those questions, the data elements that can be collected to provide those metrics, and the sources of those data. This allows the IE team to understand the scope of the evaluation effort. These elements are typically documented with a logic model, which formalizes the evaluation hypotheses and describes the details by which those hypotheses are to be tested.

When developing logical-model hypotheses that evaluate the performance of travel activities, it is also important to consider the technical and usability contexts associated with **why** the travel outcomes occur. These evaluation questions cannot be effectively answered without also answering a number of the key technical performance and user interface questions of interest to the ADP team, as knowledge of one or more of these technical performance outcomes is typically required to explain travel performance outcomes, especially when the travel performance outcomes are worse than desired. For example, the data collection associated with evaluating a hypothesis, such as “Are the target populations better able to transfer between transit lines in an underground station?” need to include not only the volume of transfers successfully made—and attempted—but also **why** the transfers that are not successfully completed were not successful. Are those failures because of a specific technology failure, or because the user interface was insufficient in some manner for communicating with the traveler? Or some other reason?

One or more performance metrics are needed to evaluate each of the evaluation hypotheses. The ATTRI program has recommended a number of key performance metrics to be used nationally. These can be found in the Federal Transit Administration (FTA) Report “Mobility Performance Metrics (MPM) for Integrated Mobility and Beyond, published in February 2020, see table 2 to table 6 starting on page 27. However, given the wide range of topics that specific ADP’s involve, this table may need to be supplemented.

Appendix A in this report provides a series of tables which can be used by IE teams to identify project goals and hypotheses that can be considered for inclusion in the evaluation, as well as appropriate performance metrics for each of those hypotheses and interactions the IE team is likely to need to have with the ADP team in order to successfully obtain the data needed and apply for those performance metrics.

As noted in step 1c, the relative importance of the potential evaluation topics (e.g., overall travel outcomes versus technical performance of the ADP) is a function of the overall goals of the IE project, but some consideration of all six evaluation contexts needs to be part of the development of the hypotheses. Discussions among all stakeholders will ensure that the maximum benefits are obtained as a result of the evaluation project.

In addition, fully involving the ADP team and IE team in the final logic model development will help ensure that all parties are aware of the data collection requirements, as well as their roles and responsibilities in collecting, quality assurance testing, and analyzing those data.

The outcomes of this interactive refinement of the initial logic models will be a final plan for the evaluation and assigned roles for the project participants.

To formulate the evaluation's hypotheses, the IE examines the ADP's ability to provide users with improvements in their ability to complete one or more travel activities, mitigate threats which can potentially occur during travel, and address the specific needs of the target population. These three sub-steps are addressed below.

Step 3a: Developing a Logic Model to Evaluate Changes in Performance of Travel Activities

The first step in developing this logic model is taking the TALs identified in step 1b, the target population identified in step 1a, and the improvements the ADP is intended to achieve from step 1c and developing hypotheses that examine the **effectiveness** of the ADP in achieving those desired outcomes for those populations. Performance metrics need to be selected at this time as well. For example, if the ADP is intended to help travelers plan trips, hypotheses need to be developed which test the improvements in planning trips the ADP target population achieves when using the ADP. Performance metrics that quantify these improvements might include increases in the number of available trip options provided, the ease of use of the system, the time required to plan the trip, and the quality (decreased travel time, decreased number of transfers, lowered cost) of the trips offered by the system. If the ADP was intended to help travelers navigate through underground transit stations, the hypotheses would be developed to determine whether the use of the ADP improved the ability of the test population to travel through underground stations, and might include performance metrics that examined the time spent navigating the station, the accuracy of the navigation directions, and the ease of use of the system.

Appendix A includes a series of tables which, using directed questions, can help guide an IE team through the process of identifying hypotheses and performance metrics for many of the key evaluation topics. Examples of effectiveness-derived questions to drive evaluation hypotheses include topics such as the following:

- Can the target population perform a TAL that is a travel barrier for the target population faster, more easily, with fewer errors, or more safely?
- Does the use of transit (or other targeted mode) by target populations increase once the ADP has been deployed, and if so, by how much?
- Are target populations more likely to travel with aid of the ADP?
- Can those populations make more spontaneous travel decisions (e.g., make trips on short notice, or change destinations midtrip)?

Another aspect of the complete trip is **efficiency,** which measures an ADP's performance by comparing how many resources the user spends in order to complete their trip before and after the introduction of the ADP. For example, using the previous example of navigating the underground station from above, efficiency would be measured in how long it takes for the traveler to pass through the station. It can also be measured in terms of the number of attempts it takes a user to perform a task (such as how many attempts it takes to find an entrance or board a vehicle.) When analyzing a diverse population, it does not make sense to attempt to measure overall trip efficiency. Instead, the IE typically focuses on the relative change in efficiency that individual users experience for specific TALs.

Examples of efficiency-derived questions to drive evaluation hypotheses can be found in appendix A. include the following:

- Can the target population perform a TAL (e.g., board a vehicle? plan a trip? find an entrance?) faster using the ADP than without using the ADP?
- Does the traveler make fewer errors when performing a TAL (e.g., follow a navigation path? correctly plan a trip?) when using the ADP?
- Does the distance traveled by the user decrease after adopting the ADP?
- Does the user require fewer transactions (or maintain fewer accounts) in order to make a multimodal or multi-agency trip after the implementation of an integrated system (e.g., trip planning or payment)?
- Does wait time for a service decrease after the implementation of the ADP?

The final element of complete trip analysis is **empowerment and equity.** That is**,** does the ADP increase the target population's ability to travel freely and spontaneously, as other groups currently travel? A two-pronged approach to empowerment and equity is recommended. First, qualitative measures of user satisfaction compared to previously measured satisfaction metrics with similar travel or travel links help identify whether uses are more empowered. Second, assess whether the technology can be used demonstrably by the primary user population without compromising some other benefit of community living in order to use the technology. (This is called a "Nonadversarial Tradeoff.") That is, are the tasks the target population is being asked to perform as part of the ADP fair? Or do they have to give up privacy, security, or make themselves standout, in order to use the ADP? The resulting hypotheses address the target populations' general satisfaction with the ADP, as well as their ability to utilize the ADP without sacrificing other benefits. Examples of empowerment and equity-derived questions to drive evaluation hypotheses include the following:

- Can users of all abilities in the target population use the ADP technology effectively and efficiently in the environments in which they will be using the ADP (e.g., in noisy and crowded conditions)?
- Do users feel confident and in control while using the technology?
- Do users feel comfortable using the technology in public?
- Does the technology afford users the ability to make choices?
- How easily and inexpensively can transit agencies and cities widely deploy the ADP (assuming that the ADP generates significant improvements for the target population)?

An initial logic model covering topics from step 1c for the example Robotics and Automation example is shown in table 3.

Example Robotics and Automation Travel Outcome Logic Model (Step 3a)

For this and the following logic model examples, multiple features (e.g., there are multiple ways to communicate with the robot) have been incorporated into a limited set of hypotheses, and the same basic outcome metrics are used to evaluate them (e.g., detailed hypotheses are not present for each of the communication methods). For a complete independent evaluation, these different communication approaches would be explicitly examined, and thus would appear in the logic model. The example's more limited approach was done to limit the size of the logic model in this report and to reduce duplication of many basic hypotheses. In addition, the logic model concentrates on the overall outcomes of the major tasks that must be performed for the robot to both successfully function and interact with the user. The detailed technical reasons for why failures occur are left to more detailed technical performance testing which this example assumes is being performed by the ADP team.

Using the working Robotic example, a portion of the logic model is shown in table 3. Table illustrates several evaluation hypotheses that designed to evaluate the wheeled robotic assistant ADP's ability to improve travel activity outcomes for individuals with low vision, who need to transfer between transit vehicles at a large, outdoor transit center. Table 4 shows an example logic model for evaluating threats those individuals face when using the ADP, and table 5 presents a logic model examining user needs.

Table 3. Example entries from a logic model to evaluate changes in performance of travel activities for a robotic assistant for low vision travelers at an outdoor transit center.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
Users are able to more quickly and accurately transfer between buses. Technical function (Outcome).	Use of the ADP decreases the time it takes users to transfer between buses within a transit center.	1. Change in travel time between bus exit and arrival at next boarding location.	1. Travel time between exiting of the first vehicle and arrival at the correct transit stop.	1. Field data collected at the site. 2. Data collected from participants smart phones via the ADP app both before and after deployment of the robot.	Discuss with the ADP team to determine whether travel time data are available from the smartphone app.
The target population is able to take more trips using transit. (Empowerment/ Technical function (Outcome).)	Use of the ADP results in more use of transit by the target population.	1. Number of trips made (per week) by target population individuals should increase. (Should be computed separately for different target populations, e.g., low-vision versus full vision but wheelchair user.) 2. Survey response about the likelihood that transit trip making will increase.	1. Number of transit trips made before/after the ADP deployment. 2. Post-deployment survey question about expected changes in trip making due to the ADP.	1. Field data. 2. Post deployment survey of test subjects.	The before/after approach assumes that "before" data can be obtained on the test population and that sufficient travel is expected to/from the test deployment area that statistically significant changes in trip making can be measured. The stated-preference response via a post-deployment survey is a backup to that measure but provides useful insight in any case.
Users are able to effectively communicate with the robot. (Usability.)	Users are able to easily call the robot and identify the robot.	1. Percentage of times users successfully connect with the robot. 2. Change in the mean number of device interactions required to communicate with the robot. 3. Change in the percentage of device interactions required to successfully communicate with the robot. 4. Mean user satisfaction rating.	1. Number of times user meets and connects with robot. 2. Number of attempts to meet with robot. 3. Number of successful and total device interactions while attempting to communicate with the robot, computed separately for crowded/noisy conditions versus noncrowded/ quiet conditions. 4. Likert Scale response to question about ease of use.	1. Smartphone data from the users' phones. 2. Field data collected by observing user behavior at the site. 3. User satisfaction survey.	Communication is needed with the ADP team to determine whether data on device interactions are available from the smartphone. Also collect data on why failures occur. Measures are computed separately for crowded/noisy conditions versus noncrowded/quiet conditions and results compared.

Table 3. Example entries from a logic model to evaluate changes in performance of travel activities for a robotic assistant for low vision travelers at an outdoor transit center (continuation).

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
Users are more confident of their ability to travel safely. (Empowerment.)	Having access to the ADP technology increases confidence in the ability of users to travel safely.	1. Mean users satisfaction rating based on Likert Scale response to questions about the users' perception of their level of travel mobility.	Likert scale survey questions to be answered includes: 1. Do you feel confident that you can make trips more safely with this technology? 2. Does having access to this technology make you more confident in your ability to take transit?	1. User surveys after the field test.	None.
ADP provides mapping/navigation to the robot and to the users. (Technical function)	The robot can successfully select and follow different routes within the station, when the needs of different users require different paths through the transit center.	1. Percentage of correct routing solutions provided by robot given different user requirements. 2. Percentage of trips with robot navigation errors. 3. Percentage of trips with fail due to robot navigation errors. 4. Percentage of trips that deliver the user to the correct waiting point for their arriving transit vehicle. (not just the correct stop, but the correct location at that stop.)	1. Number of trips made. 2. Number of routing solutions developed. 3. Number of valid routing solutions developed. 4. Number of navigation errors made by the robot. 5. Number of failed trips due to navigation errors. 6. Number of trips that deliver the user to the correct location from which to board their next transit vehicle.	1. Field test. Routes to be extracted from the database of the robot. 2. If Lab tests are performed to examine the robustness of alternative paths and the sensitivity of the path selection to user profile settings, the details of the inputs to those tests would be recorded at the time of the tests.	The field test can determine the effectiveness of the routing solutions used. Lab tests that specifically examine how those routes change due to different profile settings and the validity of those routes, is likely best done by the ADP developers, with the results summarized and given to the IE team.
The robot selects and navigates the correct path between exit and boarding stops. (Technical function/Technological robustness.)	The robot is able to identify the bus/route/stop that is the destination of the user and can successfully select and navigate different routes within the transit center when the routing needs of the different users require different paths through the transit center.	1. Percentage of correct bus stop selections made by the robot. 2. Percentage of correct routing solutions provided by robot.	1. Number of correct bus stop selections and total number of attempts to identify bus stops made by the robot. 2. Number of correct routing solutions and total number of attempted routing solutions by the robot.	1. Smartphone data from the user's phones. 2. Additional field/lab tests as needed.	Tests need to be performed for different user requirements (e.g., for users with canes versus users in wheelchairs).

Source: Federal Highway Administration.

Step 3b: Developing a Logic Model to Evaluate the Ability to Mitigate Threats

To evaluate whether users of the ADP are potentially exposed to serious negative outcomes, the IE should identify risks associated with using the ADP, and then explore how well the ADP mitigates those risks.

The first point of analysis for threat models is **stability.** The IE should determine if the ADP functions without error, per its specifications. The evaluation must consider the target population's needs and capabilities and identify the degree of risk associated with each identified possible risk area. The evaluation logic model hypotheses and tests should be designed to determine if the ADP is reliable enough to safely and reliably meet the target population's needs.

For most ADP evaluations, at least some logic model hypotheses are designed to assess the **robustness** of the ADP's failure modes. Within the threat model analysis, the focus in robustness is on not just how often failures occur, but how well the ADP recovers from those failures, and if the recovery measures in-place are enough to protect the user from unsafe outcomes. For example, an ADP that includes navigation, should be able to identify when the user is off route, and assist the user getting back on route.

The next area evaluation topic the IE team should consider including in the logic model assesses the ADP's ability to **provide contextualized assistance** to the target population users. This means that the IE team needs to examine the measures in place within the ADP in the event of a security risk to the user. The IE team should consider hypotheses that explore whether such contextualized help exists in the ADP, and if that function exists, include tests for the effectiveness of user's ability to gain outside assistance from within the ADP interface (for example, calling someone for help).

Examples of questions that can be used to help develop questions to drive evaluation hypotheses include the following:

- How often does the ADP fail, either because the hardware/software fails, or because the user is not able to operate the technology (e.g., the payment system does not work, the trip planner fails to make a trip, or produces a poor trip plan, etc.)?
- What are the safety implications of a failure?
- If a user identifies a failure, are they able to request assistance? In what timeframe and with what outcome?
- If the ADP fails in some way, are there safety implications for the user? How are those safety threats mitigated? Can those mitigations be tested within the IE?
- If a failure occurs (e.g., the user goes off-path for a navigation app), does the ADP identify this threat and provide a safe remedy for the user (e.g., a new, safe path from their current location)?

An initial logic model covering topics from step 1c for the example Robotics and Automation example is shown in table 4.

Table 4. Example entries from a logic model to evaluate the mitigation of threats for a robotic assistant for low vision travelers at an outdoor transit center.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
The target population is able to successfully follow ADP navigation instructions. (Technical function (Outcome).)	Users can successfully follow the robot through the transit center, even in noisy or crowded conditions.	1. Percentage of trips in which users go off-path. 2. Percentage of trips where users stumble or fall. 3. Percentage of trips where the user can not follow navigation instructions (navigation for the trip fails). 4. Percentage of trips where some form of navigation assistance is required. 5. Likert scale response to questions about the ability of the user to effectively follow the robot.	1. Number of trips taken. 2. Number of off-path occurrences. 3. Number of falls or stumbles during a trip. 4. Number of trips for which navigation assistance is required (trip fails, instructions must be repeated, or requests are made for the robot to stop and wait). 5. Responses to survey questions concerning: 6. Ability to follow robot. 7. Ability to understand the robot's navigation directions. 8. Impact of noisy/crowding on those abilities. 9. Method used to obtain navigation instructions (auditory, or tactile).	1. Field test. Event data to be obtained from both the user smartphone app and the robot's software database. 2. User survey after each trip (to obtain information on stumbles, falls, other). 3. Survey should also contain free form text description of any issues associated with the entire trip. 4. Post deployment test user survey. 5. Data on the level of crowding and noise during each trip.	While the app data can describe the number of navigation errors occurring, it is also important for the evaluation to learn why those errors occur. Trips should be taken in both quiet and noisy/crowded conditions, with results of these analyses compared across conditions. Performance should also be reported by type of robot following mechanism (i.e., tactile following versus auditory instructions).
The robot can lead low visions travelers safely around both moving and stationary objects. (Technical function/Threat Model.)	The robot can correctly detect, avoid, and lead users around permanent and temporary obstructions.	1. Percentage of obstructions in the path of the robot that are avoided. 2. Percentage of obstructions avoided by the user. 3. Percentage of moving objects avoided by the robot. 4. Percentage of moving objects avoided by the user.	1. Number of obstructions placed in the path of the robot and the number of them that are avoided. 2. Number of obstructions avoided by a) the robot; b) the user. 3. Number of moving objects that present a potential collision threat. 4. Number of those moving objects avoided by a) the robot; b) the user.	1. Field test. Collect data (likely manually) on the presence of fixed and moving objects, and their impact on the trip being made.	Is the robot assistant able to collect these data based on its environmental sensing system? If so, use those data. Otherwise self-reporting from users or field observations are required. This test likely requires field validation of the outcomes, as independent validation of the number of fixed and moving objects needs to be performed. This test may be best performed by the ADP team as part of the initial technology development, rather than as an aspect of the IE.
Users are notified when system failures occur and have access to help. Empowerment/Technological robustness. (Note: identified as an issue within the Threat Model.)	The technology effectively notifies the user when the robot does not correctly respond to a call, and the Help Button functionality is able to provide the user with the required assistance.	1. Percentage of equipment failures that result in help notification of the user. 2. Percentage of users that can identify and activate the Help feature. 3. Number of successful help information transfers after use of the help feature. 4. Mean user satisfaction rating of the Help feature.	1. Number of successful technology failure notifications during testing, and the total number of those tests. 2. Number of successful Help requests, and total number of Help requests attempted. 3. Number of successful actions taken given a call for assistance, and the total number of calls for assistance. 4. Likert Scale response to survey question about use of the help feature.	1. Smartphone data from the users' phones. 2. Field or laboratory data collected by observing user behavior given technology performance at the site.	It is likely that the Help feature will need to be tested in a controlled environment, with artificially generated equipment failures. The ADP team should be consulted on what those failures should be and how to generate them in a controlled but realistic environment.

Table 4. Example entries from a logic model to evaluate the mitigation of threats for a robotic assistant for low vision travelers at an outdoor transit center (continuation).

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
The ADP provides accurate connection information (scheduled arrival by route and direction, stop location). (Technical function/ System integration.)	Via an automatically updated data feed, the robot is able to identify the bus/route that is the destination of users being helped and can successfully travel to that location In the transit center.	1. Percentage of correct responses to requests for transit departure bay information.	1. Number of requests for transit schedule and departure information. 2. Number of correct responses to questions about transit schedule and departure information.	1. Lab tests performed specifically to test the accuracy of the schedule information and its updating during a schedule shake-up.	It is assumed that the accuracy of the schedule information and its updating, is best performed as a lab test over the course of a schedule shake up, given the limit number of trips occurring during those time periods.
The ADP provides accurate connection information (scheduled arrival by route and direction, stop location). (Technical function/ System integration.)	Via an automatically updated data feed, the robot is able to identify the bus/route that is the destination of users being helped and can successfully travel to that location In the transit center.	1. Percentage of correct responses to requests for transit departure bay information.	1. Number of requests for transit schedule and departure information. 2. Number of correct responses to questions about transit schedule and departure information.	1. Lab tests performed specifically to test the accuracy of the schedule information and its updating during a schedule shake-up.	It is assumed that the accuracy of the schedule information and its updating, is best performed as a lab test over the course of a schedule shake up, given the limit number of trips occurring during those time periods.
ADP is able to ingest data automatically on changing routes and vehicle arrivals via standardized input. (System integration.)	The technology automatically updates transit route and stop information based on GTFS feed.	1. Percentage of routes that update stop location and arrival times automatically when a schedule change occurs.	1. Bus stop locations for each route. 2. Scheduled bus arrival and departure times in the ADP and as published by the transit agency via their GTFS feed.	1. ADP database. 2. Transit agency GTFS data feed.	Test performed at next routine schedule change.

Source: Federal Highway Administration.

Step 3c: Developing a Logic Model to Evaluate the Ability to Address Target Populations Needs

The third portion of the evaluation project’s logic model is used to determine the extent to which the ADP addresses the user needs identified in the ATTRI User Needs Assessment Report as it pertains to the population addressed by the ADP. For this effort, the IE team can use the User Needs checklist, as it pertains to the target population as provided in table 2. From that list, the IE team can determine not only which user needs are being addressed by the ADP, but which ones are most important contributions of the ADP to the ATTRI subpopulation whose needs are being addressed, and therefore should be included in the evaluation. Hypotheses and other logic model details can then be developed to test the impact the ADP has produced in those areas of need or barriers.

An initial logic model covering topics from step 1c for the example Robotics and Automation example is shown in table 5.

Summary of Step 3

Once these tasks have been performed initially, the resulting logic model will describe the evaluation questions the IE team would like to answer. Working with the ADP team, the IE team should also determine the data sources available for answering the questions posed by the hypotheses in the logic model.

The review of available data sources determines what data can be obtained from existing sources, and what data collection must be performed specifically for the evaluation. When these data collection requirements are combined with an understanding of the work tasks required to generate the desired performance metrics and perform the required analyses, it is possible to develop a good cost estimate for performing the evaluation. When available resources are insufficient for this plan, stakeholders can then make decisions about how to refine the evaluation to keep its scope within available resources. The outcome of this refinement process is a final plan for the evaluation. This plan must include a clear understanding from all parties what their roles and responsibilities are for collecting, quality assurance testing, and analyzing data.

Table 5. Example entries from a logic model to the ability of the Accessibility Development Projects to address target populations needs for a robotic assistant for low vision travelers at an outdoor transit center.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	3. Data Sources	Comments
Technology is usable by the identified user population. (Usability.)	Users are able to easily enter their origin/destination (O/D) within the transit center into the ADP technology.	1. Time required to enter the O/D of their trip. 2. Number of data entry errors that must be corrected before O/D are correct.	1. Start and end time of O/D entry process. 2. Number of entries provided. 3. Number of entries that must be corrected.	Field test: 1. Event data from the ADP app. 2. Alternatively, this could be a lab test performed as part of the system development.	See above.
ADP allows users to create a personalized profile that affects the technology's performance. (Technical function.)	Users are able to effectively communicate with the robot ADP technology by selecting and changing their profile settings.	1. Percentage of times the user is able to call the robot to their arrival station. 2. Percentage of attempts to call the robot that successfully bring the robot to the correct arrival point. 3. Percentage of successful interactions between the robot and user at the arrival point. 4. Percentage of trips with navigation errors caused by poor communication. 5. Likert scale response to questions about the ability of the user to communicate effectively with the robot.	1. Number of robot service requests. 2. Number of times robot meets the rider at the arrival point. 3. Number of times the robot correctly obtains destination information. 4. Number of navigation directions given by the robot. 5. Number of navigation errors made by the user due to poor communication. 6. Type of communication medium selected by the user (via smartphone, or direct auditory). 7. Survey responses to questions about the effectiveness of the medium selected to communicate with the robot.	Field test: 1. Event data to be obtained from both the user smartphone app and the robot's software database. 2. User surveys after the field test given to test subjects.	The IE team needs to work with the ADP team to ensure that event data is retained in a database and made available to the IE team.
The ADP provides amenity information at transit centers. (Technical Function)	The robot is able to identify, locate, and take the user to amenities at the transit center.	1. Percentage of correctly answered requests for the presence of amenities. 2. Percentage of trips to amenities that successfully arrive at that amenity.	1. Number of requests for the presence of amenities. 2. Number of correctly answered requests for the presence of amenities. 3. Number of trips attempting to go to an amenity. 4. Number of trips to amenities that successfully arrive at that amenity.	Field test: 1. Event data from the ADP app should indicate the number requests for amenities, and if the response to that request is correct. 2. Trace data from the app can be used to determine if the users are successfully led to that amenity.	The IE team needs to ensure that these events are recorded by the ADP app, and can be extracted from the app and the robot. The field test may need to be supplemented by specific lab tests performed in the field, as the number of requests for amenities may be too small from the test subjects to effectively evaluate this ADP function.

Source: Federal Highway Administration.

Step 4: Perform the Evaluation

Once the logic model has been finalized, the evaluation needs to be performed. Tasks that typically need to be performed by the IE team are briefly described below.

Obtain the appropriate approvals for handling sensitive data. For university, nonprofit organizations, and hospital personnel this typically requires completing, filing, and obtaining approval to collect, analyze, and report sensitive information about vulnerable human test subjects through their Institutional Review Board (IRB) For private companies, similar approvals are typically required The IRB (or its equivalent) process ensures that the appropriate safeguards are in place in the study, to ensure that no harm will come to test subjects as a result of their participation in the project. IRB documentation will need to include information describing how test subjects will be recruited, compensated, and treated during the project.

- Will the IE team use the test subjects already recruited by the ADP team, or are new subjects required?
- The IE team must establish the need for individuals with different disabilities within the test subject population, as well as determine how to change the evaluation if the desired test subject sample size cannot be recruited within a target population.
- If the IE team is doing their own test subject recruitment, the IE team will need to develop participant onboarding, training and ongoing communications protocols with those individuals, as well as ensuring that each participant has access to the appropriate equipment for participating in the test (e.g., a smartphone with the ADP software downloaded and successfully tested.) This often involves working with members of the local disability stakeholder community, for whom outreach and ongoing communication plans and protocols are required.

Create a data security plan. As part of the human subjects review, the evaluation team also needs to develop a data security plan. The data security plan identifies all information about individuals being collected, evaluates the potential risks that are associated with those data, designs a program to protect that data, and puts in place a system which routinely monitors and tests those protections.

Requesting and obtaining missing background material on the ADP technology. If the IE team does not already have detailed information on the ADP technology being used, any connections that technology has to other traffic and transit management systems, and the details of tests being performed by the ADP team, this information should be gathered at the beginning of the evaluation effort so that it can be used to develop the data collection and testing protocols.

Determine the sample sizes. As part of developing the detailed plans for performing the project, the IE team needs to determine the sample sizes required from their data collection effort in order to meet the accuracy objectives for each of their performance metrics. This requires not only a detailed understanding of statistics, but an understanding of the nature of the data they will be collecting, and the analytical techniques IE team intends to apply.

Finalize the data collection plan. Once the team understands the data they need to collect, the sample sizes required, the methods that are required to collect them (e.g., are the data being collected as part of the ADP technology, does data collection need to be added to that technology, does additional automated data collection technology need to be purchased and installed, does manual field data collection need to take place?), and which specific group (e.g., the ADP team, participating public agencies, or the IE team) are responsible for collecting each data item, it is possible to finalize the data collection. This includes fully documenting the data to be collected, assigning each data collection element to a specific individual or firm, and describing how those data will be delivered to the IE team. Several key aspects to this data collection plan include the following.

- Finalizing plans for the recruitment of test subjects.
- If the IE team is sharing test subjects with the ADP team, the two teams need to coordinate their respective survey work to limit the impacts those surveys have on survey response rates and test subject attitudes towards the project. All surveys to be used should be thoroughly tested and refined prior to their distribution and use.
- The need for access to data collected through the ADP technology.
    - The IE team needs to work with the ADP team to understand the details of the data the ADP technology already collects, and how those data can be obtained by the IE team.
    - If the ADP technology does not currently collect data, the IE team requires, the IE team needs to arrange for that data to be collected, with by working with the ADP team to modify their current data collection system, or for the IE team to collect that data in some other manner (e.g., by arranging to download a data collection application to the test subject's smartphones.) These data collection systems then need to be tested to ensure that they do not create unexpected issues with the ADP technology, such as high power use that drains smartphone batteries.

Collaboration plan for participating agencies/forms/groups. While finalizing the data collection plan, the IE team should also work with all other participants in the study (e.g., the ADP team, the local transit agencies, road authorities, local communities, and disability community groups) to formalize the cooperation that needs to occur during the project. This includes understanding the project schedule in order to time data collection events, and having communications protocols in place that keep all parties aware of project progress and preliminary outcomes, while also updating the schedule when unexpected events occur.

Data quality assurance testing. The IE teams needs to routinely download data from any automated data collection systems, including those operated by the ADP team. These data then need to be subject to timely data quality testing, to quickly identify data quality issues. The IE team then needs to work with the ADP team or other stakeholders to resolve data quality or data collection issues identified as a result of those tests.

Preliminary data analysis. The IE team should plan on performing preliminary analyses using the downloaded data to identify the need for changes in the data collection protocols, or even changes in the evaluation scope of work (e.g., a specific user error is occurring for which more data needs to be collected.)

Ongoing project communication. The IE team needs to design and implement a robust communication system with all participating stakeholders. Routine communication needs to occur so that each participant is aware of the progress of the ongoing field tests, data collection activities, data analysis, and preliminary findings. Routine communication allows the IE team to stay up-to-date with project events, and learn how participants are reacting to those events. It also gives the IE team early insight into when changes in data collection or planned analyses need to occur as a result of unexpected project outcomes.

Finally, as the ADP test is completed, the IE team needs to perform final analyses, draw conclusions, and write the project reports.

Step 5: Potential Gap Analysis

Once the basic analysis described in the Logic Model has been finished, if resources allow, a “complete trip gap analysis” should be performed. The complete trip gap analysis is designed to examine in a qualitative manner, the expected impacts of the ADP on the variety of trips made by the target population and to assess the project’s contribution to travel by the target population within the context of the larger transportation network. It is designed to identify and highlight gaps in the target population’s ability to make trips, given the intended deployment of the ADP technology being evaluated. It is understood that ADPs are developed with a more singular focus but the hope is that multiple ADPs will be integrated together to form a solution for the complete trip.

The gaps identified will be used to help guide the selection of the next steps that should be taken to gain the most benefits from the ADP technology so that the target population can travel more effectively. Gaps might include the need to provide supporting databases before the ADP can be deployed, or the need for additional ATTRI technologies to support the target population reaching the TALs which is being supported by the ADP being evaluated.

The complete trip gap analysis will also be used to help guide the selection of future technology improvement projects, as funding agencies and organizations look to provide further travel mobility to the target population. Understanding the types of trips that still face major barriers, even if the technology being evaluated was deployed, will allow these agencies to prioritize their available funding and to make calls for ADPs that address these specific barriers.

As previously discussed, the IE tests can only examine a small subset of complete trips. That is, the IE will have been explicitly designed to test the performance of the ADP in a given set of travel environments and for a given set of TALs. This is necessary simply because of the scope of deployment tests, which take place within a given subset of environments.

In contrast the “complete trip analysis” is intended to give the sponsoring agency insight into the larger travel benefits that the ADP can provide, as well as insight into the remaining issues that may still prevent the target population from traveling as easily and freely as desired.

The complete trip analysis makes extensive use of the TALs and relies on the IE team’s ability to understand the target population, where the target population lives, and the types of trips members of that population need to make. Factors to consider when developing trip scenarios for the complete trip analysis include the following:

- The land use environment in which the target population lives.
- The attributes of the transportation modes and services used by the target population.
- The mobility needs of the target population.
- The mobility barriers the target population experiences.
- The types of trips that this population needs to make.

For example, an ATTRI project might be designed to help disabled veterans plan trips more easily and efficiently. Disabled veterans can be found in two very different land-use environments. Many live in rural areas where housing costs are low but where services (medical, groceries, and other retail services) are widely separated, and available transit services are limited in scope and frequency. Another large concentration of veterans live in large urban cities, where fixed-route transit is often available but not all origins and destinations within the city are accessible and where transfers between transit vehicles can be difficult for people with mobility issues.

Because of these very different environments, the new travel planning system might have very different impacts on veterans' ability to travel. In rural areas, the ability to make complete trips might be a function of the ability call for, reserve, and link different shared-ride services. "Efficient travel" might mean having the system determine when these services can be linked and telling veterans when to travel. Because rural transit options are often door-to-door, transfers would be easy. However, because travelers get dropped off and picked up in the same location, those transfers could also be potentially time consuming, and because of the distances involved and the need to meet other travelers' needs, the timing of that transfer could both be imprecise and unreliable. The simple lack of services (e.g., "no seats are available on the day you need to travel"), and the timing and unreliability of transfers might limit the ability of veterans living in these environments to travel, even if the ADP technology can tell them how a desired trip might be made.

In an urban environment, the planning system is likely to attempt to use existing fixed-route transit to the extent possible, both because it costs less than on-demand services and because it allows for more flexible departure and arrival times. However, in this case, transfer between transit vehicles might require traveling between the exit stop and the transfer boarding stop. This movement might generate a number of other barriers to travel (e.g., lack of sidewalks, safety issues, lack of bus arrival information), as the fixed-route transit vehicles would not wait for veteran using the ADP to arrive at the transfer boarding stop if they were delayed. Fear of this happening might limit the benefit a veteran might otherwise gain from the ADP trip planning system.

As can be seen in these examples, other barriers might limit our example veteran's travel opportunities, and those barriers would fall outside of the ADP being tested, which is tasked only with providing a usable trip plan. And those barriers might change from one environment and study population to another.

The goal of the complete trip analysis is to highlight these barriers—both those that have been removed by the new ADP (e.g., the example trip planning system might make reservations on the rural on-demand transit routes for the veteran, guaranteeing the trip would work) and those that have not yet been resolved (e.g., for that rural trip, the lack of sufficient service, or neither the traveling veteran nor the service providers might have good real-time knowledge of the other's location, making it possible for the transfer to fail).

The results of the complete trip gap analysis will be used to help guide the selection of future technology improvement projects, as funding agencies and organizations look to provide further travel mobility to the target population. Understanding the types of trips that still face major barriers will allow these agencies to prioritize their available funding and to make calls for ADPs that address these specific barriers.

Setting Up the Complete Trip Analysis

The IE team should start with their good understanding of the target population. The IE team should document the expected built environment for that target population, and the needs and barriers to travel that population experiences. (See the example above.)

Given the built environment and target population, the IE team should develop up to five trip scenarios representative of the types of trips that members of the target population likely need to make. The scenarios should vary by trip purpose, and where appropriate, mode. Where possible, most trips should be multimodal, either involving access or egress movements (typically walking or nonmotorized) to transit or requiring transfers between vehicles. These trips should be described in terms of the travel activities that must be linked to complete that trip.

An example is shown below for an ADP technology designed to help blind individuals navigate pedestrian and street environments. A set of scenarios to be evaluated for that ADP would include the population "blind or low vision individuals" and might include the following set of trips within a major urban area:

- Trip 1: Walking trip across downtown (from the office to a meeting, routinely performed) (TALs: 4b, 5, 6, 5, 3b, 4b).
- Trip 2: A transportation network company (TNC) trip across downtown (from the office to a meeting, a trip that is not routinely performed) (TALs: 1, 4b, 5, 3a, 10, 11, 4a, 5, 3b).
- Trip 3: Shopping trip from a suburban home to a neighborhood center via paratransit (TALs: 1, 4b, 3a, 4a, 11, 10, 3a, 4a, 3b, 4b).
- Trip 4: Commute from downtown to a suburban home (via transit) with a desired spontaneous stop at a local shopping center on the way home (TALs: 1, 4b, 5, 3a, 4a, 11, 10, 2, 1, 3a, 4a, 5, 6, 5, 3b, 4b).
- Trip 5: Commute from home to downtown, including a walk to bus and transfer to light rail, ending at a major underground transit station, and then walking to the office (TALs: 4b, 5, 3a, 4a, 11, 10, 3a, 4a, 9, 11, 3a, 4a, 10, 3a, 4a, 8, 3b, 4b, 5, 3b, 4b).

These scenarios differ from each other in the level of knowledge the travelers have about their environment, as well as the type of transit mode they use to enter or exit the pedestrian environment from, and therefore the nature and importance of different travel barriers that they face. In each of these trips, the ADP would play a significant role in helping low vision travelers navigate a portion of their trip. However, in most of these trips, a number of other travel barriers would exist for people with low vision.

The complete trip analysis is a **gap analysis.** It is designed for the IE team to identify both the types of trips that could be more readily made (assuming that the ADP successfully met its performance objectives) and to identify the significance of barriers that would remain for this target population.

Therefore, trips 1 and 3 might well become trips that low vision persons could make more confidently as a result of the new ADP technology. Trip 2 would still present a potential barrier, in that low vision individuals might face difficulties in 1) identifying their specific TNC vehicle; or 2) orienting themselves to their location when the TNC vehicle dropped them off. Trip 4 might contain barriers if the transit agency did not already operate geolocation systems that broadcast bus location information—either on a cell phone application or via in-vehicle voice announcements. Trip 5 would only be possible if low vision travelers had the ability to successfully navigate the complex underground transit station.

Examples of the gap analysis for the robotics and automation project are shown below.

Example 1: Robotics and Automation Gap Analysis

Two scenarios to be evaluated for the robotic assistant example ADP are shown for the population "blind or low vision individuals." The specific land use for these scenarios is a major urban area and its suburbs.

Trip 1: An individual with low vision needs to plan and then take a trip from their home in the suburbs to a business meeting in a part of downtown they are not familiar with. Their intent is to walk to the transit stop they routinely use outside their apartment, take that bus to the main transit center in their suburban city, and transfer to a different bus than they take to downtown, exit the bus and walk to their final destination. The trip requires a trip planner (TAL 1) to determine the best route to take and to determine what time to leave; then the traveler must travel to their stop (TALs 3 and 5), catch and ride that initial bus (TALs 4, 10, and 11), then transfer at the transit center where the ADP provides assistance (TALs 3, 4, 9, 3, and 4), ride the second bus (TALs 10, 11, and 2), exit that bus (TALs 3 and 4), and then walk to their final destination (TALs 5, 6, 3, and 4.).

When looking at the first trip, the first noticeable gap in the traveler's ability to perform this complete trip is whether the traveler has access to a robust trip planner that they can use, given their limited vision. If such a trip planner exists, is that trip planning software linked to the ADP software, so that the ADP learns which transit vehicle the rider will be arriving on, and which route (and bus) they wish to depart on? If not, barriers to use of the ADP begin to form, as the traveler must negotiate more than one trip planning activity—one to obtain their primary trip plan, and a second one for their transfer activity at the transit center.

The next gap that might be identified is whether the rider can catch and ride their initial bus. In these simple trip examples, it is stated that this portion of the trip is routine for this individual. If that were not routine, a number of barriers exist that will limit the willingness of the rider to take this trip via fixed route transit. Adopted ATTRI User Needs include the need to improve these tasks by making it easier for riders with disabilities to identify their buses, board those buses, pay for their trips, find places to sit, determine when it is time to exit, and finally, debark from the vehicle. If any of these tasks presents barriers to the target traveling population—and many will—those barriers will reduce the number of transit trips that the target population is willing to take, despite the addition of the robot at the transit center.

Once the travel arrives at the transit center, the ADP takes over. However, once the rider has boarded the second transit vehicle, they are back facing the previous noted issues associated with riding a transit vehicle. Do they know where they are? Do they know when to get off? Do they know how to travel from their bus stop to their final destination and find the door to that building? These tasks remain issues for most of the target population, especially when they are using transit routes for which they are not familiar.

Thus, many additional barriers exist that might prohibit significant gains in transit use by the target population, even if the ADP technology works easily and flawlessly.

Example 2 Robotics and Automation Gap Analysis

Trip 2: A blind individual who lives downtown wishes to attend a music concert that will be held at a large church in the suburbs. Their intent is to walk to a transit stop near their apartment, take that bus to a transit center in the suburbs, and connect to a paratransit service which will deliver them to the church. The trip requires a trip planner (TAL 1) to order the paratransit ride, then a second trip planner to determine initial transit stop they need to use, and the time they need to leave to catch their initial bus in order to reach the transit center in time to meet the paratransit service; then the traveler must travel to their stop (TALs 3 and 5), catch and ride the bus (TALs 4, 10, and 11), then transfer at the transit center to the paratransit service (TALs 3, 4, 9, 3, and 4), ride the paratransit vehicle (TALs 10, 11, and 2), exit the vehicle (TALs 3 and 4), and then find the door to the church(TALs 5, 6, 3, and 4.).

For this second trip, the rider faces the same gap of needing to use two different trip planners, as most paratransit service providers are operated independently from the fixed route system, and may not be able to provide the fixed route service planning portion of this trip.

A potential gap also exists in Trip 2 that does not exist in Trip 1, in that the blind traveler is not using a route and stop they are already familiar with. This means that the accessibility of the path from their apartment to the required bus stop needs to be considered, and navigating a path selected without an understanding of the traveler's mobility could present a barrier to successfully making the trip.

The next gap identified is that in the case of Trip 2, the ADP technology also needs to be aware of where a paratransit vehicle will pick up and drop off a blind individual within the transit center. This information is needed by the robot in order to guide the rider to that location within the transit center. The ADP evaluation did not explore whether the robotic assistant's database contains information on where paratransit vehicles wait within the transit center, but such information is often not included in GTFS feeds, which are the source of that information for the robot. Thus, an enhancement to the initial ADP design may be needed to allow the robot to lead users to nonfixed route service providers at the transit center.

Based on this quick review, it would appear that one of the best ideas for the next project (presuming that this ADP is successful) might be to work on the trip planner designed specifically for the target population. And when working on such a trip planner, making sure that the planner integrated seamlessly with the ADP designed to help travelers transfer between vehicles at transit centers.

Chapter 4. Example Use of the Framework

This chapter provides a slightly simplified example of the application of the framework. The example is designed to provide a second illustration (in addition to what was presented in chapter 3) of how an independent evaluation of a technology development project designed to improve transit accessibility would be set up and performed. Consequently, the reader should have already read chapter 3 and be familiar with the steps required to set up an evaluation prior to reviewing this chapter. To limit the size of the resulting logic model, a limited number of Travel Activity Links (TAL) were selected for use in this example. The TALs are noted along with a description of the specific travel scenario to help illustrate the travel needs and barriers the Accessible Transportation Technologies Research Initiative (ATTRI)-funded Accessibility Development Projects (ADP) is attempting to mitigate.

A description of the ADP being tested is presented first. The description is followed by a summary of the "Set up" information that is required to create an evaluation. The TALs that are applicable to the project are then identified, a subset of those are then selected for use in the example. An example Logic Model is then presented for those TALs to illustrate a part of the more complete logic model that an independent evaluations (IE) team might create as part of designing an evaluation. Finally, a simple Complete Trip gap analysis is presented to illustrate the final portion of the Framework approach. The "Refine Logic Model and Develop Data Collection" task is not discussed to save space in this report, as these steps will be outcomes of the negotiation process among the IE team, the evaluation sponsors, and the ADP team.

Introduction to the Example Project

The example evaluation is applied to **an ADP that is designed to help individuals navigate across signalized intersections and along paths which can be used by a mix of pedestrians and vehicles.** The ADP test deployment is assumed to be occurring in a scenario where travelers will be arriving via transit near a major event venue in a busy part of town, and in order to travel between their transit stop and the venue, need to cross streets at intersections and navigate paths and parking lots.

The ADP is assumed to be designed for use by people with both low or no vision as well as by individuals who use wheelchairs or move slowly and need additional time crossing streets.

The ADP technology being tested combines two major ATTRI features. The first feature is technology identifies paths the traveler can safely use given their need to navigate environments where people and vehicles share right-of-way (for example, parking lots or woonerfs.) This feature relies on a database of pathway information that was created using a new data standard. The pathway database includes surface roughness and grade information, the locations of curb ramps, as well as the location of obstacles that need to be avoided along a selected path. One of the features contained in the pathway database is an indicator of which intersections are equipped with the second major ATTRI feature, a safe intersection crossing technology. This allows the navigation feature to route pedestrians to those intersections where they can cross in a safer and more efficient manner.

This second ATTRI technology helps individuals cross busy intersections by notifying travelers when it is safe to cross the street, and by ensuring that the crossing time provided by the traffic signal controller is sufficient to allow a slow-moving individual to safely cross the street. This ADP feature works directly with the traffic signal control hardware and software. It uses a smartphone along with connected vehicle technology to identify when a user has arrived at a signalized intersection and informs the signal controller of the crossing movement the user needs to make. The signal controller then changes the signal timing plan for the traveler’s crossing and informs the user when it is safe to cross the street. Once the traveler has completed the crossing, the signal controller reverts the signal timing plan to its normal operation.

The ADP does not perform real-time moving obstacle detection, either for the street crossing or for the spaces where pedestrians and vehicles share right-of-way.

Target Population

For this evaluation, the target population comprises users with at least one of two disabilities. The first is that the users may have low or no vision. The second is that they may move slowly, often because they are using a wheelchair or some other assistive device. Therefore, most travelers who need added time to cross a street would be potential system users, especially if those individuals needed assistance selecting travel paths with specific features, such as smooth surfaces and no tripping hazards.

Low and no vision users benefit both from increased time to cross the street and from the improved safety which results from being able to obtain audio notifications of when it is safe to cross the street because the correct phase is now green. However, **the very different abilities of the large and diverse target population,** low and no vision users versus those users with adequate sight but who have other mobility disabilities—especially the need to use mobility assistive devices—**means that the evaluation will need to split some evaluation tasks into two groups of tests subjects in order test outcomes for groups with very different mobility profiles.** Some tests will be performed only for those with low or no vision, while others will need to be performed separately for those using assistive devices. In other cases, the same test can be carried out, but the results for these two groups will be analyzed separately as the outcomes for these two groups may be different, and it is important to understand these differences.

User Travel Needs

The user needs being addressed by this combined ADP are shown in table 6. Short explanations about why the questions in the Evaluation Hypothesis Checklist are answered as shown in table 6 follow the table, to document the ADP project and help develop the evaluation plan.

Table 6. Evaluation hypothesis checklist for Accessible Transportation Technologies Research Initiative-funded development projects technology in the example.

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable?	Significance to ADP Objectives (L, M, H)
All	1. Does ADP provide information in a variety of accessible formats?	Y	H
All	2. Is information from ADP interface accessible in a variety of environments (i.e., amid heavy crowds and noise, underground)?	Y	H
All	3. Does ADP perform a task that improves safety and security or that provides emergency information?	Y	H
All	4. Does ADP provide en route assistance and information?	Y	L
All	5. Does ADP provide connection information (where, who, when)?	N	
All	6. Does ADP provide estimated trip length and distance?	Y	M
All	7. Does ADP provide comprehensive travel information?	N	
All	8. Does ADP require access to equipment (phones, computers, charging, training)?	Y	L
All	9. Does ADP allow the user to create a personalized profile?	Y	M
All	10. Does ADP require coordination of information (between agencies, modes)?	Y	M
Blind and Visually Impaired (BVI) (Visual) Motor Impairment (MI) (Motor)	11. Does ADP provide real-time transportation information?	Y	H (BVI) H (MI)
BVI (Visual)	12. Does ADP provide audible mapping/directions?	Y	H
BVI (Visual)	13. Does ADP provide destination information (hours, addresses, entrances, layout)?	Y	M
BVI (Visual) MI (Motor	14. Does ADP provide transit schedule and other transit information (e.g., stop location)?	Y	L (BVI) L (MI)
BVI (Visual) MI (Motor	15. Does ADP provide information about pathway infrastructure?	Y	H (BVI) H (MI)

Table 6. Evaluation hypothesis checklist for Accessible Transportation Technologies Research Initiative-funded development projects technology in the example (continuation).

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable?	Significance to ADP Objectives (L, M, H)
BVI (Visual) MI (Motor	16. Does ADP include provision for outside assistance or attendants?	N	
BVI (Visual) MI (Motor	17. Does ADP provide amenity information (e.g., restroom, shelter, benches, food, drinks)?	N	
BVI (Visual) Cognitive	18. Does ADP provide information about, and interpretation of, signage?	N	
MI (Motor)	19. Does ADP provide transportation facility information (e.g., maps)?	Y	L
MI (Motor)	20. Does ADP provide information about weather conditions?	N	
Hearing Cognitive	21. Does ADP include information about and/or interpretation of announcements?	N	
Hearing	22. Does ADP incorporate speech-to-text or text-to-speech that enables the user to communicate more easily?	N	
Cognitive	23. Does the ADP provide information in a concise and straightforward manner?	N	

Source: Federal Highway Administration.

1. **Does ADP provide information in a variety of accessible formats?** The ADP's smartphone application is designed to allow users to obtain information both aurally and visually. This is important both because some users cannot see and other users' hands may be busy holding assistive devices or propelling their wheelchair. Still others will prefer to read text or other notifications on their device screen. Some individuals may need to interact with their phone via voice commands and listen to voice instructions to follow the recommended path, others will prefer to text input to the system, and view their navigation instructions on their phone's screen. The personal profile is used to select the communication options for the interface. Testing the effectiveness of both of these forms of communication is a key evaluation task.
2. **Is information from ADP interface accessible in a variety of environments (i.e., amid heavy crowds and noise, underground)?** The ADP's smartphone application is designed to allow users to obtain information both aurally and visually. Both systems are designed to function effectively in street environments, which are both loud and often subject to bad weather (e.g., glare and rain.) Determining whether the system works as well in difficult environmental conditions as in good conditions is another important evaluation task. The "personal profile" (see Question #9 below) is used select between these interface options.
3. **Does ADP perform a task that improves safety and security or that provides emergency information?** The ADP technology is designed to provide two specific safety improvements. The first is to improve the safety of street crossings. It accomplishes this by identifying the best locations for crossing streets as part of its path selection process, by manipulating the traffic

signal phase length for pedestrian crossing at the selected locations, and by telling travelers when to cross on the basis of the walk/do not walk indicator status for that signal. The second safety benefit comes from the combination of the infrastructure database and the pathfinding function, which identifies specific hazards for users to avoid as they travel. It applies that knowledge through both the path finding algorithm used to route travelers from origin to destination and in the location-specific navigation directions it provides to help users follow that path. Quantifying the changes in safety which result from these ADP features is a key outcome from the IE.

4. **Does ADP provide en route assistance and information?** The current version of the ADP technology does not specifically incorporate a “request help” function. The path finding algorithm does allow users to “recompute their route” if they find their current route obstructed. While the evaluation will test this functionality, it is considered of less importance to the evaluation, relative to other tasks.
5. **Does ADP provide connection information (where, who, when)?** The ADP does not provide transit connection information, other than basic schedule information describing which transit routes stop at bus stops located within the map database and the scheduled arrival time for transit vehicles at each of those stops. (This information is available through the General Transit Feed Specification (GTFS) feed.)
6. **Does ADP provide estimated trip length and distance?** The path finding algorithm does provide a distance value as part of its navigation instructions. These will be checked as part of the IE, but they are less important than the paths themselves and the ability of the user to follow those paths.
7. **Does ADP provide comprehensive travel information?** The ADP does not provide comprehensive travel information. The ADP’s technological function is restricted to finding paths within the covered geographic area and improving the safety of users as they cross streets at signalized intersections within that covered geographic area. The ADP does have a basic origin/destination (O/D) interface that helps users plan their trips, but it is limited to the geographic area covered by the pathway and signal databases. These will not be evaluated as part of this project.
8. **Does ADP require access to equipment (phones, computers, charging, training)?** The ADP technology is smartphone based, so members of the intended user population must have access to a smartphone and the charging infrastructure needed to support that phone. In addition, the ADP requires that Digit Short-Range Communications (DSRC) technology be available to users and be connected to their cell phone. DSRC functionality must also be available in the traffic signal controller hardware. The IE will not specifically evaluate the availability of these devices, as this is an early stage test and participants will be provided with the required equipment.
9. **Does ADP allow the user to create a personalized profile?** The ADP’s smartphone application has features that allow users to set their preferred method of communication (e.g., audio versus visual/text, etc.), as well as define their form of locomotion or need for assistive devices (e.g., manual wheelchair, powered wheelchair, cane, guide dog, etc.). These are then used to both estimate the travel time required to cross streets (which affects signal timing) and determine the types of surfaces that should be used or avoided in selecting the traveler’s path. The responsiveness of the ADP to this functionality will be tested in the IE.

10. **Does ADP require coordination of information (between agencies, modes)?** The ADP requires integration with traffic signal systems that can employ the safe intersection crossing technology. The ADP includes the ability to adjust signal timing, which means that the ADP software must interact directly with the signal system software. Therefore, direct coordination with the agency that controls the traffic signal system is required to implement and operate the ADP. The IE will determine if the interaction with the signal system occurs as required. The IE team will document the types of integration required for this system if it were to be more fully deployed. No specific hypotheses or tests are required for this later task.
11. **Does ADP provide real-time transportation information, including 1) real-time vehicle status; or 2) real-time travel condition/obstruction information?** The ADP technology does provide real-time location information, both to help with navigation and to alert users of the presence and status of traffic signals. The ADP also tells users when it is safe to cross a signalized intersection. However, the ADP technology does not include real-time transit system information (e.g., bus arrival information). The accuracy of the location information will be tested as part of the intersection crossing evaluation and as part of the successful navigation of the paths between the transit stop and the venue.
12. **Does ADP provide audible mapping/directions?** The ADP's smartphone application uses a very detailed database that describes path attributes. These attributes inform navigation directions that guide users through their trip. These include not only directions on where and when to cross streets but also detailed routing instructions for the pedestrian paths that lead from the street to the venue entrance, given each user's specific mobility requirements and attributes. The IE will test the accuracy of the navigation directions based on the map database, identifying limitations in the mapping and navigation instructions.
13. **Does ADP provide destination information (hours, addresses, entrances, layout)?** The underlying map database and navigation software include detail that direct users to the venue entrances that meet their needs. This will not be a focus of the evaluation.
14. **Does ADP provide transit schedule and other transit information (e.g., stop location)?** The path finding algorithm includes transit stop location information (e.g., stop locations, the route numbers serving each stop, and the schedule for those routes). This information is obtained from the transit agency's GTFS data feed, so it is updated whenever schedule changes occur. This will not be a focus of the evaluation, as it is a minor aspect of the ADP features being tested.
15. **Does ADP provide information about pathway infrastructure?** This is one of the key attributes of the path finding algorithm and database technology being tested. The IE will evaluate the accuracy and completeness of the pathway database as a function of the errors in navigation which occur, and as a function of the ability of test subjects to traverse the paths they are provided.
16. **Does ADP include provision for outside assistance or attendants?** The ADP technology is not designed so that someone other than the user would access the technology. The technology is controlled exclusively by the user and does not communicate with others.
17. **Does ADP provide amenity information (e.g., restroom, shelter, benches, food, drinks)?** This set of ADP technologies is not specifically designed to include amenity information associated with buildings along the path. The pathway database does include the locations of benches and other infrastructure features that are considered amenities, but the current interface to the system is not

designed to search for or describe these amenities other than to identify them as obstacles to avoid. All testing of the accuracy of items in the pathway database will occur as part of item 12 above.

18. **Does ADP provide information about, and interpretation of, signage?** The ADP is “self-contained.” That is, the ADP technology does not interact with external signage and therefore does not provide interpretation of signs. Navigation directions are provided on the basis of internal map databases and the estimate of the user’s current location as identified by the technology (smartphone).
19. **Does ADP provide transportation facility information (e.g., maps)?** The map database includes transit stop locations as well as street layout information. The pathway database includes the locations of street furniture in the areas around the transit stops. No other transit facility information is currently included in the path attribute and street/sidewalk map database. This feature will not be a focus of the independent evaluation.
20. **Does ADP provide information about weather conditions?** The ADP technology does not include weather information.
21. **Does ADP include information about and/or interpretation of announcements?** The ADP is “self-contained.” That is, the ADP technology does not interact with external announcements and therefore does not provide interpretation of announcements.
22. **Does ADP incorporate speech-to-text or text-to-speech that enables the user to communicate more easily?** The ADP does not provide speech to text or text to speech functionality. It does provide information in both formats, but does not translate from one to another.
23. **Does the ADP provide information in a concise and straightforward manner?** The ADP is not designed for the cognitively impaired and therefore this is not an evaluation topic.

Technical Functions Performed by the Accessibility Development Projects Technology and Its Applicable Travel Activity Links

A travelers’ initial task when using this ADP technology is to select an origin and destination for a trip. In future versions of the ADP technology, this functionality will be connected to a full trip planner. For this evaluation, the O/D input task is more limited. For the ADP testing, either the origin or the destination needs to be one of multiple venues in the city that can be selected from a prepopulated menu.

After users select the venue, they will either go to, or leave, they then enter the bus route (or transit line) they will arrive/depart on and the direction they will be coming from/going toward. The ADP then identifies the transit stops for the route that serves the selected venue and builds paths to each of those stops from the venue. Travelers then select the stop and path they wish to use. These initial trip planning tasks are used within the evaluation to set up the trips to be studied as part of the evaluation. The IE will not evaluate the performance of these tasks.

Given the transit stop selected by the traveler and the user’s personal mobility profile, the ADP computes a set of paths between the venue and that stop. To perform the required pathfinding and navigation tasks, the system relies on a standardized database of pathway information. That database includes which signalized intersections support the ADP street crossing application and are therefore considered safer street crossing locations, the surface roughness and grade associated with each path segment, the location of curb cuts and wheelchair accessible ramps, and the location of obstacles (e.g., speed bumps, curbs, and parking bollards) along each path. The database also tracks which paths are shared by motor vehicles, and which are reserved for pedestrians.

The path creation process takes into account users' personal mobility profile and provides the user with options allowing them to tradeoff the distance to be traveled with the use of safer street crossings, and their willingness to share right-of-way with vehicles. That is, a user might have the option of choosing a longer path that does not share right-of-way with vehicles (e.g., taking a detour to exclusively use sidewalks) and that uses the safe intersection crossings, or taking a shorter path that passes through a shared right-of-way (e.g., a parking lot or long driveway as the path approaches the venue is those options exist), or crosses a street without a traffic signal. Travelers select the path they wish to follow from the path options offered to them.

Once the path has been selected, the ADP tracks the location of the traveler in real time via global positioning system (GPS) and provides turn-by-turn navigation directions along the selected path. While users follow that path, the system indicates when vehicles may be present, but does not perform real-time moving obstacle detection, either for the street crossing or for the environments where people and vehicles share physical right-of-way.

If the selected path must cross a street with an instrumented signalized intersection, the ADP technology uses GPS to identify when the traveler reaches the intersection and then communicates to the traffic signal system that the traveler has arrived and that it is time to change the length of the traffic signal phase to that is needed by the user to cross the street safely. This communication occurs by having the user's smartphone communicate directly with the signal controller via DSRC. The ADP application indicates the phase to be lengthened, the time required for that phase, and the fact that the user has arrived at the intersection and is ready to cross. The signal controller then changes the signal timing plan to ensure that the user has enough time to cross the street.

The traffic signal system communicates to the smartphone via the DSRC link when the light is green for the user's movement and it is safe to cross the street. The smartphone then communicates this information to the user. The system then monitors the progress of the traveler across the street and further extends the crossing phase if necessary. Once the traveler has crossed the street, the smartphone alerts the signal controller, which releases the phase and transitions back to its previous timing pattern.

Based on the above description of the technical tasks performed by the ADP, the project evaluation will focus on TAL 6 (street crossing) and TAL 7 (mixed environments with moving vehicles and pedestrians), as these activities are at the core of the safety and travel improvements that the ADP is intended to accomplish. The logic models shown for this example only cover these two TALs.

Accessibility Development Projects Outcomes/Stakeholder Priorities

This ADP deployment actually combines two ADPs that were originally developed separately. Therefore, one of the outcomes of interest to the evaluation stakeholders is to understand whether the two systems function seamlessly together. That is, does the pathfinding algorithm correctly identify both the intersections where the street crossing assistance is provided, and pathways that are accessible to the user, given the user's mobility profile?

The first hypothesis to be tested—determining if the instrumented intersections are correctly identified and included in the overall path finding algorithm—is likely best studied in the lab prior to deployment starting. It may be performed by the ADP team as part of that team's initial deployment testing, in which case the results of those tests can simply be provided to the IE team. However, the IE team might also perform these tests. The use of volunteers to enter a large number of different O/D pairs within the deployment area is suggested, to test the performance of the path building algorithm, and the resulting navigation instructions. By examining the ADP application's output to these O/D requests, it is possible to determine if good path options are being provided by the ADP application, as well as whether the instrumented intersections are being correctly identified and used in those paths. Finally, the tests can be used to determine whether the key safety attributes (e.g., presence of the intersection safety system, and presence or lack of vehicles on specific path sections) are described in the path options provided to the users, so that travelers can make informed decisions between alternative paths.

Data collection for this evaluation task would be performed by the individuals performing or supervising the lab tests. Performance metrics should describe the number and percentage of errors discovered in the navigation paths constructed as part of the testing effort. That is, given how many paths are built as part of the testing, the percentage of those paths that are both valid and correctly incorporate the features that should be in the database, and the percentage that have one or more errors.

Logic model entry for these initial topics is shown in table 7 below.

Table 7. Example logic model for a safe intersection and mixed environment accessible transportation technologies research.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
Combined ADP successfully incorporates signal crossing locations into the combined application.	The ADP correctly identifies signals equipped with the street crossing ADP and incorporates that information in the path building.	1. Percentage of navigation paths correctly computed. 2. Percentage of navigation paths that use instrumented signals that identify those intersections. 3. Percentage of paths that should use instrumented signalized intersections that do not use or identify those intersections.	1. Number of paths computed. 2. Number of paths correctly computed. 3. Number of correctly computed paths that contain instrumented signals. 4. Number of paths computed that cross streets, but do not contain instrumented signals. 5. Number of paths computed that incorrectly identify, or fail to identify, instrumented signalized intersections.	Lab test: 1. Manual entry of O/D paths and output of multiple paths for each O/D pair. 2. Ground truth path determinations for each entered O/D pair.	Either the IE team or the ADP team can perform this test. It is designed to check the functioning of the path building database and algorithm, not the actual performance by test subjects traveling that path. Separate paths should be constructed for low/no vision individuals and sighted individuals with other mobility disabilities.
ADP allows users to create a personalized profile that customizes the technology's performance. (*Technical function/System integration.*)	1. Both the safe intersection and path finding software are able to access the same user profile database. 2. The user profile settings affect the time provided for the user to cross the street. 3. The phase time for the crossing intersection is set for each crossing based on the user's profile.	1. Percentage of paths correctly computed given changes in personal profiles. 2. Percentage of devices which correctly compute intersection crossing times given changes in personal profiles. 3. Time allocated by the signal controller to cross the street is equal to or greater than the profile time for that user. 4. Percentage of intersection "ready to cross" arrival times correctly predicted given profile changes, in order to correctly adjust phase timing.	1. Number of paths computed. 2. Number of correct paths computed, given personal profile attributes. 3. Predicted travel times needed to cross intersection for each crossing. 4. Actual travel time required to cross intersection. 5. Predicted arrival time and predicted phase to change (on which cycle is the adjustment to be made?). 6. Actual arrival and required phase. 7. Initial phase time for the cycle phase used for crossing an intersection. 8. Subsequent phase time after crossing request is made. 9. Actual time required to complete crossing. 10. Arrival time at the intersection. 11. Signal phase to be used to cross the street.	1. Lab test of system performance. 2. Field test of system performance: 3. Phase number and phase length from the intersection controller (entire day, by phase time). 4. Phase length from the timing plan. 5. Phase length as set by the ADP. 6. Phase length applied on the phase when the crossing actually takes place. 7. Actual phase length with extensions. 8. ADP app event data.	Signal controller data. Many of these tests may best be performed by the ADP team as part of development testing. The IE team should obtain and summarize this information, as well as carry out the field test which determine if the outcomes (the phase changes occurs after the user arrives at the intersection, and safely crosses the intersection) are as desired. If signal controller data is not available, it may be possible to collect phase length data by direct observation. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.

Source: Federal Highway Administration.

Note that these lab tests simply determine the validity of the base infrastructure map and the path building algorithm. Field tests described shortly will determine whether the database includes all obstructions which need to be included in the infrastructure database.

Travel Activity Links Performance

The next set of evaluation outcomes of importance to the stakeholders are the changes observed in the travel activity outcomes that users experience when given safe intersection crossing and path guidance. That is, do users cross streets more safely, and do they follow safer paths through environments that are used by both pedestrians and vehicles? An important follow-on to the basic travel activity outcomes is whether the availability of the ADP encourages users of the ADP to travel with greater confidence and comfort and thus to allows them to travel more often. Specific evaluation topics of interest to the project stakeholders include the following:

- Are users able to follow the navigation directions?
- How often do users make errors when attempting to follow their selected route?
- Does the pathway database identify all of the obstacles that make travel difficult, or do additional features need to be added to the pathway database?
- Do they use the instrumented signalized intersections more often?
- How often do users select the “best” path versus picking one that is faster but less safe?
- Do the users find value in the system?
- Do users feel safer and more confident in their ability to travel?
- Do users travel more than they used to as a result of their improved perception of safety and ease of travel?

Unlike the initial hypothesis test which examined the ability of the two initial ADPs to function together, this portion of the IE requires field data collection. The hypotheses developed to answer the above travel outcome questions require collection of data that describe how individuals use the system in the field.

The key denominator in many of the performance metrics used to respond to these hypothesis tests is the number of trips actually made. Performance metrics of interest focus on the percentage of those trips that succeed without incident, and the percentage of trips that experience one or more specific types of navigation or safety failures. The “failures” need to be categorized by the cause of the failure, for example differences in the infrastructure database and the actual infrastructure conditions, or incorrect navigation directions, or poorly time delivery of those directions, or an inability of the user to understand the directions. These performance outcomes also support the threat model hypotheses discussed later in this chapter.

Because this project has a multi-faceted target population, it is important that these tests be performed for each of the separate target population subsets. That is, it is crucial that one set of tests be performed for individuals with no vision, another set for individuals with low vision, and another set for individuals with mobility disabilities but good vision. While a summary of results for all of these subpopulations combined is important, understanding how each of these target sub-populations interacts with, and the degree to which they benefit from or have difficulties with, the ADP technology is important.

The data required to test these performance hypotheses are best collected from the ADP technology itself. Thus, the IE team should work directly with the ADP team early on to determine if these data are available through the ADP, or can be collected automatically through the ADP technology itself—and then shared with the IE team—if minor changes to the ADP software are made. For example, in this test, the IE team would need to ask the ADP team if the ADP smartphone application has the ability to capture 1) all of the alternative paths offered to the travelers; 2) the timestamped navigation directions provided for the path the traveler selects; and 3) high-density trace data showing exactly where the traveler actually goes. This set of data allows the IE team to compare the taken against the directions provided, and thus determine the ability/willingness of the user to follow those directions, and where deviations occur from those directions. The trace data also allow the computation of the time spent waiting at an intersection?

If high-resolution traffic signal controller data can also be obtained, it is possible to automatically determine how the traffic signal system interacts with the traveler as they arrive at the intersection. If these data can be captured automatically, the cost of the data collection will be greatly lowered. If not, the IE team will need to budget for staff observation of the field tests, which both raises the cost of the evaluation, and likely decreases the number of person trips that can be performed and tested within the evaluation.

In addition to the physical travel outcomes that need to be measured and evaluated in this portion of the IE, stakeholders are also interested how the individuals using the system perceive the performance of the system. Do they like to use the system? Do they feel safer when using it? More comfortable? To obtain this information, the IE team needs to perform attitude surveys of the test subjects. If the IE team is using the same test subjects used by the ADP team, these surveys will need to be coordinated with any surveys being used by the ADP team, to avoid over burdening the test population.

Table 8 provides a first cut at hypotheses for evaluating the basic performance outcomes for this combined street crossing and path finding ADP. It includes a preliminary set of performance metrics and data sources.

Table 8. Example logic model for evaluating travel activity outcomes for a safe intersection and mixed environment accessible transportation technologies research initiative-funded development projects.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
ADP provides mapping/ navigation to the users? (Technical function.)	Users are able to safely follow the navigation instructions.	1. Percentage of trips completed without navigation error. 2. Percentage of trips completed safely but where one or more navigation errors occurred. 3. Percentage of trips completed but where an unexpected obstacle was encountered. 4. Percentage of trips where an unsafe event occurred. 5. Percentage of trips where navigation assistance was required.	1. Number of trips taken. 2. Number of trips completed without navigation error. 3. Number of trips completed safely but where one or more navigation errors took place. 4. Number of trips where at least one safety issue occurred. 5. Self-reported obstacles encounters. 6. Self-reported safety issues.	Field test: 1. Trace data from the ADP app indicating the path actually taken by the user. 2. Data on the computed navigation path, taken from the ADP app. 3. Post trip survey data from the user indicating if difficulties were encountered, and what those difficulties were.	The IE team, the ADP team, and the sponsor will need to agree on how to define “the correct path.” A key in the field test will be to identify if any obstacles are encountered by test subjects that are not included in the pathway database. These tests need to be performed and reported separately for individuals with low or no sight, versus individuals with mobility disabilities.
ADP provides information users find valuable. (Closing Information gaps.)	The pathfinding technology encourages users to take safer paths (e.g., those with less likely vehicle conflicts—including choosing sidewalks over environments shared with motor vehicles, and intersections with the ADP over intersections without the ADP.)	1. Percentage of trips taken where users select instrumented intersections versus other street crossing locations. 2. Change in percentage of trips using signalized intersections. 3. Percentage of paths selected that are “safer” when faster paths exist. 4. Percentage of trips that take slower paths but that do not share right-of-way with vehicles. 5. Change in percentage of trips avoiding shared use paths.	1. Number of trips taken. 2. Number of trips taken that include instrumented intersections. 3. Number of trips taken that follow slower but safer paths (e.g., include the equipped intersections, pedestrian only environments) versus faster paths that do not include those features. 4. Number of trips taken that follow faster but less safe paths. 5. Number of trips that follow paths without shared vehicle access segments when faster paths exist that share right-of-way. 6. Number of trips taken that use shared right-of-way, when slower paths exist which do not share right-of-way.	Field Test: 1. Trace data from the ADP app indicating the path actually taken by the user. 2. All trip paths offered by the ADP as options for each specific trip to be taken.	If possible, these tests should be performed as a before/after experiment to observe changes in behavior. This assumes that test subjects can be tracked prior to the deployment and that those subjects make a sufficient number of trips. To perform many of these tests, it is also required that the ADP store all of the optional trip paths identified for each trip, so that the IE team can compare the paths selected versus the other options offered.
Users expect to continue to use the ADP after the test concludes. (Empowerment.)	Travelers see value in the use of the ADP.	1. Mean rating of interest in continuing to use the technology.	1. Likert Scale response to question about interest in continuing to use the ADP after the test.	Post field test user survey.	None.

Table 8. Example logic model for evaluating travel activity outcomes for a safe intersection and mixed environment accessible transportation technologies research initiative-funded development projects (continuation).

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
The target population wishes to, and is able to, take more trips to more destinations as a result of the technology. Empowerment/ Technical function. (Outcome.)	Having access to the ADP technology increases the interest/willingness to make trips.	1. Actual or user predicted change in the number of trips taken per week by users before and after deployment of the ADP.	If available both before/after: 1. Number of trips per week. 2. Distance traveled. 3. Number of unique destinations. 4. Answer to post deployment survey question, "Expected number of trips taken per week."	The best approach would be to have measures of trip making behavior, perhaps from GPS traces from user cell phones. If that is not possible, rely on a post deployment survey to measure user perceptions, and user expected ravel changes.	This assumes that it is not possible to perform a before/after analysis of the test population but will rely on a survey of their stated travel likelihood. If a before/after analysis of actual user behavior is possible, that would be the preferred technique.
Users are more confident of their ability to travel safely. (Empowerment.)	Having access to the ADP technology increases confidence in the ability of users to travel safely.	1. Change in mean street crossing time. 2. Change in the mean number of cycles a traveler waits before crossing the street at instrumented intersections. 3. Mean Likert scale response to questions about the users' perception of their level of travel mobility.	1. Time required to cross street. 2. Number of signal cycles user waits before crossing street. 3. Survey question to be answered includes: a. Do you feel confident that you can make trips safely by yourself with this technology?	Field test of user performance: 1. With/without technology tests measuring. a. Crossing time. b. Number of cycles a pedestrian waited until they chose to cross. User surveys after the field test, given to both the users and their caregivers.	Crossing time provides insight into whether the users are calm and confident in their movements or are stressed, and hurrying. The number of cycles the user waits is an indication of their confidence that the road is safe to cross. The survey yields perceptual information that also provides insight into why measured outcomes are occurring.

Source: Federal Highway Administration.

User Needs

The next set of evaluation topics are designed to evaluate the effectiveness of the ADP in meeting identified ATTRI user needs. This includes both the outcome of tasks associated with the technical performance of the system and the ability with which users are able to close information gaps by easily interacting with the ADP technology. Some of these tasks—particularly the usability tests of the ADP technology—will likely have been performed by the ADP team as part of earlier system development work. The IE team would need to obtain those results, summarize, and report them. However, during the field trials that are the source of much of the IE work, limitations in the ability of users to successfully interact with the technology in crowded, noisy, or poor environmental conditions might come to light as the source of technical performance problems that are observed (if any are observed.) Thus, even if detailed usability tests are not within the scope of the IE, the IE team needs to be aware of usability issues which might explain overall performance outcomes.

The initial hypotheses presented below for this example are basic usability tests. They start with the ability of users to personalize the ADP so that it interacts with them in ways that allow them to travel more successfully. The evaluation tests then examine whether these personalizations correctly cascade through the software algorithms so that user preferences and capabilities are reflected not just in the ways the user interacts with the ADP hardware, but that the paths the system computes and the directions given to the users also change to reflect the abilities and preferences of the users.

Consequently, this section examines whether the paths being computed are correctly design for each traveler’s abilities? Can the users follow those paths? Do the intersection crossing times change to reflect the abilities of each user? Many of the outcomes these tests also serve to answer issues that will be raised in the Threat Model analysis, as usability issues can easily result in potential safety issues.

In this example, many of the desired tests require the collection of data from the ADP technology. For example, when testing whether users with different disabilities can easily use the technology, the evaluation tests expect to be able to obtain data on how often a user has to interact with the device. For example, can they hear the audio directions? Do they need to have those directions repeated multiple times? Tracking these outcomes in the field requires either that the ADP technology records these device interactions, or that independent field observations of travelers are made. The former requires that the ADP technology record these interactions. The later requires a considerable increase in the staffing needed to perform the evaluation. The degree to which these hypotheses are tested will thus be a function of availability of that data from the ADP and the budget for the evaluation. Given that this example ADP is designed for multiple, very differently abled target sub-populations, these tests need to be performed and reported for each these distinct sub-populations.

Table 9 provides an initial set of evaluation hypotheses and performance metrics for evaluating the user needs for the combined street crossing and path finding ADP.

Table 9. Example logic model evaluating user needs for a safe intersection and mixed environment Accessible Transportation Technologies Research Initiative-funded development projects.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
ADP allows users to create a personalized profile that customizes the technology's performance. (Technical function/System integration.)	Both the safe intersection and path finding software are able to access the same user profile database. The user profile settings affect the time provided for the user to cross the street. The user profile affects the path selection.	1. Percentage of paths correctly computed given changes in personal profiles. 2. The phase time for the crossing intersection is set for each crossing based on the user's profile. 3. Percentage of devices which correctly compute intersection crossing times given changes in personal profiles. 4. Time allocated by the signal controller to cross the street is equal to or greater than the profile time for that user. 5. Percentage of intersection "ready to cross" arrival times correctly predicted given profile changes, in order to correctly adjust phase timing.	1. Number of paths computed. 2. Number of correct paths computed, given personal profile attributes. 3. Predicted travel times needed to cross intersection for each crossing. 4. Actual travel time required to cross intersection. 5. Predicted arrival time and predicted phase to change (on which cycle is the adjustment to be made?). 6. Actual arrival and required phase. 7. Initial phase time for the cycle phase used for crossing an intersection. 8. Subsequent phase time after crossing request is made. 9. Actual time required to complete crossing. 10. Arrival time at the intersection. 11. Signal phase to be used to cross the street.	Lab test of system performance. Field test of system performance: 1. Phase number and phase length from the intersection controller (entire day, by phase time.) 2. Phase length from the timing plan. 3. Phase length as set by the ADP. 4. Phase length applied on the phase when the crossing actually takes place. 5. Actual phase length with extensions. 6. ADP app event data.	Signal controller data. Many of these tests may best be performed by the ADP team as part of development testing. The IE team should obtain and summarize this information, as well as carry out the field test which determine if the outcomes (the phase changes occurs after the user arrives at the intersection, and safely crosses the intersection) are as desired. If signal controller data is not available, it may be possible to collect phase length data by direct observation. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.
ADP provides mapping/navigation to the users? (Technical function.)	The ADP correctly selects the right path, based on the user's mobility profile and the pathway information, and the pathway information should be based on sidewalk and not road network.	1. Percentage of correct navigation paths computed. 2. Summary of user comments about the quality, efficiency, and safety of the path selected. 3. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.	1. Number of paths computed. 2. Number of correct paths computed. 3. Free text responses to post-trip survey, asking about the quality and safety of the path provided by the ADP.	Field test: 1. Trace data from the ADP app indicating the path actually taken by the user. 2. Data on the computed navigation path, taken from the ADP app. Post trip survey data from the user indicating if difficulties were encountered, and what those difficulties were.	The IE team, the ADP team, and the sponsor will need to agree on how to define "the correct path." A key in the field test will be to identify if any obstacles are encountered by test subjects that are not included in the pathway database.
Technology is usable by the identified user population. (Usability.)	Users are able to easily enter their O/D locations into the ADP technology.	1. Time required to enter the origin/destination of their trip. 2. Percentage of data entries that must be corrected before O/D are correct. 3. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.	1. Start and end time of O/D entry process. 2. Number of entries provided. 3. Number of entries that must be corrected.	This might best be performed as a lab test as part of the system development. Alternatively, a field test might be conducted, although that may require additional software development. 1. Event data from the ADP app.	The ADP technology needs to be capable of recording event data, including when each task starts and ends. The IE team should perform this test, but the cost may require this test to be done as part of the ADP, and those results summarized by the ADP team and given to the IE team.

Table 9. Example logic model for a safe intersection and mixed environment Accessible Transportation Technologies Research Initiative-funded development projects (continuation).

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
Technology is usable by the identified user population. (Usability.)	Users are able to follow the navigation instructions.	1. Percentage of correct crossings taken by user compared to ADP-suggested routes. 2. Percentage of intersection crossings that take place on the correct phase. 3. Percentage of times the user starts crossing within three seconds of the ADP announcement that it is time to cross the street.	1. Number of times the user crosses the correct street during the appropriate phase. 2. Number of times the user crosses the wrong street. 3. Number of times the user crosses the correct street, but at an incorrect time.	Field test: 1. Location and event data from the ADP app, to indicate when street crossings occur. 2. Planned navigation path description. (O/D, and path). 3. Signal system event data, reporting when each phase change occurs.	Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.
The target population can use the technology, even in crowded and noisy conditions. (Usability.)	Users can interact with, hear, read, and understand the ADP technology in a crowded, noisy conditions.	1. Percentage of successful device interactions. 2. Number of safety related events that occur as a result of incorrect device interactions. 3. Percentage of trips with safety events occurring as a result of communication errors. 4. Percentage of navigation instructions that must be repeated. 5. Percentage of trips with navigation errors. 6. Mean value of Likert scale (0 to 5) satisfaction level reported about using the technology.	1. Number and timing of device interactions by type of device interaction. 2. Compute the number of repeated navigation instructions. 3. Number of trips with navigation errors. 4. Number of safety events (e.g., incorrect attempts to cross a street when it was not safe, number of “time to cross” announcements missed by user) which occur due to failed communication. 5. Survey response to satisfaction questions: 6. Ability to hear the device. 7. Ability to read the text (if using that option). 8. 6. Effectiveness of device input (text versus audio).	1. Event data collected from the app. 2. Trace data from the ADP app. 3. Navigation path data to be followed. 4. Post technology deployment survey of test subjects.	IE team must work with the ADP team to ensure that the app data can collect and report event data. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.
Users feel comfortable using the technology in public. (Empowerment.)	Users are comfortable carrying and using the ADP while moving (crossing the street, traveling down a crowded pathway).	1. Mean of Likert scale (0 to 5) satisfaction level reported about using the technology. 2. Results analyzed and reported separately for low/no vision users versus individuals using assistive devices.	Questions to be answered include: 1. Are you comfortable using the phone on a bus or while waiting at a bus stop? 2. Are you comfortable using the device while in motion crossing the street? 3. Are you comfortable using the device while in motion in the mixed environment where vehicles may be present?	1. Post technology deployment survey of test subjects.	The IE team should work with the ADP team and sponsor to develop the appropriately worded questions for the survey, to determine if the users can carry the device, interact with that device (input, receive, and request information) while in public, and if the use of that device is socially acceptable to the user.

Source: Federal Highway Administration.

Threat Model

The next section of the logic model focuses on the threats to traveler safety and how the ADP responds to those threats. Traveler safety is essential in any setting where people cross streets or navigate paths in shared environments. The ADP technology evaluated in this example provides real-time information about the current location of travelers, the relative location of specific infrastructure features, and real-time navigation instructions, including specific instructions about when it is safe to cross streets.

Traveler safety and security will be directly affected if the information about the path selected is inaccurate, notifications of the status of the traffic signal phase are incorrect, or navigation instructions are incomplete or misunderstood. Safety will also be affected if the phase timing changes are insufficient to allow travelers to cross the street or if travelers cross a different street than the ADP technology expects. Safety is also compromised if navigation directions are in error—for example because of an error in a database entry for an infrastructure feature or because of a physical change to the features has occurred without an update to the database (e.g., a construction project blocks a sidewalk)—if navigation directions are followed incorrectly because of poor wording or delivery, or if errors in navigation occur because of the timing of the delivery of those navigation instructions. Therefore, the number of errors caused by incorrect or insufficient database information must be evaluated, as well as the ability of users to understand and follow the navigation instructions.

Because the project is also evaluating the pathway database, if users have trouble navigating their selected path, it is important for this evaluation to identify the causes of those errors, not just that the errors occur. In particular, it is important to determine whether errors occur because specific types of features are missing from the database, because of errors in converting the path to navigation instructions, because the timing of those instructions, or because users simply do not understand or follow the instructions. As part of the evaluation effort, the IE team must be able to identify any new items that should be added to the pathway database.

The following issues need to be included in the evaluation, either through a review of the ADP team test results or by the IE team performing tests as part of the independent evaluation. Important user safety topics include the following:

- Do users more often complete their crossing movements before the signal changes?
- The degree to which collisions and near collisions occur or are avoided on the path between the transit stop and the venue.
- The ability of each user to safely follow the path selected, given each user's capabilities (the safety risk is that the identified path is not safe for that specific individual because it contains features that are beyond the capabilities of that user).
- The ability of users to effectively obtain and understand the navigation instructions.
- The timing of those instructions, to ensure that travelers following those instructions do not make turns or other movements either earlier or later than intended.
- The reliability of the instructions given to indicate when users should cross a signalized intersection.
- The reliability of the traffic signal timing changes (i.e., do users complete their crossings before the signal phase ends?).

It is assumed that considerable detailed analysis of the technical functioning of the combined ADP systems (e.g., the interaction between the ADP application and the signal controller), as well as detailed user experience testing, will be performed by the ADP team before the IE project. Therefore, within the threat model, the IE will only focus on whether users are able to safely and successfully travel between their transit stop and the venue. The IE team will include in their report a summary of the detailed evaluation tests performed by the ADP team, but the evaluation hypotheses shown below as well as earlier will focus on the travel outcomes of the users, and not on either the detailed technical performance of the hardware/software or the effectiveness of the system's interface design. These factors will only be examined in the broader context of the travel outcomes.

Because the ADP does not directly identify moving vehicles, the safety of users is a function of two basic factors, 1) when crossing at an instrumented intersection, users cross during the correct signal phase and that phase remains green until they have crossed the street; and 2) when traveling between transit stop and the venue, they are aware of a) when they are in shared right-of-way; and b) where specific obstacles are located in their path, so that they can identify and avoid them.

The second of these safety concerns are addressed in the evaluation tests presented earlier in this chapter which include safety performance metrics along with those metrics that explore the performance of the path finding algorithm. As a result, the hypotheses shown in the table below examine only the safety aspects of the intersection crossing aspect of the ADP which have not been previously examined. This includes whether users are correctly identified as arriving at an instrumented intersection, that the correct phase has been selected for modification, and the timing for that phase is modified correctly. The analysis then examines whether users successfully cross the intersection before the appropriate phase ends.

Table 10 provides an initial set of evaluation hypotheses and performance metrics for evaluating the threats to users of the combined street crossing and path finding ADP.

Table 10. Example logic model for a safe intersection and mixed environment Accessible Transportation Technologies Research Initiative-funded development projects.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
ADP improves safety and security. (Technical Function.)	The technology correctly identifies that users have arrived at the crossing point and notifies the signal controller.	1. Percentage of times the ADP technology correctly identifies that users have arrived at the safe crossing point. 2. Percentage of times the signal system receives and correctly responds to the arrival notification.	1. Number of trips and intersections crossed by test subjects. 2. Number of times the ADP technology correctly identifies that users have arrived at the safe crossing point. 3. Number of times the signal system receives and correctly responds to the arrival notification.	Field test: 1. Location and event data from the ADP app, to indicate when test subjects arrive at specific intersections. 2. Signal system video, to confirm the location of the test subject. 3. Signal system event data, reporting when each phase change occurs and what each phase length is initially, and if it is extended to assist the user.	Also see the usability tests for the device in crowded and noisy conditions, as important safety tests are performed there. These tests are not repeated in this logic model row, but they are applicable here.
ADP improves safety and security. (Technical Function.)	The ADP correctly computes signal phase times for the user crossing the street.	1. The percentage of street crossings where users do not complete the crossing before the light changes to yellow.	1. Number of street crossings. 2. Number of street crossings where the phase ends prior to the user reaching the sidewalk.	Field test: 1. Location and event data from the ADP app, to indicate when street crossings occur. 2. Signal system video, to confirm the location of the test subject when the light changes. 3. Signal system event data, reporting when each phase change occurs.	The test will require a combination of high-resolution traffic signal controller data, and time stamped video image data (or manual observation of controlled tests) to determine the interaction of the test subjects and the signal system.

Source: Federal Highway Administration.

System Integration

The final set of evaluation topics in which stakeholders are interested involves determining whether the ADP technology effectively connects smartphones and DSRC technology, whether the combined system reliably interacts with the traffic signal system, and because this ADP must directly interact with traffic signal control systems, the resources needed to integrate the ADP into traffic signal systems. These evaluation topics come from examining the System Integration context. They are a key outcome for stakeholder agencies looking to implement these technologies.

In summary, the evaluation needs to explore topics such as:

- How easy is the system to implement?
- What maintenance activities are required to operate the system?
- Can the ADP integrate with all traffic signal systems?

These evaluation topics are designed to understand the overall impact of the ADP on the traffic signal systems to which it will eventually need to be connected. The hypotheses built off these topics are designed to provide the IE team with the information needed by roadway agencies around the country so that they can understand the system integration requirements of the technology, the prerequisites that apply to their signal system before they can consider adopting and deploying the technology, and the resources needed to deploy, operate, and maintain the technology.

The evaluation therefore examines the degree to which standards used by the ADP allow it to be deployed, the degree to which the system (as currently configured) can work with traffic signal systems in use around the country, and the resources required to maintain the system.

One key integration topic from this project that is not being explored is the required integration of the DSRC communications device with the smartphone and the traffic signal controller. This example assumes that this very important, but technical topic is being performed by the ADP team and is outside the scope of the IE.

For this example, it is assumed that the IE team is responsible only for evaluating whether the communication successfully occurs. It does not examine, whether communication failures occur, or how and why those failures occur. (That is, if the signal controller does not change phase length, the IE team does not know if that failure is due to a disconnect between the phone and the DSRC device, a failure of the DSRC communication with the signal controller, or a failure in the traffic signal controller software.) The IE team only cares that the phase length did not change. It is assumed that more detailed technical testing of the system hardware and software which would identify the specific failure point, is the responsibility of the ADP team.

To understand and report on the key integration requirements is another area where considerable coordination is required between the IE team, the ADP team, and in this case, the agency that controls the traffic signal system (or even the company that maintains the traffic signal system being used.)

Specific integration topics for the evaluation are shown in table 11 below.

Table 11. Example logic model for a safe intersection and mixed environment accessible transportation technologies research.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	Data Elements	Data Sources	Comments
Safe intersection crossing software needs to connect to the available signal controllers and central traffic signal systems. (System integration.)	The safe intersection system software follows standard protocols in order to connect to routinely used traffic signal control systems.	1. The percentage of currently marketed traffic signal systems that the ADP works with.	1. The names of signal software systems the ADP can connect to. 2. The ADP uses National Transportation Communications for Intelligent Transportation System Protocol standard calls for interacting with signal controller software. 3. Geographic information system map of signal locations available.	1. Technical specifications for the ADP.	A key evaluation outcome is to understand the degree to which the ADP can be adopted around the country, or whether it only works with specialized traffic signal system software.
ADP technology is affordable by roadway agencies. (System integration.)	ADP technology is affordable by roadway agencies.	1. Deployment cost versus available budget.	1. Number of devices needed per intersection. 2. Cost of devices per intersection. 3. Other system costs.	1. ADP Team supplied information.	As this ADP requires direct connection/ interaction with signal controller hardware, a key evaluation outcome is to identify the costs required to implement such a system.
ADP technology can be maintained by roadway agencies. (System integration.)	ADP technology can be maintained by roadway agencies.	1. Comparison of technical requirements against agency skillset.	1. Technical specifications of ADP technology devices.	1. ADP Team supplied information combined with interviews of agency information technology (IT) staff.	None.
ADP can be easily deployed by agencies. (System integration.)	ADP technology has an ability to import intersection phase maps via a standard interface from agency databases.	1. Intersection layout import process exists for station layout. 2. Standards exist and are used for intersection layout and integration to the signal controller hardware (i.e., phase length changes can be made by the ADP).	1. Technical specifications for details needed for identifying phase numbers for specific intersection movements, and for making and releasing "calls" to signal controllers to change phase length timing.	2. Technical specifications specify an automated intersection layout input format.	As this ADP requires direct connection/ interaction with signal controller hardware and software, a key evaluation outcome is to identify the mechanisms by which the ADP will determine what phase lengths to change, and how those changes will be made within the signal controller software.

Source: Federal Highway Administration.

Finalize the Evaluation Scope

Once the logic model has been developed, the next step is to determine if the resources available to perform the evaluation are sufficient to perform all of the work described in the logic model. If not, either additional resources will be needed, or modifications to the logic model are required.

Perhaps the most important portion of that review as far as this example goes is a review of the availability of the data the IE team requires to produce the performance metrics identified in the logic model. In the example logic models presented above, there are a large number of datasets that are assumed to be available through the ADP application. For example, the GPS traces of all trips made, the navigation directions for those trips, and the alternative navigation paths which the test subjects decided against using.

Without access to these data through the ADP application, the evaluation activities described above would need to be revised. For example, all data on the time required to cross the intersection would need to be collected manually by staff in the field. This would mean that for data collection purposes, test subjects would need to restrict their trip making to the time periods when research staff could monitor their actions. This is a very different testing protocol than simply allowing test subjects to use the ADP technology for 3 months while using the data they automatically generate to quantify those travel activities. Alternatively, if data cannot be collected, (either because they are physically not available or because the budget is insufficient, the logic model, and potential the evaluation outcomes, need to be revisited.

Data availability not only covers the existence of data, but whether the details of the available data meet the needs of the IE team. For example, GPS trace may be available, but the individual data points that make up those traces may not occur frequently enough to compute some required statistics. For example, if the traces occur only once every minute, that level of detail is not accurate enough to determine the time required to cross the street at an intersection. It may also not provide sufficient detail to understand the full path an individual took, or where they left the expected path.

Thus, in this task, the IE team needs to work directly with the ADP team and determine;

- What data are actually available?
- What the meta data are for those data, and whether the details of the current data collection process meet the needs of the initial performance metric computations.
- Whether there is a cost to obtaining those data (e.g., new code must be written by the ADP team.)
- What steps are needed to obtain those data?
- Where logic model hypotheses could be answered by another data collection methodology.

These discussions also allow the IE and ADP teams to coordinate their survey activities. This both reduces the cost to both groups and decreases the survey burden on the project’s test subjects.

If these data are readily available through the ADP application or through work already being performed by the ADP team (e.g., they already have arranged for test subjects and plan on performing a survey of those subjects), the cost of the data can be more readily obtained. If the ADP does not collect data required for the evaluation, the IE team needs to explore available methods for performing those tasks, determine the cost of those data collection efforts, and then determine the relative importance of the different evaluation hypotheses, in order to ensure that the evaluation focuses its resources on what is most important.

For this example, there are several key data items that the logic model expects to obtain. The first is trace data from the ADP. If detailed, high frequency GPS traces are already captured by the user's smartphones, it is possible to compute a large number of the key performance metrics from that raw data (e.g., the time spent waiting at each signalize intersection, the time required to cross each street.) However, if the current ADP technology does not collect these data, the IE team would either need to work with the ADP team to build such an addition to the ADP software, or develop another data collection method.

Some data desired may simply not be possible. For example, the hypotheses examining whether people change their behavior (e.g., do they cross at safe intersections, or use paths that do not share right-of-way with vehicles more often now that they have better information on available paths?) can only be answered with confidence if 1) the IE team can obtain information on the user's path selections prior to the study; and then 2) measure a sufficient number of trips after the deployment of the ADP to determine a statistically significant change in behavior. This may simply not be feasible due to difficulties in collecting data prior to the ADP's deployment as well as difficulties in obtaining a large enough sample size of test subjects that both have the disabilities of interest and that already travel to the venues in the test area. Thus, the inability to obtain some data may prevent some desired hypotheses from being explored.

The other major electronic data source needed for this project is data on the activities of the traffic signal system. High-resolution traffic signal data would allow the determination of when travelers equipped with DSRC devices were detected by the signal system, what changes occurred to the signal timing based on their arrival, when phases started and ended, and why the key crossing phase ended. Having these data (and ensuring that the clocks on the traffic signal system and the phones were synchronized) would allow the IE team to evaluate how the signal system interacted with the test subjects. Without such a data source, some data can be collected manually, but some performance metrics (e.g., the number of cycles through which an individual waited), would need to be estimated.

Another key issue for this evaluation is that the ADP technology is designed to help a very diverse target population. Tests for each subpopulation need to be performed. This will significantly increase the cost of the evaluation, both because of the number of tests which need to be performed within each sub-population in order to have statistical confidence in the evaluation conclusions, and because of the need for a large and diverse group of test subjects which must be recruited and for which test equipment must be procured, distributed, and who's travel then needs to be tracked.

Perform the Evaluation

Once the logic model has been finalized the evaluation needs to be performed. For this example, this task requires a number of tasks to be performed by the IE team. The basic tasks are described in the "Perform the Evaluation" subsection of chapter 3. where specific tasks are needed for this example, they are discussed below. More general instructions can be found in chapter 3.

Obtain the appropriate approvals for handling sensitive data. This would include obtaining Institutional Review Board (IRB) approval. Specific attention would need to be given to how individuals with visual disabilities will be kept safe when crossing streets, if the ADP does not work as intended.

Create a data security plan. This is particularly important since the evaluation expects to collect trace data and individual profiles for test subjects.

Collaboration plan for participating agencies/forms/groups. This project requires considerable coordination between stakeholders. In particular, it is necessary to communicate frequently with the ADP deployment team that is in charge of the field test, but also the development teams that produced the original two ADP systems (the safe intersection crossing and path finding systems), to ensure that questions that arise during testing of the joint software systems can be quickly addressed. The project team also needs to work with the city traffic engineering office, as the IE team needs to obtain, understand, and be able to evaluate detailed traffic signal control data associated in order to evaluate the interaction of the signal system and the ADP. A review of the logic models presented earlier also shows an array of questions that need to be answered for this project that address whether tests performed by the ADP team are sufficient to meet the needs of the independent evaluation. For example, whether tests that demonstrate the ability of the ADP to identify and incorporate into the navigation path finding algorithm the presence of those intersections that are equipped with the safe intersection crossing hardware. If the ADP team's tests are sufficient, then these results need to be provided to the IE team. If the ADP team's tests are insufficient, then these tasks need to be included in the data collection plan.

Determine the sample sizes. The research team will need to work with the project stakeholders to analyze the implications of the test sample sizes. This includes trading off the ability to recruit test subjects, the number of trips each individual can be expected to perform, whether those tests will be done in a controlled fashion (i.e., staged testing) or whether the test subjects will simply use the ADP as part of their day-to-day activities. (In this example, not all test subjects can be expected to walk through the instrumented intersections routinely. Thus, if "natural use" of the system is selected, the time required to obtain a significant number of intersection crossings could be high. Alternatively, if the test subjects were asked to make multiple trips across the intersection within a limited timeframe (e.g., several trip on a given day) the users might become too familiar with the intersection itself, resulting in bias in the measured outcome from the system.) Thus, the stakeholders will need to work together to tradeoff these different testing outcomes, versus the available study budget, and available time for performing the field tests. This information, taken in context of the importance of the various hypotheses being tested and the available budget will be used to select the desired sample size—and testing protocols for each test.

Finalize the data collection plan. To finalize the data collection plan for the project, the IE team will need to work closely with the ADP deployment team and the city traffic engineering office. The most important decisions to be made concern the recruitment of the test subjects, and whether these will be the individuals already working with the ADP deployment team, or need to be recruited directly by the IE team. This decision will then drive how surveys are performed and whether the IE team will ensure that smartphone data collection activities are put in place by the IE team, or by the ADP deployment team working in coordination with the IE team. This decision will also have a considerable impact on the schedule for the evaluation, as the use of the ADP team's test subjects means that the IE team needs to work within the existing plans of the ADP team for their own testing.

The end product of this cooperative effort will be a written understanding of what data are being supplied by each participant in the project, when those data are to be provided, and the details of how those data will be shared. For example, the IE team will receive from the city, a weekly upload of high-resolution traffic signal controller data via a secure ftp site.

Data quality assurance testing. Using the data collection plan as a guide, the IE team will then need to develop and perform quality assurance tests on the data they receive.

Preliminary data analysis. Once the data have passed quality assurance testing, preliminary analyses will need to be performed to confirm the data collection protocols are functioning as intended and that the assumptions made for statistical analysis are valid.

Ongoing project communication. The IE team will hold project status meetings with the project stakeholders once every two weeks.

Once the field test has been completed, the IE team will complete the required data collection, quality assurance testing, and analysis. Based on the results of those analyses, the IE team will develop their project conclusions and write them up.

Perform Gap Analysis

Before developing recommendations based on the conclusions from the evaluation analyses, the IE team needs to perform a gap analysis which explores what trips the ADP technology—as it currently functions (i.e., given the results of the analysis)—can allow to occur completely, where and how it can be deployed, and what enhancements to the technology are needed to more fully expand the number of complete trips which individuals in the target user population can make.

For this example, we will assume that the ADP evaluation shows that the ADP technology can:

- Correctly identify the user arriving at an intersection.
- Change the traffic light phase to a length appropriate for that user.
- Communicate to the user that it is safe to cross the intersection.
- Hold that phase until the user finishes crossing.
- Correctly identify and provide safe navigation paths for the user to follow leading to and from that intersection and their destination (either a venue or transit stop).
- These tasks can be successfully performed by both individuals with low or no vision and individuals with mobility disabilities that require them to use an assistive device, and therefore move slowly.

However, to perform a complete trip, a variety of TALs besides street crossings and moving through mixed environments come into play. Two example trips consisting of multiple TALs will be used to illustrate important gaps that need to be filled in order for the ADP being evaluated to allow a far greater number of trips to be successfully completed by the target population for this ADP.

Trip 1: A wheelchair user needs to travel from their suburban condo to a downtown music venue. The required TALs would be TAL 1 (trip planning), TAL 3 (Identifying the proper exit from their condo), TAL 4 (Entering the sidewalk environment), TAL 5 (using the sidewalk environment), TAL 6 (crossing streets to get to the transit center), TAL 9 (finding the correct stop in the transit center), TAL 3 (entering the bus), TAL 10 (riding the vehicle), TAL 11 (paying for the trip), and TAL 3 and TAL 4 again, to determine the exit stop and exit the vehicle. Only after this list of TALs has been accomplished does the ADP technology take over.

Trip 2: Is the same as Trip 1, except that it involves a low-vision user and a heavy-rail transit trip, and that rail trip includes a transfer at an underground subway station before exiting at a second subway station near the venue, where the ADP takes over.

When examining these example trips and comparing them to the ADP that is being evaluated, the first key gap is in trip planning. For this evaluation, a limited set of origins and destinations has been selected to be part of the deployment test. A simplified user interface to the ADP was developed to facilitate the ADP tests. But the current ADP is only designed to test the street crossing and path finding improvements, it does not contain a fully function trip planning tool. Therefore, even if the ADP was widely deployed, it would be necessary to either connect the ADP to a commonly available, fully functional trip planner (e.g., Google Maps or Open Trip Planner) and that trip planner needs to be capable of planning trips that are accessible to a wheelchair and low-vision user, or build and distribute a full function trip planner with the ADP's capabilities and all of the traditionally available navigation capabilities. Thus, **the first gap identified is the lack of a widely available trip planner that can be used by the target population to help them take advantage of the ADP in the locations where it has been deployed.**

Access to a wide geographic area is why Google Maps has been so successful as a vehicle navigation platform. As a trip planner, it can provide usable navigation instructions to a vehicle driver from just about anywhere they are to anywhere they need to be. A similar, robust trip planning application is needed by the traveler undertaking Trip 1, except they need information not available in Google Maps, and information that is not universally available (the second gap.)

The traveler needs a planner that knows that the ADP technology exists, and which intersections are equipped with it, and also has the detailed path detail required to provide effective wheelchair routing. Thus, **the second gap is a mechanism needs to exist which can populate a standardized database that contains the pathway data included in this ADP, as well as indicate the location of intersections with the instrumented intersection data.** It is likely that such software will need to be built outside of current commercial systems, using open source trip planners (e.g., Open Trip Planner) which are then updated locally as the required pathway infrastructure data is collected. **A third gap is then to perform the required outreach which informs the target population that this functionality exists and is available for their use.** Once these data become commonly available, private sector trip planners can be expected to add these features, resulting in even more widespread use of the ADP functionality.

A fourth gap is identified in the outcome of the evaluation itself. Currently, the safe intersection crossing software only works with the Surtrac Intelligent Traffic Signal Control software. While this is excellent, state-of-the-art traffic signal control software, it has a relatively modest share of the traffic signal control market. Thus, the fourth gap identified with this ADP would be to expand the ADP's ability to work with other traffic signal systems, thus increasing the number of cities which could easily adopt the system without having to change out their traffic signal control system.

The Trip 2 scenario points out a fifth potential gap and that is if the ADP is deployed in a city with underground transit stations which the target population may find difficult to traverse. If low- or no-vision travelers or wheelchair users have difficulty making transfers in underground stations, that difficulty may limit their transit trip making such that they are not able to take advantage of the ADP, because they cannot reach the areas of the city/region where the ADP has been deployed.

This gap, while of fairly significant importance in dense urban areas with underground transit stations, has less impact on the overall ability of the individuals with disabilities to make complete trips elsewhere in the Nation. A wide variety of other gaps of modest importance can be identified, but it from a deployment perspective, the first four gaps are those that will impose the most significant barriers to widespread improvement in trip making ability from the successful deployment of this ADP.

Appendix A. Evaluation Contexts

When evaluators assess a new technology deployment, they can study a variety of attributes. To ensure a thorough evaluation outcome, it is important that the evaluation consistently consider the full breadth of these attributes, ranging from basic technology performance, to the usability of the technology, to the impacts that technology has on travel behavior, to the ability of the technology to integrate with other hardware and software systems.

To help the independent evaluations (IE) team cover these various of topic areas, the evaluation framework divides evaluation subjects into six specific **contexts.** These are six generalized topic areas that need to be considered in designing the evaluation and which, when taken together, ensure that the evaluation produces a complete and cohesive outcome.

These contexts and the performance characteristics they evaluate are as follows:

- **Technical Function:** Does the Accessible Transportation Technologies Research Initiative (ATTRI)-funded accessibility development projects (ADP) technology function according to its design specification?
- **Technological Robustness:** Does the ADP technology have a high level of product quality; is it safe, reliable, durable; and does it preserve users' privacy and security?
- **Usability:** Can the ADP technology be easily and effectively used by the specific subpopulations for which it is intended? In addition, this context also measures the extent to which the technology affects travelers' access to services and/or access to the pedestrian environment that provides the interface between services, as well as whether sufficient customization and flexibility have been built into the system to accommodate heterogeneity within subpopulations.
- **Communication and Closing Information Gaps:** Does the ADP supply information to users that resolves their information gaps?
- **System and Service Integration:** Does the ADP impose requirements or capabilities (e.g., equipment or services) on agencies for the ADP technology to function as intended?
- **User Empowerment and Social Acceptance:** Is the ADP technology designed to support a traveler-empowering social model of disability? That is, does it give them more control over their life and ability to perform activities. Does it include allowing for personalization, increasing user choice, and user control (both actual control and perceived control)? Does the ADP technology facilitate users' expression of intentionality and freedom through the removal of barriers posed by environment, information gaps, user learning curves, adaptation, and adoption, and does the technology provide for equitable use among different traveler subpopulations?

Each of these contexts is described below. In addition to further discussion of the concept behind each topic area, a table is presented for each context that lists directed questions that can be considered when an evaluation is developed for a given ADP technology. These questions, when applicable to that technology, may lead to suggested evaluation hypotheses that can be tested, performance metrics to use in testing those hypotheses, and items that the project sponsor, the IE team, and the ADP team should discuss and agree upon as part of developing that evaluation.

Evaluation Contexts

Technical Function

The main goal of the technical function evaluation context is to address the question, “Is the product functioning according to the ADP design specification?” That is, does the technology work as desired or advertised? It also explores whether the product improves users’ travel efficiency (e.g., decreases travel time) and increases the efficiency with which users complete specific travel activity links (TAL). In the evaluation process, this context may require collecting data about traveler performance before the technology is used, at initial introduction of the technology, and after several weeks of use to determine whether functional metrics (e.g., number of trips taken successfully) for the target traveling population have changed through the application and use of this technology.

Many of the technical function hypotheses within this context are ADP specific. Consequently, they must be drawn from a review of the technology being evaluated, the tasks that users must perform with that technology, the outcome of the tasks the ADP is intended to perform or assist, and the effectiveness of the ADP technology’s ability to address the targeted user needs or overcome targeted barriers to travel. In an example of a robotic assistant helping a blind individual navigate a transit station environment, the technical function questions might include the following:

- Can the users successfully call and/or identify the robot?
- Can the users successfully communicate their intended destination within the station to the robot?
- Can the robot successfully communicate with the traveler?
- Can the robot successfully navigate through the station to the intended destination?
- Can the users successfully follow the robot through the station?
- Do the travelers understand when they have arrived at their intended destination?

Therefore, setting up the evaluation to test the technical function of the ADP requires the IE team to clearly understand the technical tasks that the ADP technology is designed to perform. It is also important for the IE team to understand the expected standards against which the technology’s performance must be judged.

Table 12 provides IE teams, the project sponsors, and the ADP teams with example-directed questions, evaluation hypotheses, and performance metrics for determining technical functions that could/should be evaluated. It also describes specific items they should agree upon as the evaluation is scoped.

Table 12. Directed questions, hypotheses, and performance metrics for examining technological function.

Primary Project Goals and Context	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
For every technology task or function: Does the ADP perform the task or functional goal?	Examples: Users can create accounts. Users can log into their accounts. Users roles are defined, and they cannot modify app attributes outside of their roles. Users are able to modify their user profiles. Users are able to start a new trip plan. Users are able to save a trip plan. Users are able to select a destination.	1. Percentage of computational function completion. 2. Percentage of completion with no errors, crashes or warnings. 3. Percentage of detectable defects (unintentional program responses to a given input or state). 4. Percentage of branch coverage (the percentage of instruction branches that were tested). Functional coverage—The portion of the functionality of the technology that was actually tested.	The granular functions that they want tested routinely. Error tolerance for each type of error in the routine functions. Acceptable level of functional completion. Definitions for major and minor defects.
If the technology incorporates sensors (e.g., cameras, sonar for collision detection, infrared, ultraviolet, temperature, global positioning system (GPS) location, etc.): Does the sensor operate to acceptable accuracy of gold standard, or acceptable difference from current standard best sensor under all TAL conditions, in isolation and when integrated into the complete system?	Examples: GPS location sensor reports user location within X meters in outdoor environments. Heart rate sensor reports user's heart rate within one standard deviation of the Respironics Pulse Oximeter (the industry standard). Short range Bluetooth receiver receives packets from nearby SR Bluetooth transmitters at X% accuracy Y% of the time.	1. Accuracy. 2. Reliability. 3. Tolerance. 4. Repeatability under different environmental and situational travel conditions. 5. Confidence measures.	Acceptable accuracy and tolerance appropriate for the use case. (For example, an app that just needs to know what area of town the traveler is in requires different GPS location tolerance than the app that alerts the user s/he needs to turn right to stay on the chosen path.)
For every technology task or function: Does the ADP perform the task or functional goal?	Examples: Users can create accounts. Users can log into their accounts. Users' roles are defined, and they cannot modify app attributes outside of their roles. Users are able to modify their user profiles. Users are able to start a new trip plan. Users are able to save a trip plan. Users are able to select a destination.	1. Percentage of computational function completion. 2. Percentage of completion with no errors, crashes or warnings. 3. Percentage of detectable defects (unintentional program responses to a given input or state). 4. Percentage of branch coverage (the percentage of instruction branches that were tested). Functional coverage—the portion of the functionality of the technology that was actually tested.	The granular functions that they want tested routinely. Error tolerance for each type of error in the routine functions. Acceptable level of functional completion. Definitions for major and minor defects.
If the technology infers user's condition/ state, or environmental condition/state: Do inference algorithms predict, label, classify, or segment condition or state with sensitivity X and specificity Y? *Do the inference algorithms demonstrate any biases?*	Examples: When balance monitoring is enabled, technology infers when user is not properly balanced with X% specificity and Y% sensitivity, tested on 100 annotated examples both when stopped and in motion. The computer vision algorithm classifies "walk" versus "do not walk" state of pedestrian signals from user's phone imagery correctly with X sensitivity and Y specificity, tested on Z annotated examples from various pedestrian signals. The robotic assistant at the transit station correctly infers the user's destination using user's voice input with X word-level sensitivity and Y word-level specificity, and within D meter location accuracy, tested on input from Z individuals from various gender, age, and ethnic backgrounds. The app infers which is the upcoming bus stop from the user's location and General Transit Feed Specification (GTFS) feeds with X% accuracy and Y false predictions, tested on input data collected from various Z bus trips, collected in different weather conditions and in different areas of the city.	1. Sensitivity (based on gold standard or supervised labeled data). 2. Specificity. 3. Receiver operating characteristic curve. 4. To the best ability, test for bias in particular travel scenarios that specifically affect travelers with disabilities (for example, if inference algorithm is particularly bad at detecting construction sites while travelers with disabilities are disproportionately affected by construction).	What level of sensitivity and specificity is appropriate for the use case? For example, if the technology is using biometrics such as heart rate monitoring to gauge user's discomfort or anxiety, is it sufficient to provide heart rate at +/-5 bpm error? At what levels should the technology respond? In general, IE and ADP teams should discuss acceptable accuracy and tolerance appropriate for the use case. Teams should also discuss what algorithmic biases might specifically affect the particular subpopulation of interest.

Source: Federal Highway Administration.

Technological Robustness

The main goal of the technological robustness context is to address the question, “Is the ADP technology product of high quality, reliable, safe, durable, and does it preserve security and privacy?” ADP technologies are designed to overcome problems, risks, vulnerabilities, or threats that users typically face in day-to-day travel scenarios. As a result, ADP technologies must provide a level of safety and security guarantee, or risk mitigation, to their users. Within this context, the effectiveness of these guarantees and risk mitigation approaches can be examined.

This context allows examination of whether the ADP technology works as intended, or errors in software or failures in hardware cause the technology to fail periodically, and if it fails, how often, under what conditions, and what happens as a result? That is, can users rely on this technology to work as advertised all of the time, and if not, where and when does it not work, and what happens to users when those events occur? This context also includes how data security and privacy are protected by the ADP technology.

An important aspect of developing the evaluation plan is understanding the ways in which the technology can fail its intended users and studying whether and how those potential events and outcomes need to be included in the evaluation project. Possible failure mechanisms include both technical failure modes (e.g., loss of communication, dropping the device on the ground, etc.) and information security associated with modern Internet connected devices (e.g., How does the technology prevent a stranger from getting a secret access code if it is part of wayfinding instructions for a user?).

Therefore, under this context, the IE and ADP teams should work together to develop a threat model. Most ADP teams have thought explicitly about the protections their technology does and does not offer, and a discussion between the IE and ADP teams should enumerate the steps taken by the ADP team in this regard. A more difficult task for both teams is to explore potential threats to users of the technology. This step is important for understanding the security of the technology and the users’ expectations of the new technology. Do the steps taken by the ADP team conform to the expectations of users? If the technology does not meet the security expectations of users, then any other protections it does offer do not matter—users will either choose to not use the tool or will be potentially misled into thinking that they are safe against real and significant threats.

Building a threat model may reveal new safety, security, or privacy features that the developers should consider implementing to better accommodate their users, and will likely result in recommendations for small changes that help users engage with the technology in ways that will produce desired safety, security, and privacy outcomes. However, the task that typically is foremost in the evaluation is determining how the ADP technology communicates with users about security and privacy and the features which ensure that users obtain important security and privacy outcomes (i.e., text, images, and help documentation), along with explicitly setting users’ expectations for privacy and security outcomes.

Table 13 presents a series of directed questions to help IE teams work through this process, along with evaluation hypotheses, performance metrics for examining those hypotheses, and specific items the IE and ADP teams should agree upon as the evaluation is scoped.

Table 13. Directed questions, hypotheses, and performance metrics for examining technological robustness.

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Safety: Is the use of the product or service safe? Does the technology adhere to robust safety measures?	Example: The technology is compliant with a common safety standard.	Pass/fail acceptance Tests; Pass/Fail User Story Test (or Use Case Testing focused on user behavior, not technical implementation).	1. What safety standard the technology adheres to. 2. What qualifies as acceptance testing against that standard. 3. The major safety risks the ADP should address. 4. Agree on one to two User Stories that will be tested by the ADP internally pertaining to the risks identified in (3).
Error rate and error handling: What is the frequency with which this technology fails, requires repair, or requires maintenance?	Example: The technology can run for at least two hours without crashing. The technology can run well with many users logged on.	Performance test pass/failure: 1. Run tests to identify breaking points and confirm that breaks are handled gracefully. 2. Time technology runs before it crashes or errs out. 3. Verify technology runs with no errors, crashes, or warnings in edge cases. 4. Percentage of detectable defects.	This level of testing might be deemed outside the scope of the IE, but the team should agree on this—in particular, in scenarios (such as safe intersection crossing technologies) in which poor performance puts users in danger.
Failure modes: Every technology requires complete-system level verification, in particular to ensure that there are comprehensive failure modes to prevent accidents or injuries. (Note that there is some dependence on the TALs for which the ADP is designed to function.) Does the technology work reliably end-to-end in edge cases and edge environments?	Examples: The wayfinding app is able to tell users that they need to call for assistance and/or to recover from GPS location services being turned off or display consistently low confidence readings. Travelers are not left unaided and with no ability to call for assistance in the middle of a high-risk travel environment.	1. Technology runs to completion. 2. Functional coverage—What portion of the functionality of the technology was actually tested in the integration test?	Decide what broader components of the technology will be tested in the IE. Decide on are the extreme but real environments in which the technology would need to be system tested. (Some environmental examples are not obvious, such as poor satellite coverage areas.) The ability to test things in real environments and see how the technology is integrated into the actual scenario, is very important.
If the technology gathers or generates data, and in particular if the technology is connected via communication protocols (ex: Bluetooth, Wi-Fi, etc.) to other devices: Does the technology maintain adequate levels of data privacy and security?	Examples: All networked paths protect data through a firewall. All data are transmitted using privacy-preserving data exchange techniques, such as differential privacy.	Use accepted security evaluation testing. This is a heuristic evaluation on the: 1. Sufficiency of the security protocol used by the technology. 2. Whether the protocol is executed properly.	Acceptable security evaluation test framework. Many such frameworks exist. The teams need to agree on: 1. Whether the security protocol used by the technology is sufficient. 2. Whether use of the protocol is executed properly. 3. Which testing method is relevant in the context of the ADP technology.
The technology should have built-in factors that mitigate risks identified by a technology threat model: Is the technology offering to prevent the stated risk as part of strategic and operational planning?	Examples: Users are successfully prevented from talking to or following strangers by in-app reminders about previous training, notifications, and offering of direct communication access to trusted caregivers. The cane beeps while the user is crossing the road and will continue until the user has reached the other side.	1. Qualitative data about the level of risk mitigation matching user expectations. 2. Pass/fail of specific tests for risk mitigation (User Story Tests).	Agreement on user stories that are relevant and testable.
Mitigating risks identified by the threat model: Is the technology designed to detect when the risk or threat actually occurred?	Example: When the technology infers the user is off-balance, the appropriate parties are contacted.	Risk mitigation pass/fail tests.	Agreement on user stories that are relevant and testable.

Table 13. Directed questions, hypotheses, and performance metrics for examining technological robustness (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Technology should behave acceptably vis-a-vis "nonfunctional requirements." Does the technology pass general performance testing? Does it pass stress load tests? If the technology is a phone application, there is a need to verify that the technology does not exceed reasonable memory usage, central processing unit (CPU) usage, network data usage, and battery usage, hence rendering the phone unusable for other purposes, or for very short trips because of battery drain.	Example: The technology does not take unreasonably long to respond, especially for just-in-time demands (e.g., determining whether the crosswalk light is on or off requires low latency so that pedestrian does not walk into moving traffic).	**Response time:** Total time to send a request and get a response. **Wait time:** Also known as average latency—How long it takes to receive the first byte after a request is sent. **Average load time:** The average amount of time it takes to deliver every request—a major indicator of quality from a user's perspective. **Peak response time:** The measurement of the longest amount of time it takes to fulfill a request; slow peak response may indicate an anomaly that will create problems. **Error rate:** The percentage of requests resulting in errors in comparison to all requests. **Concurrent users:** How many active users at any point, the most common measure of load. **Requests per second:** How many requests are handled. **Transactions passed/failed:** The total number of successful or unsuccessful requests. **Throughput:** Kilobytes per second, shows the amount of bandwidth used during the test. **CPU utilization:** The time the CPU needs to process requests. **Tracked outliers:** Extreme measurements reveal possible failures. **Memory utilization:** How much memory is needed to process the request. **Calculate averages to arrive at actionable metrics.**	There are many tests that developers may perform. The IE should be involved in understanding how the results will affect users, not just test environment servers. The IE should discuss the following with the ADP team: 1. Involving developers, information technology (IT), and test subjects early in creating an adequate performance testing environment. 2. Developing a test environment that takes into account as much user activity as possible. (Performance tests are best conducted in test environments that are as close to the production systems as possible and limited resources may restrict choice even further). 3. Determine baseline measurements that provide a starting point for determining success or failure. 4. Consider the audience when preparing reports that share performance testing findings.

Source: Federal Highway Administration.

Usability

The main goal of the usability evaluation is to address the basic questions, “Can the ADP technology be efficiently, effectively, and conveniently used by specific subpopulations of interest? If so, which ones? Are there subpopulations that cannot use the technology but that would benefit from what the technology provides?”

Assessment in this context measures whether—for the populations for which the device is intended—the technology is comfortable, attractive, and easy and enjoyable to use. This context also examines the ability of users to customize the technology so that they can more easily interact with it and obtain its intended benefits. Assessment in this context also measures the extent to which the technology impacts travelers’ access to services and/or access to technology. Evaluation in this context will likely be the core of joint work by the IE and ADP teams, with agreement needed between those teams on the performance metrics to be used and the extent of testing that will be performed by each team.

The evaluation tests are approached by asking a number of questions that drive the selection of the evaluation hypotheses, as shown in table 8. This table provides key questions and directions that will assist the IE team in the creation of evaluation hypotheses for this context, as well as insights into the performance measures that can be used to evaluate those hypotheses.

Table 14. Directed questions, hypotheses, and performance metrics for examining usability.

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Matching real world: Does the system speak the users' language, with form factors, words, phrases, and concepts familiar to users, rather than system-oriented terms? Does the system follow real-world conventions, making information appear in a natural and logical order?	Examples: The button order in an interface matches the order in the ticket machine that the user may have used in the past. Landmark representation is identifiable and recognizable. The technology clearly communicates (visually, audibly, or via tactile response) identifiable information about a path or an environment. All programmed travel instructions (for example, “not your stop” or “now is the time to get off”) are understood effectively by users in this subpopulation.	Percentage correct: Users can answer simple questions about a path represented in the path visualization. This should be measured prospectively (TAL 1) or on-location (TAL 2). (Example questions: 1. Is the portion of the path represented in the app a sidewalk or a crossing? 2. At this point in your trip, does the app represent the same street you are on? 3. Can you match the landmarks in the app to the environment?)	1. What real-world state should users be aware of? 2. What would qualify as acceptance testing against that state? 3. What are the major user abilities and use cases (e.g., noisy environments, high intensity lighting, etc.) that the ADP should address when communicating world state to users? 4. IE and ADP agree on what constitutes in-lab heuristic user testing.
Consistency and adherence to design and platform standards and guidelines: Does the system represent information in a way that is familiar to users and does not defy conventions or expectations? Does the technology leverage users’ knowledge from other technologies or world interactions?	Examples: The technology does not force the user to learn new representations or tool sets for all or each task. When things mean the same or perform the same operation, the technology represents them in the same way.	1. Percentage of views that follow platform conventions. 2. Percentage of error messages that follow platform conventions. 3. Percentage of buttons that follow platform conventions. 4. Percentage of errors made by users that require backtracking to previous state or previous screen. 5. Perform a qualitative “talk-through” analysis of users’ first few times working with device. Note the number of interactions with the technology when performing directed tasks (not explorations) that users wish to backtrack.	IE and ADP must agree to what extent this population has seen a similar technology. IE and ADP must agree on what comparable technology or platform standards (if any) to use as a baseline for this technology. IE and ADP must agree on how much of this task-based analysis would be done with actual users in field testing.
Error prevention: Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state. Does this exit require users to go through an extended dialogue? Does the technology support *undo* and *redo*?	Examples: The technology supports users to undo their last click, choice or move. The technology supports backtracking with a prominent placement, so it is easy to detect.	1. Percentage of user interactions that can be undone (even if takes several actions). 2. Percentage of user interactions that can be undone with a single action, in a visible way. 3. Percentage of user interactions that request users’ verification before committing to an action.	IE and ADP have to agree on what portion of this will be purely evaluated with in-house testing versus in the field. IE and ADP would have to agree on which user interactions require confirmation with users before committing to an action with the technology (for example, “delete all prior planned trips” would require user confirmation).
Error recognition, not recall: Does the technology make backtracking options, objects, or actions clearly visible? Does the technology assume users remember something from a previous dialog or state of the technology? Are instructions for use easily accessible when they are needed?	Examples: The technology minimizes users’ memory load by making interactions clearly accessible. Each interaction with the technology stands on its own, not assuming user background or memory of previous interactions or state. Each screen contains instructions, or access to instructions for use of that screen.	Each system state that requires user input has sufficient information for users to provide informed interaction with no assumed knowledge of system state or technology-specific language. **This will typically be evaluated with heuristic evaluation, and in rare cases, some field testing.*	IE and ADP should agree on what the evaluated system states will be. For example, in a mobile app, which user screens are considered unique for evaluation purposes?
Error reporting, diagnosis and recovery: Does the system handle errors gracefully and in ways that are understood by users?	Example: The technology expresses/displays error messages in plain language, not codes, that clearly indicate the problem and offer constructive solutions.	Every system error provides simple clear messaging, sufficient explanation of the problem, and suggests a means of exiting or a solution. **This will typically be evaluated with heuristic evaluation.*	ADP should expose to IE all error states and any associated error handling available on the system.

Table 14. Directed questions, hypotheses, and performance metrics for examining usability (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Keep it simple: Does the system offer a simple aesthetic and minimalist design?	Example: The technology offers clear dialogues with no extraneous information or visual distractions.	Every system dialogue provides simple messaging and sufficient explanation of relevant information. Every system state will be evaluated for extra units of information that may diminish from the relative visibility of relevant information. **This will typically be evaluated with heuristic evaluation.*	Agreement on the system state or screens that will be under evaluation.
Documentation: Help and documentation should be available for those who require it. Is the documentation method sustainable so that the ADP can maintain it and it enables continued user support as the system matures?	Examples: There is a clear, identifiable way to find help and documentation when using the technology. System documentation is maintained in a sustainable, easily manageable way.	Number of system functions that are found to not be defined in the help/documentation. Documentation of the system function is focused on users' tasks. (This is a binary yes/no evaluation.) Documentation and help narrative provide concrete steps that need to be carried out by the user. (This is a binary yes/no evaluation.) Documentation and help narrative are not long or confusing. (This is a heuristic evaluation.) Documentation can and is easily maintained by the ADP. (This is a heuristic evaluation.) **This will typically be evaluated with heuristic evaluation.*	ADP should provide its system documentation for evaluation by the IE.

Source: Federal Highway Administration.

There are two instances in which it is important to consider more than a heuristic expert usability test: (1) if the population of interest involves people with cognitive disabilities; and (2) if funds are sufficient to run a user study for assessing usability. In these cases, the usability field experiment should examine all of the above and add the questions and hypotheses presented in table 15.

Table 15. Directed questions, hypotheses, and performance metrics for additional usability field tests.

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Intuitive design: Is the design simple to understand?	Example: The ADP technology offers a nearly effortless understanding of the layout (physical or digital) and navigation of the technology.	User testing of the technology experience and design. Task-based user study collects analytics in which users are asked to perform system-specific subtasks. Metrics: 1. Binary (yes/no) user ability to achieve subtask successfully with no tester intervention. 2. Time to perform a task the first time. 3. Time to perform a task the first time / average time to perform the task on repeated instances. 4. Number and type of interaction errors made en route to performing the subtask successfully. 5. Five-point Likert scale rating by users about the ease of use of the specific subtask. 6. Overall: Percentage of users who navigate the interface with no need for intervention or documentation.	IE and ADP should agree on the tasks that users will be asked to perform with the technology to test simplicity.
Ease of learning.	Example: The technology facilitates a user who has never seen the technology before to accomplish basic tasks.	User testing of the technology experience and design. Metrics: 1. Time to perform a task the first time. 2. Time to perform a task the first time/average time to perform the task on repeated instances. 3. Percentage of uses before a user performs agreed upon tasks with no need for backtracking, intervention, or documentation.	IE and ADP should agree on the tasks that users will be asked to perform with the technology to test ease of learning.
Efficiency of use: How fast can experienced users accomplish tasks?	Example: The technology is intuitive and fast for experienced users.	Experienced user testing of the technology. Metrics: 1. Time to perform a task the first time. 2. Time to perform a task the first time / average time to perform the task on repeated instances. 3. Time to perform tasks with the technology past the first visit.	IE and ADP should agree on the tasks that users will be asked to perform with the technology to test efficiency of use.
Memorability: After visiting the site, can users remember enough to use it effectively in future visits?	Example: The technology is easy to remember.	Experienced user testing of the technology. Metrics: 1. Time to perform a task the first time. 2. Time to perform a task the first time/average time to perform the task on repeated instances. 3. Time to perform tasks with the technology past the first visit. Binary (yes/no) success of performing the task another time at the end of the user session without tester intervention or documentation.	IE and ADP should agree on the tasks the users will be asked to perform with the technology to test efficiency of use.

Table 15. Directed questions, hypotheses, and performance metrics for additional usability field tests (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Interface use error frequency and severity: How often do users make errors while using the system? How serious are the errors? How do users recover from the errors?	Example: The technology offers usable error recovery.	User testing of the technology experience and design. Task-based user study will collect analytics in which users are asked to perform system-specific subtasks. Metrics: 1. How often do users make errors while using the system? 2. Segregate errors into sub classifications of errors and count by sub classification. 3. Does the system support adequate recovery from the error (binary yes/no)? 4. Percentage of errors that can be recovered from gracefully. 5. Average time to recover from errors.	IE and ADP should agree on the tasks and scenarios in which users will be asked to use the technology.
Subjective satisfaction: Does the user like using the technology?	Example: The system increases user delight during travel.	Five-point Likert scale rating of: 1. Satisfaction of performing each subtask. 2. Satisfaction of using the entire system. 3. Likelihood user will recommend use of the system to a friend.	IE and ADP should agree on the tasks and potential scenarios users will be asked to perform with the technology to test system satisfaction.

Source: Federal Highway Administration.

Communication and Closing Information Gaps

This context examines whether the ADP technology can effectively communicate with the user population. Successful communication includes allowing the user both to request specific information when and where it is needed and to receive a response to those requests or receive other needed information. A key outcome of the questions asked under this context is to determine whether the information being provided to users actually closes their information gaps. Therefore, the evaluation tests performed within this context examine whether users' information needs are successfully met by the ADP technology, whether those information needs are "Where is the bus stop?" or "Where am I now, and what direction do I need to walk in to reach my destination?"

In summary, the evaluation tests under this context look at whether users:

- Can request information.
- Can perceive, process, and comprehend the responses to their requests.
- Receive sufficiently detailed information that actually resolves their information needs related to the travel environment or travel functions (e.g., how to buy a ticket) to facilitate effective travel.

This context is a particularly important part of the evaluation process for ADP technologies that are developing data standards, as it is within this context that gaps in the data standards will be exposed. It is within this task that any differences between the information users need (by population and TAL) and the information included in the data standard will be most directly exposed. Use of the TALs will also allow the IE team to examine whether, given the environment (TAL) in which information is being requested or needed, users are able to provide the inputs required and manipulate the technology in the manner required to accomplish the tasks necessary for that TAL.

Table 16 provides key questions and directions that will assist the IE team in creating the evaluation hypotheses for this context, as well as insights into the performance measures that can be used to evaluate those hypotheses. It also describes specific items they should agree upon in scoping the evaluation. These questions should be used with respect to particular TALs and subpopulations when forming the evaluation hypotheses.

Table 16. Directed questions, hypotheses, and performance metrics for communication and closing information gaps.

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Visibility: Is the system keeping the user subpopulation properly informed about what is going on in the travel environment? Is the type of feedback appropriate (addressing the visibility of system status)? Is the feedback provided in an appropriate manner (evaluating the use of feedback, such as the timeliness of it)? Does the system always keep users informed about what is going on, through appropriate feedback within a reasonable time? (e.g., Did the system actually catch the button press or was it ignored? Or did my request go through?)	Examples: The technology has appropriate feedback about system status. The technology provides feedback within reasonable time.	1. Percentage of user actions (click, push, etc.) that report back: 2. Action was received and not ignored. 3. Feedback or system status demonstrates that system response occurred. 4. Number of milliseconds it takes to receive system feedback is appropriate and reasonable. *This is not only testing that the **system state** is visible end-to-end but does so taking into account user abilities and different TAL scenarios.	IE and ADP must agree on: 1. The system status that users should be aware of in each TAL. 2. What would qualify as acceptance testing against that visibility. The major user abilities and use cases in different TALs (noisy environments, high intensity lighting, etc.) that the ADP should address when communicating system status to users.
Comprehensibility: Is the level of detail of information being communicated appropriate for each task?	Example: The technology allows informative exploration of an environment rather than just remaining on a predetermined path (i.e., allows user to probe information about the environment representing the path, either prospectively, without having to be present in the current on-location environment).	The ATTRI Needs Assessment document offers detail on the type of information each subpopulation of interest requires (see table 1). For each subpopulation included in the ADP analysis, provide a five-point Likert scale assessment by the heuristic evaluation team for each line item under the "information" category in table 1. In particular, the evaluation must adequately test that the appropriate format of the information is communicated adequately for the subpopulation, and also test that format against each TAL to ensure it is appropriately perceived in each travel environment.	IE and ADP must agree on the overall ADP system goals or travel use cases that need to be addressed for each subpopulation of interest.

Source: Federal Highway Administration.

Table 16. Directed questions, hypotheses, and performance metrics for communication and closing information gaps (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Visual elements: Is variability in users' sight properly addressed? Does the technology provide a voice-reader-accessible text alternative to all communicated information? Does the technology provide high contrast text or alternative text to all media content and messaging? Does the technology provide sufficiently detailed information to close the information gaps in the complete travel chain for people with no or low vision?	Example: The technology properly addresses the needs of individuals with no or low vision.	1. Percentage of messages, buttons, links, information media, including pictures and video, that are not alternatively tagged with text. 2. Percentage of messages, buttons, links, information media, including pictures and video, that are alternatively tagged with inappropriate or overly general text. 3. Percentage of messages, buttons, links, information media, including pictures and video, that do not have high-contrast or high-contrast alternative views. 4. Percentage of messages, buttons, links, information media, including pictures and video, that do not have magnified alternative views. 5. In-lab assessment detailing whether the technology is fully accessible to a voice-over reader. 6. In-lab assessment detailing whether the technology is accessible to a braille reader. In an ideal IE, field tests would collect data on how often and how successfully information is consumed and acted upon by users with no/low vision (e.g., number of times information was accessed successfully, number of times information was accessed, but not used in practice in the travel demonstration, number times a misinformed choice was made—for example, number of times a path with no accessible pedestrian signals crossing was chosen when an alternative existed). 7. For each TAL associated with this ADP, identify information that persons with no or low vision require in that environment and assess whether the ADP supplies this information (yes/no) and to what level of detail (assessed by expert on a five-point Likert scale). 8. All of the following metrics will be assessed on a five-point Likert scale evaluation, assessed by the heuristic evaluation team: 9. The field of view of the subpopulation of interest is met adequately; highest marks for customizability. 10. The color contrast of the subpopulation of interest is met adequately; highest marks for customizability. The acuity of the subpopulation of interest is met adequately; highest marks for customizability.	IE and ADP teams must agree on the subpopulations addressed by the ADP.
Hearing: Is variability in users' hearing properly addressed? Is information about noise properly communicated to address different noise sensitivities? Is information about noise properly communicated to address varying types of hearing loss? Does the technology provide no-hearing-accessible text or image alternative to all media content and messaging? Does the technology provide sufficiently detailed information that closes the information gaps in the complete travel chain for people with no or low hearing?	Examples: The technology properly addresses the needs of individuals who are deaf or limited in hearing. The technology informs people about noisy environments. The technology communicates detailed information about background noise in public places to allow individuals with sensitivity to background noise make more informed travel decisions.	1. Percentage of audio messages, and information media with sound, that are not alternatively captioned with text. 2. Percentage of audio messages, and information media with sound, that are not alternatively captioned with inappropriate or overly general text. 3. In-lab assessment detailing whether the technology is fully accessible with only captioning. In an ideal IE, field-tests would collect data on how often and how successfully information is consumed and acted upon by users with no/low hearing: 4. Number of times information was accessed successfully, number of times information was accessed, but not used in practice in the travel demonstration. 5. Number of times a misinformed choice was made (for example: the number of times an establishment with no closed-loop captioning was chosen when an alternative existed). 6. For each TAL associated with this ADP, identify information that persons with no or low hearing require in the TAL environment and assess whether the ADP supplies this information (yes/no) and to what level of detail (assessed by an expert on a five-point Likert scale). 7. For each TAL associated with this ADP, identify information that persons with noise sensitivities require in the TAL environment and assess whether the ADP supplies this information (yes/no) and to what level of detail it is provided (assessed by an expert on a five-point Likert scale). 8. All of the following metrics will be assessed on a five-point Likert scale evaluation, assessed by the heuristic evaluation team: 9. Sound detection of the subpopulation of interest is met adequately. 10. Speech discrimination of the subpopulation of interest is met adequately. (addressing background levels of different noises, etc.). 11. Localization of the subpopulation of interest is met adequately; highest marks for customizability.	IE and ADP teams must agree on the subpopulations addressed by the ADP.

Table 16. Directed questions, hypotheses, and performance metrics for communication and closing information gaps (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Cognition: Is variability in users' cognition properly addressed? Is information communicated as simply as possible? Is information accessible and easy to understand? Is enough information communicated to inform safe and comfortable travel? Does the technology provide a simple text or image alternative to all media content and messaging? Does the technology provide sufficiently detailed information that closes the information gaps in the complete travel chain for people with cognitive limitations?	Examples: The technology properly addresses the needs of individuals with various cognitive limitations. The technology informs people about high sensory environments. The technology communicates information simply and effectively allowing people with limited cognition to make more informed travel decisions.	1. Percentage of messages, and information media, that pass readability thresholds. (e.g., 90–100 on Flesch Reading Ease score, or <8 on Gunning Fog index). Other indices that may be used: Flesch-Kincaid Grade Level, Coleman-Liau Index, SMOG Index, Automated Readability Index, or Linsear Write Formula. 2. In-lab assessment detailing whether the technology is fully accessible to the target population. 3. Readability factors to measure that the technology works within the attention capabilities of the population of interest and that users recognize and remember what they have read: a. Comprehension. b. Speed of perception. c. Perceptibility at a distance. d. Perceptibility in peripheral vision. e. Visibility. f. Reflex blink technique. g. Rate of work (reading speed). h. Eye movements. i. Fatigue in reading. 4. For each TAL associated with this ADP, identify information that persons with limited cognition require in the TAL environment and assess whether the ADP supplies this information (yes/no) and to what level of detail (the nine points above assessed by an expert on a five-point Likert scale). 5. In an ideal IE, field-tests would collect data on how often and how successfully information is consumed and acted upon by users with low cognition (e.g., number of times information was accessed successfully, number of times information was accessed, but not used in practice in the travel demonstration). 6. Number of times a misinformed choice was made (example: number of times the participant off-boarded at the wrong stop).	IE and ADP must agree on the readability thresholds that must be met for the technology to be accessible to the intended subpopulation. For reference, some states have explicit requirements that legal documents and health care documents must meet. Four online resources are: https://readable.io https://datayze.com/readability-analyzer.php https://www.online-utility.org/english/readability_test_and_improve.jsp http://www.readabilityformulas.com/free-readability-formula-tests.php Example thresholds: Messages score 100.00–90.00 on the Flesch Reading Ease Score (fifth grade level). Gunning Fog index less than 8 for universal understanding.
Mobility: Is variability in users' mobility properly addressed? Are there assumptions about dexterity, strength, flexibility, reach, and traversal that are implicit to the manipulation of and communication with the technology?	Examples: The technology is usable without requiring use of hands in a skillful, coordinated way to grasp and manipulate objects. The technology is usable without demonstrating small, precise movements.	1. Percentage of technology interactions and communication that are accessible within a reasonable time threshold for people who scores within some range appropriate for the population of interest on the Nine-Hole Pin Test.	IE and ADP teams must agree on whether mobility is an ADP feature that requires testing and what the functional range of the population of interest is. (For example, normative for elderly population is different than for minimally functional post-hand trauma population.) In general, use guidelines as outlined by Aaron et al., 2003.[1]

Table 16. Directed questions, hypotheses, and performance metrics for communication and closing information gaps (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Tactility/touch: Is variability in users' proprioception properly addressed?	Example: If tactile representations are used, the technology properly addresses the variability among individuals who require tangible interfaces.	This will have to be studied by the ADP team and results presented to the IE team for approval. If tactility represents information, variability in users' proprioception is properly addressed.	IE and ADP teams must agree on whether tactility is an ADP feature that requires testing.
Irritants: Are crowdedness, pollen, air-, water- or noise-pollution adequately communicated?	Example: The technology communicates about crowded locations to avoid adverse reactions by the population of interest.	All of the following metrics will be assessed on a five-point Likert scale evaluation, assessed by the heuristic evaluation team: **Crowdedness** information is provided to the user with granularity sufficient to be informative to the subpopulation of interest. **Pollution level** information is provided to the user with granularity sufficient to be informative to the subpopulation of interest. **Water toxicity** information is provided to the user with granularity sufficient to be informative to the subpopulation of interest. **Air toxins** information is provided to the user with granularity sufficient to be informative to the subpopulation of interest. Highest marks for customizable interfaces that take into account personal requirements/sensitivities.	IE and ADP teams must agree whether sensitivities in general, and which irritants in particular, are ADP features that require testing.

Source: Federal Highway Administration.

[1] See table 17.

Table 17. Guidelines for determining the time threshold for the nine-hole pin test.

Score by Functional Level	Dominant Injured Hand (sec) Range	Dominant Injured Hand (sec) Mean	Nondominant Injured Hand (sec) Range	Nondominant Injured Hand (sec) Mean
Functional	16–25	21	18–27	21
Moderately Functional	26–33	29	28–45	43
Minimally Functional	34–50	38	46–55	43
Nonfunctional	>50	93	>55	102

Source: Federal Highway Administration, 2003.

System and Service Integration

This context is concerned with the steps in the travel chain, which agencies are potentially affected by the ADP technology, and how agency stakeholders (other riders, operators, management, etc.)are affected by the technology. Topics to be examined for inclusion in the evaluation within this context include the following:

- What types of inspection/regulation of the ADP technology are required by agencies?
- What is the agency-related cost/labor/time burden for the required maintenance and service of the ADP technology?
- What kinds of expertise and training are required for agency-workers to adequately service and maintain the ADP technology?
- What is the environmental impact of the ADP technology (e.g., measurable trash production or vehicle efficiency improvements)?

These topics are important in that they describe the resources required to implement the ADP technology, provide users access to that technology, and sustain the operation of the technology over the long term.

Table 13 provides key questions and directions that will assist the IE team in the creation of evaluation hypotheses for this context, as well as insights into the performance measures that can be used to evaluate those hypotheses. It also describes specific items they should agree upon in scoping the evaluation. These questions should be used with respect to particular TALs and subpopulations in forming the evaluation hypotheses.

Table 18. Directed questions, hypotheses, and performance metrics for system and service integration.

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Travel chain impact: What steps in the travel chain are potentially affected? How are agencies affected by the technology?	Examples: The technology affects every bus onboarding along the trip. The technology affects every pay-point along the trip.	Is there a dissemination plan in-place? With adequate resources, the expert report should address the following questions: 1. What are potential efficiencies of, inefficiencies of, or impacts to sustainable service and other implications of integrating this technology into the full travel chain? a. Does technology increase safety or reduce the likelihood of collision, thereby reducing agencies' liability and other collision costs? b. What measurable costs are impacted? 2. What type of inspection or regulation are required by agencies? a. What is the agency-related cost/labor/time burden of maintenance and service? b. What kind of expertise and training are required for agency workers to adequately service and maintain the product? c. What labor impacts are implied by integrating this technology? d. Do agencies realize any operations or maintenance savings with this technology? e. Does supporting this technology require any agency purchases or retrofitting of facilities and equipment? 3. Does this technology affect agency workforces, which may result in shifting duties away from vehicle operation? 4. Other engagement impacts: a. Does this technology encourage novel engagement models between agency and traveler? b. Does adoption of the technology by an agency measurably affect equity of its service delivery? c. Are agency partnerships required to support this technology? Does adoption of this technology encourage better agency integration over the complete travel chain? (For example, many on-demand services are offered conditionally at nighttime for low lighting conditions. If the technology improves nighttime ridership for people with visual disabilities, this may confer cost effective improvements by shifting some trips from the on-demand system to scheduled route services.)	IE and ADP teams must agree on: 1. What travel chain impacts may occur within each TAL. Whether inter-agency coordination is necessary with respect to other stakeholders (e.g., other riders, operators, managers, customer support, etc.)

Table 18. Directed questions, hypotheses, and performance metrics for system and service integration (continuation).

Directed Questions	Evaluation Hypotheses	Performance Metrics	IE and ADP Teams Need to Agree On
Burden of use: What is the burden of use on users and the users' support teams and caregivers?	Example: The technology requires an anticipated number of hours of set-up and support to be usable.	Experienced user reports. Metrics: 1. Type and time of technical service and preventive maintenance required over the trial period. 2. Type and time of required maintenance required by users or their team (example: calibration, set-up, and remembering to charge). 3. More detailed type and time of daily, weekly, or monthly maintenance activities required. 4. Type of inspection required. For severely disabled users: 5. Does maintenance require intervention of other individuals (binary yes/no)? 6. What assumptions does this technology make about access to other technology, technical know-how, or capital investment on the part of users? What measurable cost burdens do these assumptions entail?	IE and ADP teams must agree on the subpopulation of interest and the assumed resources available to users.
Environmental impacts	Example: Use of the technology measurably cuts the use of private vehicle travel.	IE team might be able to support measuring the following: 1. Decreased trash production. 2. Vehicle efficiency improvement. Reduction in private vehicle use.	IE and ADP teams must agree on the potential environmental impacts to be measured.
MOD integration impacts	Example: Use of the technology measurably increases integrated transportation options for users.	Expert evaluation examining the following: 1. Whether relevant public sector and Federal requirements, regulations, and policies support or impede the adoption of this technology into the transit sector and mobility on demand (MOD). 2. Whether the technology improves transit industry preparedness for MOD. 3. Whether the technology assists the transit industry to develop the ability to integrate MOD practices with existing transit service. Document MOD best practices that may emerge from the metrics and demonstrations of impact.	IE and ADP teams must agree on whether impacts on MOD are a priority in the context of this technology.

Source: Federal Highway Administration.

User Empowerment and Social Acceptance

This context helps the IE team examine the ATTRI ADP technology within a traveler-empowering social model of disability. The social model of disability views “the problem” of disability as structural disaccommodation. Therefore, the onus of change lies with society (social institutions and the built environment), not the individual with the disability.

As a result, the empowerment context of the evaluation effort examines the degree to which the ADP technology gives users “involvement, control, and the ability to make choices” (Mir and colleagues, 2001). Tasks within this context assess whether the ATTRI ADP technology is consistent with users’ intentionality and freedom. There are several layers to this evaluation context. A number of criteria are applied to assessing whether the ADP appropriately typifies a traveler-empowering social model of disability. This includes an assessment of the use of narrative, language, cultural sensitivities (particularly disability culture sensitivities), affordances of choice and personalization, equitable use among different traveler subpopulations, and user control (both actual control and perceived control). This context also applies metrics to assess how the technology facilitates users’ expression of intentionality and freedom through the removal of barriers posed by the environment, informational gaps, user learning curve, adaptation, and adoption.

Table 19 provides key questions and directions that will assist the IE team in the creation of evaluation hypotheses for this context, as well as insights into the performance measures that can be used to evaluate those hypotheses. It also describes specific items they should agree upon in scoping the evaluation. These questions should be used with respect to particular TALs and subpopulations in forming the evaluation hypotheses.

Table 19. Directed questions, hypotheses, and performance metrics for empowerment and social acceptance.

Directed Questions	Evaluation Hypotheses	Performance/Evaluation Metrics	IE and ADP Teams Need to Agree On
Cultural competence in values and attitudes: Are help and documentation culturally competent? (Help may include nonwritten assistance, including helpline, customer service, personal attendant, assistive services or other training services.) Does the system use value-sensitive language, with form factors, words, phrases, and concepts representative of and appropriate for diverse users? Does the system specifically use nonmedicalizing terms? Does a review of the help and documentation manual show a mission statement, goals, policies, and procedures that incorporate principles and practices that honor and respect people's abilities, age, beliefs, language, interpersonal styles, and behaviors, as well as promote disability diversity and the notion that disability is not a person's defining feature but rather a design mismatch between the built environment and a person? Does the system avoid value-laden statements and assumptions about users' independence?	Examples: 1. System's messaging is accepting of different levels of self-help skills, making no assumptive statements about person's abilities. 2. System promotes person-first language. 3. System messaging does not make disability the responsibility of an individual. 4. System "promotes consumers and careers becoming active partners in the support they receive" (National Council on Disability, 2003). 5. The system avoids stereotyping. 6. The system avoids promoting a rigid definition of independence or promoting doing things on one's own as a preferred mechanism.	1. Review of help and documentation materials. 2. Review of helpline interactions. 3. Review of customer service manual and training materials. 4. Review of customer service interactions in the field. 5. All reviews will be evaluated against the agreed-upon criteria in the next column.	IE and ADP teams must agree on what constitutes cultural competency criteria. For example, for appropriate use of language, the following guidance could be considered an accepted style guide: https://ncdj.org/style-guide/
Physical environment, materials and resources: How does the material design of the technology artifact match up against the three-dimensional qualities necessary to accommodate variable users, as reflected by use of surfaces, dimensions, and shadows? How does the material design of the technology artifact account for different environments (variable lighting, wind/air flow, portability, surfaces, and other features of the built environment)?	Example: 1. System can accommodate users' variable age, height, weight, stature, and point of view (for example, laying down).	1. Evaluation of the use of the system by outlier users—extreme height, weight, stature, variable positioning devices (including walkers, exoskeletons, standers, and laying down devices). 2. Evaluation of system portability in variable environments.	IE and ADP teams must agree on the definition of extreme users and extreme environments.
Community participation and involving others: How does the system navigate a traveler calling for help or assistance? Calling authorities? Calling family and caregivers?	Examples: 1. System involves the family and caregivers in decisions about support when desired, as desired, and in an appropriate manner. 2. System incorporates a strategy to involve caregivers or support network, recognizing the strength of a personal network and the individuality of users' comfort in involving their network.	Heuristic evaluation: Five-point Likert scale on the relative control of the system afforded a caregiver versus the intended user (range varies from an equally active partnership for the role of user and role of caregiver, to a large differential in the privileges of user and caregiver). User testing: 1. Five-point Likert scale concerning whether users indicate they wish to change the role of the caregiver.	IE and ADP teams must define, within the context for the subpopulation of interest, the appropriate triggers for calls for help, and the appropriate balance between caregiver and user control.
Messaging: Does the system treat all travelers as entities with agency and identify any barriers in travel as neither the burden or fault of the individual, nor something inflicted on "suffering" travelers? *Rather, are any barriers experienced by the traveler said to be caused by a structural disaccommodation?*	Example: 1. System strategy is to involve travelers actively in their travel, including requiring active participation in confirmatory actions, making outbound calls, invoking triggers, etc. 2. System names barriers in travel using person-first language and implicating mismatched environments and structures as the source of the barrier.	Qualitative heuristic assessment of any messages communicated by the system, rated on a five-point Likert scale rating whether agency is provided to the individual and the onus for travel barriers is assigned to the traveler or alternative sources.	IE and ADP teams agree on which messaging to evaluate.

Table 19. Directed questions, hypotheses, and performance metrics for empowerment and social acceptance (continuation).

Directed Questions	Evaluation Hypotheses	Performance/Evaluation Metrics	IE and ADP Teams Need to Agree On
User empowerment: Does the technology afford users control of the device, the technology and their travel? Does the technology afford users the ability to make choices? Do users feel confident and in control while using the technology?	Examples: 1. The technology provides users choices in how the technology operates. 2. The technology presents travelers with travel choices that are appropriate for them.	User testing of the technology experience and design. Task-based and scenario-based user study to collect analytics in which users are asked to change their preferences and control system-specific subtasks. Metrics: 1. Can the system accommodate user preferences? 2. Can the system support changing preferences on-the-fly? 3. Do users indicate controls they want that the technology does not support? 4. User ratings of controls on a five-point Likert scale along metrics of usefulness of controls, clear availability, clear definition of controls, visibility of controls, and intuitiveness.	IE and ADP agree on the definition, within the context of TALs and certain subpopulations of interest, the types of controls the technology should allow users.
Customizability/personalizability: Does the technology promote user engagement that gives users control and freedom through an interface that negotiates compatibility and establishes trust?	Examples: 1. System controls and preferences are in clear view and promoted as a key feature of the technology rather than hidden. 2. System controls accommodate the full variation in the subpopulation of interest. (For example, full visual acuity can be accommodated if the technology is addressing the needs of the community with low vision.)	Evaluation of preferences and personalization abilities. Heuristic evaluation metrics along the following queries: **Predictability:** 1. Can users predict the consequences of their actions when specifying system personalization? 2. Do users understand how user profiling and the tailoring of system output work? **Comprehensibility:** Can the system be fully customized to users' abilities? (This requires the IE and ADP teams agreeing on the full breadth of the population that the technology must serve.) **Controllability**: Can users control their user profile and the generation of personalized output, and do users feel in control? **Unobtrusiveness:** Can users complete their tasks without being distracted by personalization features? **Privacy:** Do users have the feeling that the generation of a user profile infringes on their privacy? **Breadth of experience:** Do users lose the possibility of discovering something new because output only complies with their user profile? **System competence:** Do users have the feeling that the system creates an invalid user profile or does not personalize output successfully? **Settings:** Measuring how much customization is allowed for: 1. Users' diverse abilities. 2. Perceptions of time. 3. Preference for direct or indirect communication. 4. Preference for verbose or pictorial communication. 5. Importance placed on support from others (including calling for help, physical and emotional support). Evaluation of whether the user feels the system is "credible," "trustworthy," and "supportive."	IE and ADP teams must agree on the specific range of abilities that the technology must serve. For example, this goes further than merely stating that the population served has "visual disabilities" but expresses the visual acuity, and range of low vision users that the technology is expected to accommodate.

Table 19. Directed questions, hypotheses, and performance metrics for empowerment and social acceptance (continuation).

Directed Questions	Evaluation Hypotheses	Performance/Evaluation Metrics	IE and ADP Teams Need to Agree On
Spontaneity: Does the system enable travelers with disabilities to experience greater levels of parity in changing a scheduled trip or responding to real-time travel changes or conditions (e.g., impasses, construction, etc.)?	Example: 1. System allows for just-in-time trip planning when a particular aspect of the infrastructure is not accessible to the user.	User testing of the technology experience and design. Task-based and scenario-based user study will collect analytics in which users are asked to perform system-specific subtasks and certain trips. Testers will then pose hypothetical impasses or scenarios that will require users to either (1) respond to hypothetical changes in the infrastructure (e.g., Pretend this transit elevator is broken, what would you do to recover?); or (2) respond to hypothetical changes to the itinerary (e.g., Your friend asks you to divert from your way home and meet at a nearby coffee shop). Metrics: 1. How much can the system accommodate changes on-the-fly? 2. What and how many types of information are available through the technology to users that would not otherwise be available? 3. What and how many infrastructure pieces are available through the technology to the user that would not otherwise be available (e.g., Can the technology roll over stairs safely whereas before users were unable to negotiate stairs)? 4. Segregate new pieces of information or infrastructure into subclassifications and rate their importance based on the Needs Assessment report a. Does the system support spontaneous functionality for each information need named in the Needs Assessment report 9 (binary yes/no)? b. Does the system support spontaneous functionality for each barrier named for the subpopulation of interest in the Needs Assessment report (binary yes/no)? Average time to respond to spontaneous changes under examination.	IE and ADP teams define, within the context of TALs and subpopulations of interest, the events through which the technology should allow spontaneous changes to a preset itinerary.
User empowerment: Does the technology afford users control of the device, the technology and their travel? Does the technology afford users the ability to make choices? Do users feel confident and in control while using the technology?	Examples: 1. The technology provides users choices in how the technology operates. 2. The technology presents travelers with travel choices that are appropriate for them.	User testing of the technology experience and design. Task-based and scenario-based user study to collect analytics in which users are asked to change their preferences and control system-specific subtasks. Metrics: 1. Can the system accommodate user preferences? 2. Can the system support changing preferences on-the-fly? 3. Do users indicate controls they want that the technology does not support? 4. User ratings of controls on a five-point Likert scale along metrics of usefulness of controls, clear availability, clear definition of controls, visibility of controls, and intuitiveness.	IE and ADP agree on the definition, within the context of TALs and certain subpopulations of interest, the types of controls the technology should allow users.

Source: Federal Highway Administration.

Appendix B. User Needs

The Accessible Transportation Technologies Research Initiative (ATTRI) completed an activity and issued a report that solicited feedback and information from stakeholders on user needs for ATTRI's user groups. Several different stakeholder engagement and outreach activities were conducted as part of this project including a literature review, a series of webinars, presentations at several conferences with "listening sessions," and an in-person workshop. Each of these efforts was designed to provide information on many of the same topics—though using different mediums and techniques. Results from each activity were used as initial inputs for the remaining activities, allowing for the development of common themes and gathering of detailed information from stakeholders. The literature review, which was conducted first, provided insights into the historical perspectives of stakeholders as reported on by other researchers. This information directly informed the development of discussion topics and focus areas for a series of webinars. In turn, the webinars provided the second layer of context and information that was utilized and explored as part of an in-person workshop. The final report, "Accessible Transportation Technologies Research Initiative (ATTRI) User Needs Assessment: Stakeholder Engagement Report," is available through U.S. Department of Transportation's (USDOT) National Transportation Library, at https://rosap.ntl.bts.gov/view/dot/31320.

Among those persons in the general population with a disability, the user needs are often strongly associated with their type of functional disability, which was a common theme throughout the literature review. Key needs of persons with disabilities identified consistently in the literature include:

- Those with **a visual disability** have a heavy reliance on public transportation including on-demand and paratransit services and, therefore, need availability of public transportation resources including fixed service, on-demand and paratransit service; information about the public transit system in accessible form that can be used with screen readers; and audio announcements on and around public transportation facilities such as transit hubs, platforms, etc. and vehicles to provide both static and real-time information.
- Those with a **hearing disability** require non-audible means to receive information such as visual displays of upcoming stations/stops and emergency notifications. Additionally, these persons need transit operators to have basic non-audible communication skills or have the ability to achieve communication with the operator through advances in technology such as text-to-speech. These types of information mechanisms need to be included in passenger vehicles at an affordable cost to the user alleviating the requirements for special licensing.
- Those with a **cognitive disability** require information in a fashion that is simplified and made easier for their consumption. Simplified navigational guidance/instructions and information on operational aspects of public transportation are two particular areas that need improvement. Technologies that provide real-time, interactive assistance are needed by this user group as additional training for transit operators.
- Persons with **mobility disabilities** are typically constrained by the physical environment within the transportation system. Physical and architectural components of public transportation facilities need additional design considerations to make them more accessible, while assistive technologies that provide real-time updates of inoperable transit system components and alternatives have been identified as needed. Priority labeling of seats, waiting areas, etc. are typically called out as needs in the literature review for this segment of the population.

The literature review provides a perspective into the characteristics of those subsegments of the populations that are of interest to ATTRI. Regardless of the sub-segment of the population, many of the fundamental user needs for these groups of persons were found to be very similar across the literature review, webinars, and workshop. There are many themes and conclusions that emerged as a result of this research project. First and foremost, based upon the literature review, listening sessions at conferences, webinar(s) responses, and the discussion among participants at the workshop, persons with disabilities, Veterans with disabilities, and older adults clearly have significant needs and barriers to mobility. Although needs and barriers vary by sub-population and type of disability, several themes were observed to emerge regarding these needs and barriers, and potential technology solutions to address them. These overarching themes include:

1. **Information** for travelers with disabilities is a critical component for mobility. Having access to information prior to and during a trip was overwhelmingly supported by participants in project outreach activities, as well as documented in the literature review. Although this was the most cited need and barrier to improved mobility among persons with disabilities, it is interesting to note very few of the current “best practices” for transit identified through the literature review as having been adopted within the international and domestic transit community address the inaccessibility, relative unavailability, and lack of comprehensibility of information sources for travelers with disabilities. At the same time, however, participants in this research project were readily able to provide numerous specific examples of the types oftechnologies that could be utilized to address this need.

   The topic of “Information” is an area that is ripe for pursuit by the USDOT and others as relatively immediate technology insertion and would have a significant impact on improving mobility. Wayfinding and Navigation technologies tend to be the type of technologies that are most identified as being capable of meeting these needs and overcoming thesebarriers. Assistive Technologies and Enhanced Human Services Transportation technologies are also very relevant and should be pursued. There are certainly challenges with these technology options that include the reported lack of access to and utilization of technology information dissemination devices such as mobile phones, tablets, portable computers, etc. among persons with disabilities. As the trend in technology adoption by the general public continues upward, so too should adoption among persons with disabilities.

2. Travelers with Disabilities need travel **Options** before and during their travel. Many user needs and barriers can be directly attributed to the lack of or the perceived lack of travel options. In particular, travelers want to be able to have choices when planning their travel as well as have choices available during their travel to accommodate service conditions and other factors encountered. Veterans with disabilities and older adults, of which a significant percentage reside in rural areas, often have only a limited selection of transportation options available. Automation and Robotics technologies may provide solutions to these needs by allowing more utilization of personal automobiles by persons with disabilities. Data Integration and Enhanced Human Services Transportation are technology focus areas that can provide technology solutions to enhance options by addressing, through coordination and data sharing, first/last mile integration and the lack of available transportation options.

   Travelers with disabilities also want options when it comes to the facilities and amenities. They want to know prior to their trip what the transportation facility configuration is and the status of services. If unsuited or inoperable, they want alternative options for travel.

Travelers with disabilities who are on a trip when services and/or amenities change status or become challenging for them to manage, want to be notified and provided with options so they may continue their travel. Receiving real-time assistance through technology options, such as a virtual personal assistant or an electronic guide dog, are examples of technology options cited as Wayfinding and Navigation as well as Assistive Technologies.

Having options with respect to travel, improves the overall perception of the transportation system as well as the overall quality of life for persons with disabilities. For older adults, it means that they have a much more likely ability to age-in-place as they can retain more of their independence by remaining mobile. All of the technology focus areas have a role to play in potentially providing travel options either before or during a trip and should continue to be explored.

3. More travel **Assistance** could be given to travelers with disabilities during their travel. In many cases, this assistance may be most easily provided by transit operators and transportation providers/staff. However, many of these individuals need increased levels of training that is specifically focused upon providing assistance to persons with different types of disabilities. Training of transportation providers alone will likely not resolve the need that persons with disabilities often have for assistance on topics such as wayfinding, trip planning, and notification of arrival at their destination/connection. Technology solutions such as mobile tactile or audible based applications for Wayfinding and Navigation would potentially provide solutions to the need for travel assistance. Other solutions that address the barrier of a lack of customer support/service include technologies that deploy Data Integration, such as crowdsourcing and social media applications, would also be beneficial.

   Technology-based solutions that provide real-time assistance to travelers with disabilities include the use of a virtual electronic guide dog, wearable devices to provide course corrections, and even semi-autonomous technologies that detect when a traveler with disabilities has departed from their expected path and automatically begins to provide real- time guidance assistance through “on-demand virtual concierge service” technology, mobile applications, or other methods.

4. **Access** to transportation facilities could be enhanced through technology solutions, but current “best practices” have been effective in meeting this user need. Many of the current “best practices” within the transit industry involve improving the physical configurations and layouts of transportation facilities, so that they are more “accessible” to persons with disabilities. Curb cuts, raised strips indicating the edge of a transit platform, and connected and continuous pathways are examples of the types of best practices that have been deployed to improve access for persons with disabilities. These types of best practices and corresponding changes within the transportation system seem to be working. Needs and barriers associated with a lack of access were consistently cited lower than all other needs and barriers for mobility.

   There were a relatively small number of potential technology solutions that were identified as part of this User Needs: Stakeholder Engagement project that could contribute to improving access, although these solutions also could be typically be considered as “assistive” technologies. For example, assistive devices, such as a white cane with environmental sensors to detect hazards (e.g., black ice) or tactile navigation/information device for individuals who are deaf-blind, or proximity-based public announcements through text-based messaging are the type of technologies that were identified. Autonomous systems such as fully autonomous vehicles would also improve access to the transportation system.

Table 20. User needs reported by Accessible Transportation Technologies Research Initiative stakeholders.

Category	User Need
Information	• Amenity information (e.g., restroom, shelter, benches, food, drinks). • Real-time transportation information. • Safety, security, and emergency information. • Transit schedule and other transit information. • Destination information (hours, addresses, entrances, layout). • Mapping/directions. • Roadway/pathway real-time conditions. • Information in a variety of accessible formats. • Connection information (where, who, when). • Weather conditions. • En route assistance and information. • Trip length/distance. • Signage. • Transportation facility information. • Personal list of travel needs (e.g., oxygen, emergency number). • Coordination information (between agencies, modes). • Profile of traveler (for agency to accommodate traveler). • Landmarks and orientation identifiers.
Options	• Accessible payment options and trip cost information. • Flexible and/or spontaneous travel options. • Vehicles with accessible equipment. • Technology, communications devices, and recharging for traveler use. • “Last mile” transportation options.
Assistance	• Traveler help line/customer service. • Personal care attendant or other assistive/training services.
Access	• Connected, continuous, accessible pathways. • Accessible parking locations and availability.
Adverse perception of travel	• Heavy crowds and noise. • Perceived or real safety or cleanliness. • Weather.

Source: Federal Highway Administration-JPO 2016.

Table 21. Barriers identified by accessible transportation technologies research.

Barrier Category	Barrier Type
Cost	• High cost or lack of resources/funds.
Inadequate infrastructure, signage, or wayfinding tools	• Inadequate crosswalk infrastructure or signal times. • Inconsistent accessible pathway infrastructure. • Lack of first- or last-mile options. • Lacking or inaccessible signage/maps/landmark identifiers/announcements.
Inadequate transportation options and amenities	• Lack of accessible service, facility information (or not current). • Lack of available transportation (limited hours, vehicles, service area, etc.). • Lack of first- or last-mile options. • Limited or no accessible amenities (restrooms, benches, shelter, water fountains). • Nonflexible transportation options (no same-day service, same-day changes). • Trip lengths/duration too long. • Unreliable transportation (fleet, equipment, on-time performance). • Vehicle/facility configuration or policy does not meet need.
Lack of technology access	• Lack of access to technology (phones, computers, charging, or lack of training). • Lack of real-time travel information. • Navigation difficulties (do not know when to arrive, transfer time, distance).
Lack of travel support/ customer service	• Lack of assistance or attendants. • Lack of coordination/comprehensive travel information.

Source: Federal Highway Administration-JPO 2016

Table 22. Disabilities affecting travel mobility.

Disability Category	Description of Issue/Concern	Activity Concerns
Motor	Requires Assistive Device: • Cane. • Crutches. • Walker. • Support animal.	• Limited ability to hold a device. • Limited distance or speed of movement. • Limited stair climbing, restricted use of some paths. • Limited slope that can be traversed. • Possible need for storage space for assistive device.
Motor	Requires Assistive Device: • Manual wheelchair. • Power wheelchair.	• Inability to use stairs, limiting route choice. • Limited ability to traverse a slope. • Limited ability to carry devices. • Limited distance that can be traveled. • Need to be locked down on transit vehicles and may require driver assistance to perform this task.
Motor	• Inability to use stairs. • Loss of balance. • Low stamina. • Limited grasp. • Difficulty walking a quarter of a mile. • Difficulty lifting 10 lbs. or more.	• Inability to use stairs, limiting route choice. • Limited ability to carry or **<u>use</u>** devices. • Limited routes based upon distance between transit options. • Limited use of public transportation for activities such as grocery shopping.
Hearing	• Deaf. • Partial hearing acuity.	• Inability to hear approaching vehicles, creating safety risk. • Limited information transfer from auditory clues (e.g., inability to hear location announcements). • Inability to hear auditory payment receipt cues.
Speech	• Communication disorder.	• Difficulty being understood when asking for directions or assistance (vehicle boarding or attempting to identify exit stop or transfer boarding location or vehicle).

Table 22. Disabilities affecting travel mobility (continuation).

Disability Category	Description of Issue/Concern	Activity Concerns

Vision	• Blind.
	• Partial vision can impact:
	  ○ Acuity.
	  ○ Field of vision.
	  ○ Depth perception.
	  ○ Scanning efficiency.
	  ○ Color perception.
	  ○ Contrast sensitivity.
	  ○ Glare/light sensitivity.
		• Need for screen readers or other technology to work with Internet applications.
		• Cannot see and identify approaching transit vehicles.
		• Difficulty tactically navigating because of difficulties in detecting and avoiding objects.
		• Difficulty navigating because of inability to determine exact position and location relative to surroundings.
		• Difficulty with situational awareness because of lack of sensory information, resulting in difficulty finding entrance/egress points, identifying transit stops, identifying specific transit routes or vehicles, and safety risk.
		• Difficulty paying for transit because of an inability to identify the payment location.

Learning/ Cognition	• Dementia. • Processing disability. • Intellectual disability. • Learning disability. • Mental or emotional disability. • Developmental disability. • Alzheimer's. • Senility. • Autism.	• Limited short-term memory, making it difficult or impossible to remember directions. • Easily confused by changes in environment. • Difficulty reading signage. • Social anxiety. • Limited problem-solving capacity.
Multiple Disabilities	• Intersection of two or more of the above categories. • Example: • Deaf and blind.	• Varies depending upon the specific disabilities. • Example: ○ Lack of auditory and sensory input significantly limit the ability to interact with the travel environment, often requiring assistance. • Information delivery must be performed via tactile interfaces, requiring data formatted for that purpose.

Source: Federal Highway Administration.

Appendix C. User Needs Checklist

The table on the following page can be printed and used by independent evaluation teams to document and help prioritize the user needs that a given Accessibility Development Projects (ADP) is designed to address.

To use this checklist, indicate for each user need whether the ADP being evaluated includes features that address this specific need by placing a "Y" or an "N" in the third column of the table.

When a given need is being addressed by the ADP, also indicate the significance of that specific feature to the overall objectives of the ADP being evaluated. Where a given user need applies to two very different types of disabilities, the significance can differ between those target populations. For example, the ADP may include providing auditory announcements that a bus is arriving, and that the arriving bus is a northbound, route 7. Having this type of feature in the ADP would result in responding "Yes" to the user need "*Does ADP provide real-time transportation information?*" The significance of that feature to the evaluation objective as applied to users with visual disabilities might be scored as "High" while the importance that same feature to individuals with motor disabilities might be scored as "Low."

Once filled out, this table serves as a good starting point for both the development of a logic model for the evaluation, and a means where the priorities of the evaluation can be discussed with both the ADP team and evaluation project's sponsor.

Table 23. User needs checklist.

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable? Y/N	Significance to ADP Objectives (L, M, H)
All	Does ADP provide information in a variety of accessible formats?		
All	Is information from ADP interface accessible in a variety of environments (i.e., amid heavy crowds and noise, underground)?		
All	Does ADP perform a task that improves safety and security or that provides emergency information?		
All	Does ADP provide en route assistance and information?		
All	Does ADP provide connection information (where, who, when)?		
All	Does ADP provide estimated trip length and distance?		
All	Does ADP provide comprehensive travel information?		

Table 23. User needs checklist (continuation).

Subpopulation	Which User Needs Are Addressed by the ADP?	Applicable? Y/N	Significance to ADP Objectives (L, M, H)
All	Does ADP require access to equipment (phones, computers, charging, training)?		
All	Does ADP allow the user to create a personalized profile?		
All	Does ADP require coordination of information (between agencies, modes)?		
Blind and visually impaired (BVI) (Visual) Motor Impairment (MI) (Motor)	Does ADP provide real-time transportation information?		
BVI (Visual)	Does ADP provide audible mapping/directions?		
BVI (Visual)	Does ADP provide destination information (hours, addresses, entrances, layout)?		
BVI (Visual) MI (Motor	Does ADP provide transit schedule and other transit information (e.g., stop location)?		
BVI (Visual) MI (Motor	Does ADP provide information about pathway infrastructure?		
BVI (Visual) MI (Motor	Does ADP include provision for outside assistance or attendants?		
BVI (Visual) MI (Motor	Does ADP provide amenity information (e.g., restroom, shelter, benches, food, drinks)?		
BVI (Visual) Cognitive	Does ADP provide information about, and interpretation of, signage?		
MI (Motor)	Does ADP provide transportation facility information (e.g., maps)?		
MI (Motor)	Does ADP provide information about weather conditions?		
Hearing Cognitive	Does ADP include information about and/or interpretation of announcements?		
Hearing	Does ADP incorporate speech-to-text or text-to-speech that enables the user to communicate more easily?		
Cognitive	Does the ADP provide information in a concise and straightforward manner?		

Source: Federal Highway Administration.

Appendix D. Travel Activity Links

To help IEs create Accessibility Development Projects (ADP)-specific evaluations, the evaluation process relies on a structured view of trip-making behavior. That structure separates the specific activities that individuals perform when traveling into manageable pieces called Travel Activity Links (TAL). Every trip (from Origin A to Destination B) requires that travelers perform one or more of these activities. The number of TALs performed and the order in which they are performed vary from trip to trip. TALs are defined as follows:

- TAL 1: Trip planning (both pre-trip and midtrip).
- TAL 2: Accessing trip itineraries midtrip and assessing trip progress.
- TAL 3: Identifying entry/egress:
  - 3a to/from a transit vehicle.
  - 3b to/from a travel environment.
- TAL 4: Entry/egress:
  - 4a to/from a transit vehicle.
  - 4b to/from a travel environment.
- TAL 5: Pedestrian-only environments.
- TAL 6: Street crossings and intersections.
- TAL 7: Mixed environments with moving vehicles and pedestrians.
- TAL 8: Indoor and underground transit facilities.
- TAL 9: Outdoor transit facilities (e.g., transit transfer centers).
- TAL 10: Riding a vehicle.
- TAL 11: Transit payment (includes identifying payment location).

Each TAL is briefly described below.

Travel Activity Link 1: Trip Planning

This TAL relates to how travelers plan trips from origin to destination. The task can be performed in advance, in real time, or in the middle of the trip. This TAL is the strategic planning portion of the Smart Wayfinding and Navigation technology development area. It can also be important for ADPs that are designed to support the Pre-Trip Concierge and Virtualization development areas. Specific tasks that fall within this TAL include, but are not limited to, travelers doing the following:

- Identifying the origin, destination, and start time or desired arrival time of a trip.

- Determining one or more feasible paths (strategic navigation), using one or more modes of travel from that origin to that destination.
- Learning the financial cost for each of the trip options.
- Selecting the preferred path for making that trip.
- Identifying and reporting key waypoints (e.g., transfer points).
- Where appropriate, making reservations for rides to accomplish the selected trip.

Travel Activity Link 2: Accessing Trip Itineraries Midtrip and Assessing Trip Progress

This TAL includes the tasks associated with travelers identifying their current location and associating that location with their current travel plan to determine the following:

- Their current trip status relative to that plan.
- Whether that plan is still valid (e.g., determining whether they will be able to make a planned transfer they are expecting to make).
- Whether they need to perform a specific task in accordance with that plan (e.g., prepare to exit a transit vehicle).
- Whether they need to alter their plan because it is no longer valid.

This TAL is part of the Wayfinding and Navigation development area. It includes the strategic aspects of the situational awareness tasks that travelers must perform. Some tasks performed within this TAL can also be addressed with pre-trip concierge services.

Travel Activity Link 3: Identifying Entry/Egress to/from Vehicles and Environments

This TAL addresses tasks associated with travelers transitioning from one travel environment to another. These tasks include the following:

- Identifying the location of an entrance/exit that they can use to move from their current environment to the desired environment.
- Identifying a path, which they are capable of navigating, that leads to that door or entry/exit point.
- Determining whether a door or other barrier exists at the entrance/exit.
- Determining how to open that door or pass through that barrier.

For people with specific disabilities, some of these tasks can be quite difficult. An example is a wheelchair user who needs to board the Route 16 bus at a busy downtown stop. Multiple bus routes use that stop location. The traveler might well be waiting at the correct stop, but three buses arrive at roughly the same time. Because Route 16 is third in line, it has not stopped at the sign where the traveler waits. The traveler must determine whether her bus is in the queue of buses, whether she needs to move to the bus, or the bus will make a second stop, and, once that is decided, where she needs to be located. Once she has identified the correct vehicle, she must communicate to the driver that they need to board that bus and that the driver should open the door and activate the wheelchair lift/ramp.

A somewhat different, but equally complex, process can occur with travelers entering a building from the sidewalk. In this example, a traveler who uses a wheelchair must first identify the location to which he is going (e.g., which entrance among several on that block is the destination) and whether that destination has a wheelchair accessible entrance. If so, what path does he need to follow to reach that entrance, and is that path currently accessible? If that entrance is not accessible to him, where is an entrance that is accessible? Finally, once they have identified an accessible entrance, they must determine whether the door to the building has a push button to start an automated door opening system, or if there are other mechanisms used for opening the door, as that mechanism, by itself, may prove to be a barrier to his entrance to the building.

Finally, in working with paratransit services, this TAL includes the following additional list of tasks:

- How drivers identify the individuals they are to pick up.
- Where they are to meet those individuals.
- Where they need to park to pick up those individuals.

For individuals being picked up, in addition to knowing where the paratransit vehicle will park, they need to be able to identify the correct vehicle or individual providing the paratransit service.

This TAL may be appropriate for ADPs in the Smart Wayfinding and Navigation, Pre-trip Concierge and Virtualization, or Safe Intersection Crossing development areas, depending on when the traveler accesses information and the environmental circumstances where that access occurs.

Travel Activity Link 4: Entry/Egress to/from Vehicles and Environments

This TAL includes the tasks associated with travelers transitioning from one physical environment to another. The tasks are second nature for many travelers, but they can cause significant difficulty for individuals with mobility disabilities. The activities in this TAL include the physical tasks associated with moving into or out of vehicles and buildings, as well as communications tasks to elicit assistance from vehicle operators or systems and to be informed that those requests have been received and are being acted upon.

Tasks associated with this TAL start where TAL 3 stops. Continuing the example of the wheelchair-using bus rider from the previous TAL, after the rider has identified the correct bus and signaled to the driver, then:

- The driver must initiate the technology that the traveler will use to board the vehicle (e.g., stop other riders from boarding, if necessary, request that riders currently on the bus move to free up space designated for wheelchairs).
- The driver must open the bus door.
- The driver must activate the wheelchair lift/ramp.

- The traveler must then position herself to use that lift/ramp.
- The traveler must use the lift/ramp, sometimes with driver assistance, sometimes without assistance, depending on many factors.
- A set of tasks is required to navigate from the lift/ramp to the designated on-vehicle parking location.
- Finally, the traveler must secure the wheelchair in position, again, with or without assistance from the driver.

Similar tasks need to be performed, in reverse order, to exit the vehicle.

Similar complex information gathering, and technology interactions are often necessary for some individuals to enter or exit a building and successfully reach the sidewalk. Many of these tasks fall into the Robotics and Automation development area. Others are part of the Wayfinding and Navigation task. Others do not fall neatly into any of the four development tasks but are found in the policy and organizational changes being worked on within the Accessible Transportation Technologies Research Initiative (ATTRI) program.

Travel Activity Link 5: Pedestrian-Only Environments

This TAL includes the tasks associated with travelers moving through above-ground, pedestrian-only environments. (It does not include intersections, indoor environments, or underground environments.) Specific activities include travelers doing the following:

- Gaining situational awareness (identifying where they are in space, what objects are nearby that need to be avoided, whether objects are moving or stationary).
- Predicting the path of moving objects and forecasting the speed at which those objects will move along that path.
- Identifying paths that can be navigated within the pedestrian environment (e.g., can travelers actually walk on that surface, or should a different path with a better surface be taken?).
- Selecting the specific travel path, they wish to take through that environment to their next travel waypoint (tactical navigation).
- Physically moving through that environment, including obstacle avoidance and mobility assistance (e.g., balance support).

Technologies that may be developed within the ATTRI program to address this TAL are those that assist individuals moving through pedestrian environments. These include aspects of the Wayfinding and Navigation and the Robotics and Automation development areas. This TAL differs from indoor or underground pedestrian facilities in that it assumes that assistive technologies that might be available indoors or underground are not available. It also assumes that global positioning system (GPS) is functional, giving at least some real-time location awareness.

Travel Activity Link 6: Street Crossings and Intersections

This TAL is specifically designed to address the complex tasks required to safely navigate across high-risk arterial environments. It includes crossing streets both at signalized intersections and where no signalization is present. Tasks that must be accomplished in this TAL include, but are not limited to, the following:

- Identifying safe crossing locations.
- Identifying crossing locations where it is physically possible for users to cross (e.g., whether there are wheelchair ramps, or where barriers in the street might prevent a user with specific physical characteristics from crossing).
- Determining when traffic conditions allow for safe crossing, including identifying moving motor vehicles and other moving objects (e.g., bikes, eScooters, other pedestrians).
- Navigating across the street.
- Navigating around moving and stopped vehicles, bikes, and pedestrians.
- Identifying the location for exiting the street environment (e.g., whether there are wheelchair ramps on the far side of the street and, if not, how the traveler can reach the sidewalk).

These tasks must be performed in concert with an understanding of the traffic control environment or a lack of traffic control (e.g., Is the light green? How long will the light stay green? Is there a stop sign? Do all directions of traffic have stop signs?).

A specific ATTRI development area, Safe Intersection Crossing, applies to this TAL.

Travel Activity Link 7: Mixed Environments with Moving Vehicles and Pedestrians

This TAL shares many of the characteristics of the Pedestrian-only TAL and the Street Crossing and Intersections TAL. It differs from those two TALs in the nature of interactions between travelers and the other objects found in the environment, as well as the expectation of information about that environment.

For example, this TAL applies to travelers with disabilities walking across a parking lot. Vehicles can be present and moving in this environment along with the travelers; however, they can be moving in any direction (forward, backward, turning, or moving straight), often without indications of their expected movement (e.g., sudden turns). In addition, other moving objects (e.g., pedestrians, bikes) are also frequently encountered moving through the environment, often on unmarked paths, and thus without well-defined headings.

A second example of a mixed environment is an outdoor urban mall that is primarily a pedestrian area, but where bikes, delivery carts, and other small vehicles are common and navigate through the environment at slow speeds. At the same time, this environment is filled with a number of fixed objects (e.g., seating, planters, advertising signs, public art), as well as curbs or other vehicle control devices, that must be avoided while navigating.

Tactical navigation through these environments can be difficult for many travelers because of a lack of navigation clues and the difficulties they face in collision avoidance. Path identification in these environments is particularly important, both because of the physical barriers to travel (e.g., curbs, planters, construction closures) and because of the potential level of activity or crowding. Both tactical navigation and path identification in some of these environments can require a more precise location referencing system than is expected from GPS.

ATTRI technologies within this TAL are associated with the Wayfinding and Navigation development area, although some Robotic Assistance applications also apply to this TAL.

Travel Activity Link 8: Indoor and Underground Transit Facilities

This TAL addresses the complex tasks required to safely navigate subway stations and other indoor transportation facilities. This TAL falls within the Wayfinding and Navigation ATTRI development area.

This TAL differs from the outdoor transit facility TAL and the pedestrian-only environment TAL in two ways. The first is that in this transportation environment, travelers very likely have to navigate in three dimensions. That is, they must be able to take stairs, escalators, or elevators to move between levels of a station or building, and they often need guidance as to which level they are on and the level they need to access. The second major difference is a lack of GPS signals, which significantly limits the effectiveness of many location-referencing systems that work perfectly well outdoors.

Moving vertically requires that travelers have access to information on the location and operating condition of the various options for traveling from one level to another within the station. Individuals with different abilities often require different paths for navigating within underground stations. For example, wheelchair users cannot navigate paths requiring the use of stairs or escalators. They may need to exit at the preceding or following stop on a transit line if the elevators at a given station are out of order. Other users may prefer to take an elevator but may also be physically capable of using stairs or an escalator, and therefore may choose that option if the only other option is to travel to a different station.

Navigation becomes considerably harder with a lack of GPS location finding technology. Technologies that substitute for a lack of GPS are a common part of navigation systems intended for indoor and underground applications.

Other issues related to navigating these facilities arise from the need for detailed indoor mapping of the underground station, in all three dimensions. This includes detailed location information concerning the following:

- Transit stops by line and direction.
- Locations of exits to/from the station.
- Locations of key transit services (e.g., fare payment kiosks, ticket vending machines, staffed kiosks) and the degree to which those service locations are accessible to travelers with specific mobility issues (e.g., whether travelers must use an elevator, whether ticket vending machines inside the station are located on a level that is accessible by elevator or tickets must be purchased before entering the station).
- The attributes of entrances/exits to/from the station (e.g., whether there are fare gates to the station and whether they are compatible with wheelchairs and/or other assistive devices).
- The relationship of each of the station exits to services in the surrounding city.
- The locations of retail stores and other activities within the station.

Finally, in terms of basic navigation data, the need for up-to-date information is often critical to the success of travelers with disabilities in these environments, as one or more elevators or escalators may be out of service, meaning that a path that is typically navigable is not viable at a specific time.

These attributes, along with the detailed maps of the paths (and the path's attributes) within the station, must then be effectively presented to travelers in ways that allow them to be effectively used. These communication tasks are frequently made more difficult by the facts that indoor and underground transit facilities can be both crowded and noisy and may or may not have good cellular or Wi-Fi coverage. These built environment factors can make the use of many technological solutions far more difficult for people with disabilities.

Travel Activity Link 9: Outdoor Transit Facilities (e.g., Transit Transfer Centers)

This TAL is specifically designed to address the complex tasks required for travelers to safely navigate outdoor transit facilities, where there are multiple transit bays and transit transfers are common, often with little time available between arrival and departure. This TAL differs from the indoor transit facility TAL in that it assumes that GPS location signals are available. It does share the potential need for travelers to change levels by using elevators, stairs, or escalators; however, for these outdoor facilities, at most one level change (e.g., accessing an aerial rail platform) must occur. This is generally a far simpler task than navigating a complex subway station.

The TAL also shares many detailed tasks with the Mixed Environment TAL and the Street Crossing and Intersection TAL. Travelers often need to cross bus access roads to reach their desired transit bay. This means they need to identify moving vehicles, predict their travel paths and speeds, and make decisions about whether they can cross those paths safely. They also have to move from pedestrian/sidewalk environments into street environments, and back again, while also identifying where those transitions can occur. This TAL falls primarily within the Wayfinding and Navigation ATTRI development area, although some tasks are within the Safe Intersection development area.

Finally, travelers must be able to identify which transit stop, bay, or platform they need to reach and to determine a navigable path to do so. This task then connects to the Identifying Entry/Egress TAL, which includes the task of identifying the actual vehicle that travelers need to board and the specific location from which they will board that vehicle. It differs from that TAL, which is oriented around identifying where the door is located and accessed, and instead focuses on the general location in the station where they need to position themselves for the arriving vehicle, or how they orient themselves after arriving at such a stop on a vehicle. For example, this TAL might identify that the path to the southbound light rail platform is upstairs on the left, and that the elevator to reach that platform is behind the traveler, whereas TAL 3 would identify where the train doors open on that platform. Similarly, this TAL would relate to helping travelers understand that they have arrived at this station on a raised platform, and that the bus to which they must transfer is on the ground level. The elevator to the ground level is at their right, an escalator is to their left, and the bus bay they need is the third bay to the left as they exit the rail station.

Travel Activity Link 10: Riding a Vehicle

This TAL includes tasks required by travelers actively riding on a transit vehicle, whether those vehicles are buses, trains, or Transportation Network Company (TNC) vehicles (e.g., taxis or autonomous vehicles). This TAL is within the Wayfinding and Navigation development area. It includes a variety of tasks, some of which overlap with other TALs. The two primary tasks associated with this TAL are as follows:

- Providing situational awareness (e.g., answering the question, “Where am I?”) in real time.
- Providing action keys or assistance (e.g., “Now is the time to request a stop,” or “This is the station at which you need to exit,” as well as, “How do I signal for a stop?”) that connect accurate situational awareness to the travelers’ specific itinerary.

Travel Activity Link 11: Transit Payment

This TAL includes tasks associated with paying for the transit ride. It can be found in three of the four development areas: Pre-Trip Concierge, Wayfinding and Navigation, and Robotics and Automation. The specific tasks in this TAL include the following:

- Identifying where to pay.
- Identifying how to pay.
- Paying for the trip, including multiple different payment systems.
- Identifying and accessing subsidies for trip costs.

This TAL covers both the technical aspects of paying for the trip and the policies associated with reducing the cost of trip-making for specific populations, reducing the cost of providing services to travelers by implementing better systems, and encouraging intergovernmental and public-private relationships.

Combining Travel Activity Links to Create Complete Trips

Complete trips require different combinations of these TALs. For example, a simple walking trip from an office in downtown to a restaurant across the street, made by an individual every day, would include only three TALs, although some of the TALs would occur more than once during the trip. The TALs would occur in the following order: TAL 4b entry/egress, TAL 5 Pedestrian-only environments, TAL 6 street crossings and intersections, TAL 5 again, and finally TAL 4b for a second time. (Note that for this example, the trip is considered routine, the traveler knows the locations of the appropriate building entrances and exits, so we did not include TAL 3a and TAL 3b.)

An example of a more complex trip might be if that same individual needed to take a transit bus across town to meet a client at a restaurant. That trip might include TAL 1, TAL 4b, TAL 5, TAL 6, TAL 5, TAL 3a, TAL 4a, TAL 11, TAL 10, TAL 3a, TAL 4a, TAL 5, TAL 3b, and TAL 4b. In this case, the traveler would need to plan her trip—determining which bus to take, where to catch that bus, how to strategically navigate to that bus stop, and when she needed to arrive at that stop in order to catch that bus. Then she would leave her location, travel to the stop, identify that the correct bus was arriving and where the door to the bus was located. Next, she would need to board the bus, pay for the transit trip, and find her seat on the bus.

Next, the traveler would need to identify when it was time to disembark, signal the driver, and successfully exit the bus onto the sidewalk when the bus stopped. Finally, she would need to navigate to the restaurant, identify how to enter the restaurant, and complete the trip.

Different users (populations of individuals) experience a given trip (set of TALs) differently as a result of their different abilities. TALs that are easy to perform for some populations can be very difficult for others to perform. Linked together, they allow the independent evaluations (IE) to explore the combined impacts of a travel population's disabilities on their travel mobility, and the effects that a new ADP technology will have on that overall mobility.

By breaking down the complete trip (from office building to restaurant) into these travel activity components, TALs are extremely useful in identifying specific issues or user needs associated with a specific population's difficulties, such as 1) exiting the building and entering into the sidewalk environment; 2) navigating to the intersection; 3) getting across the street; 4) navigating to the proper door of the restaurant; and finally 5) navigating into the restaurant itself. The impact of the ADP technologies being studied can then be examined for those TALs to which the ADP applies. In addition, using the entire string of TALs involved in a complete trip makes it easier to identify when other factors—outside of the ADP technology—still limit the ability of specific user populations to travel freely and spontaneously. This examination of different TALs for different complete trips, given the effects of the ADP, is the qualitative complete trip analysis that can be performed at the end of the ADP evaluation.

Appendix E. List of Acronyms

ATTRI	Accessible Transportation Technologies Research Initiative
ATTRI-ADP	ATTRI-Funded Accessibility Development Projects
ADP	Accessibility Development Projects
BVI	Blind and Visually Impaired
CPU	Central Processing Unit
DSRC	Digit Short-Range Communications
FHWA	Federal Highway Administration
FTA	Federal Transit Administration
GPS	Global Positioning System
GTFS	General Transit Feed Specification
IE	Independent Evaluations
IRB	Institutional Review Board
IT	Information Technology
MI	Motor Impairment
MOD	Mobility on Demand
MPM	Mobility Performance Metrics
O/D	Origin/Destination
TAL	Travel Activity Links
TNC	Transportation Network Company
USDOT	U.S. Department of Transportation

U.S. Department of Transportation
ITS Joint Program Office—HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free “Help Line” 866-367-7487

www.its.dot.gov

FHWA-JPO-20-784

U.S. Department of Transportation