Better, Quicker, Together: Enabling Public Transport Service Quality Co-monitoring Through a Smartphone-Based Platform

by
Corinna Lubin Li

B.A. in Ethics, Politics, and Economics
Yale University (2011)

Submitted to the Department of Urban Studies and Planning
and the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degrees of

Master in City Planning
and
Master of Science in Transportation
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author_____
Department of Urban Studies and Planning
Department of Civil and Environmental Engineering
May 19, 2016

Certified by _____
P. Christopher Zegras
Associate Professor of Urban Studies and Planning
Thesis Supervisor

Certified by _____
Jinhua Zhao
Assistant Professor of Urban Studies and Planning
Thesis Supervisor

Accepted by_____
P. Christopher Zegras
Associate Professor of Urban Studies and Planning
Chair, MCP Committee

Accepted by_____
Heidi Nepf
Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Better, Quicker, Together: Enabling Public Transport Service Quality Co-monitoring Through a Smartphone-Based Platform

By

Corinna L. Li

Submitted to the Department of Urban Studies and Planning and the Department of Civil and Environmental Engineering on May 19, 2016
in partial fulfillment of the requirements of the degrees of

Master in City Planning
and
Master of Science in Transportation

at the
Massachusetts Institute of Technology

Abstract

Public transport system is an important part of cities, and the quality of public transport service – passengers' perceived performance – is a key urban indicator. Customer satisfaction surveys has been the traditional methods and metrics for monitoring and evaluating public transport service quality, but they come with a number of weaknesses. They are administered too infrequently and ask subjects to provide only general ratings. The infrequency results in potential delay for agencies to receive feedback, and the abstractness reduces the possibilities of associating satisfaction levels to specific trips and their attributes, as well as using the data to inform service improvement decisions.

Given these shortcomings with conventional surveying practices, there is great value in engaging riders as additional sources of information. This reflects the concept of "co-monitoring" – agencies using public feedback to supplement the official monitoring and regulation. This is aided by the growing ubiquity of Internet-connected mobile devices, which enables citizens to generate and submit feedback without time or geographic constraints. From the data collection perspective, this would make the process more dynamic, low-cost, and in real-time. Equally importantly, it is poised to enhance public transport agencies' relationship with their customers – conveying to customers that their experience and feedback are valued. The service sector today is increasingly striving to be more responsive to the customers' needs and experiences, seeking to strengthen the relationships with customers. The benefits of co-monitoring may help public transport agencies adapt to these current service paradigm shifts towards "real-time" and "on-demand."

This thesis documents the creation and piloting of a smartphone-based platform for engaging customers in becoming co-monitors of the local bus service quality. Working with

a team of academics and software engineers, the author leads the effort to adapt a smartphone-based travel survey system, Future Mobility Sensing (FMS), to collect real-time customer feedback as well as objective operational measurements on specific bus trips. The system (FMS-TQ) uses a combination of GPS, Wi-Fi, Bluetooth, and cellphone accelerometer data to track transit trips, while soliciting users' feedback on trip experience with built-in questionnaires. FMS-TQ has been piloted in partnerships with public authorities in Singapore and Boston. The pilots have demonstrated the platform's capability to collect trip-specific performance data, as well as value for public transport operators and regulators.

The significance of this effort is three-fold. First, it embodies one of the first successes in making public transport service quality data associable, attributable, and actionable. One can associated this information to individual trips, and attribute performance excellence or shortfalls to specific infrastructure, personnel, or service operations. As a result, the data may reveal more actionable information for service quality monitoring. Second, the new kinds of data open up possibilities for new academic inquiries on travel satisfaction. Finally, the system's public deployment signal the beginning of a mentality shift in customer-engagement and relationship-building in the public transport sector. Collectively, the methodology and institutional innovations aim to contribute towards a better public transport service for the city and its people.

Thesis Supervisor: P. Christopher Zegras
Title: Associate Professor of Urban Studies and Planning

Thesis Supervisor: Jinhua Zhao
Title: Assistant Professor of Urban Studies and Planning

Acknowledgements

Over the last 33 months, there has not been a moment when I have not felt extremely privileged to be where I am, working with and learning from the most inspirational and supportive people. I owe everything to these professors, mentors, colleagues, friends, and family -- no words are adequate to describe my gratitude to you, but I take a stab here anyways.

I owe so much of my growth in graduate school to my advisers Professors Chris Zegras and Jinhua Zhao. Thank you for being the two best teachers, advisers, mentors, and thinkers that one could ever wish for.

> To Chris: Thank you for bringing me to DUSP and the memorable journey with the FMS-TQ project over two continents. I am very grateful for all your support and guidance through every challenge and hurdle. Thank you for always believing in me, even during times when I doubt myself. You have shown me that in research it is okay to embrace the unknowns and take risks, and that a professor is can spend a day driving a pick-up truck all over Boston, hauling a ladder, and getting hands (literally) dirty for a project. Thank you for all the time and attention that you have given to this thesis.

> To Jinhua: Thank you for the freedom to pursue my interest in emerging smartphone technology for urban transportation. I am grateful for the invaluable academic training that you have coached me through: from the e-hail paper to the FMS-TQ project, the many grant proposals, peer reviews, conference presentations, and research meetings. Thank you have pushing me to achieve things that I did not previously know that I could.

When I look back on my time at MIT many years later, I am sure that I will also remember the mentorship from Fred Salvucci and Mikel Murga. I am not only super lucky to be their student, but also to have worked along their side as a teaching assistant. To Fred and Mikel: thank you so much for what you've inspired in me. You have shown me how to build consensus among stakeholders, leverage technical expertise, and work tirelessly towards our visions for our cities (and, the 19-Step Process and 4-Step Model, of course).

I am deeply indebted to the team that has made FMS-TQ a technical reality. I sincerely thank Dr. Fang Zhao, Zhengquan (Robert) Qin, Kalan Vishwa Nawarathne, Dr. Francisco Pereira, Professor Moshe Ben-Akiva, Paola Garbagnoli, and Andrew TONG Kwok Cheong. I especially thank Fang and Zhengquan for their expertise, diligence, and patience, and apologize to Zhengquan for the many late nights that my questions and requests have put him through. I also thank my fellow MCP student and Ayesha Shahid for the comradery and help on the Boston pilot, and Andrew Carrel for his wisdom and advice.

This project would not have been possible without our many institutional partners. I thank our colleagues at the Singapore Land and Transport Authority (LTA), including George

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1   From Snowmaggaddon, an idea was born

The winter of 2015 was an unforgettable one for Boston. Ferocious blizzards slammed the city, dumping more than 110 inches of snow and shutting down this northern city that had been so used to the cold. The region paralyzed along with the transportation system. Travel bans were imposed, buses and subway trains broke down, and the entire public transport system shut down for a total of four days. Even after the shutdowns officially ended, the system struggled with rampant delays. Horror stories of typical 45-minute commutes turning into four-hour journeys abound. The MBTA (Massachusetts Bay Transportation Authority), the region's public transportation operator, became a subject of public criticism and apprehension. The day after the second shutdown ended, the MBTA's General Manager resigned.

This episode highlights a conundrum for public transport operators and regulators. They work hard to keep the bus and trains running, but that scarcely make riders happy. For major systems that serve millions of rides per day across hundreds of routes on multiple modes, it takes monumental coordination and diligence to dispatch and safely run every vehicle and adjust to extenuating circumstances. When service runs smoothly, there are scarce praises or recognition, but when hiccups occur, however minor, complaining voices surface loudly. Sure, there are many possible reasons for this phenomenon – image of public transport as a public good and a right, unpleasant emotions associated with commuting (i.e., going to work), economic impact of a bad trip (e.g. arriving late to work or an important meeting), heuristic biases that render negative experiences more salient than positive ones, etc.. Though many other service sectors encounter similar imbalance of

customer feedback between the positive and negative poles, the phenomenon poses extra issues for public transport operators and planners. First, public transportation affects so many people's lives every day, so the relevant agencies need to have a good understanding of the true state of customer experience in order to bridge any service quality shortfalls. Second, there is scarcely another service that is featured so saliently in the public and political spheres. Given that service quality and customer satisfaction are carefully-tracked metrics used to demonstrate performance and lobby for financial support, public transport agencies have to be able to obtain reliable, representative data on the state of the system with little time lag.

      As the digital age brings shifts in customer preferences and ushers in new possible business models for urban mobility, these changes pose further challenges and opportunities for the public transport sector (Prahalad & Ramaswamy, 2004). Today's customers, growingly more accustomed to the on-demand economy and customer-centric user experience, continue to hold services to ever higher standards. To remain sensitive to consumers' needs, service providers need to increasingly focus on proactively engaging customers to understand their service experience and build stronger relationships. These evolving new norms in the service sector challenge traditional methods and metrics for monitoring and evaluating public transport service quality. Customer satisfaction surveys are administered too infrequently and ask respondents to provide only general, overall ratings. This abstractness reduces the possibilities of associating satisfaction levels to specific trips and their attributes. It also shortchanges the potential to use riders as sources of information, hindering agencies from obtaining more spatially and temporally precise results.

      The research behind this thesis was born out of the desire to help the public transport sector innovate its tools and organizational paradigm. I aspire to develop new tools that would collect actionable data to guide improvement efforts, as well as to showcase a model for empowering customers in co-monitoring the public transport service quality. The unprecedented richness of data would also enable researchers to examine interesting questions related to travel behavior, customer satisfaction, and service performance. With the advance in new sensing and analytics

technologies, I look to leverage distributed data collection to accomplish these purposes. These tools, as I hope, will both enable and be supported by positive institutional changes in how the public transport sector engages and build relationships with its customers.

The development and deployment of the tools, as described in this thesis, has been a team effort. Though I initiated the idea and have managed the whole process, numerous professors, research scientists, engineers, and public transport agency colleagues have been closely involved in the design, development, and pilot of the smartphone platform. Hence, from time to time in this thesis, I will use the pronoun "we" when talking about endeavors that involving collaborators. This highlights the enormous efforts that the team has dedicated to this project, and their valuable contribution that cannot be unacknowledged.

1.2   Research questions

This thesis presents work from the last 17 months in pursuit of following key research questions:

1. How do public transport agencies currently assess and interpret service quality and customer satisfaction metrics? What are the strengths and shortfalls?

2. Can we develop a system that solicits real-time, high-resolution feedback on public transport services, both objectively and from the customers, down to the level of a specific trip, while minimizing user burden?

3. Are there any differences between passenger's real-time satisfaction (reported during the trip) and recalled satisfaction with service (reported long after the trip)?

4. How do passengers' reported satisfaction relate to objective trip attributes (e.g. wait time, travel time, smoothness of the ride)?

1.3 Organization

Chapter 2 synthesizes how the public transport sector currently assesses the quality of their service delivery. It defines the key concepts of quality of service and customer satisfaction, and highlights the connections and differences between them. It also synthesizes prior academic efforts in understanding public transport service quality and customer satisfaction, as well as current practices within the industry. The chapter ends by discussing the opportunities and challenges for smartphone-based innovations in service quality assessment.

Chapter 3 examines the theories, challenges, and potential of engaging customers and citizens in public service delivery. Along with Chapter 2, this chapter completes the two halves of the theoretical foundation for this thesis. Public service agencies can think of soliciting customer feedback as three inter-related objectives: to collect better data, to further institutional learning, and to strengthen relationship with the customers. I will draw examples from both public and private sectors to illustrate these objectives, and discuss the particular value and challenges for the public transport sector.

Chapter 4 describes the technology of the foundational Future Mobility Sensing (FMS) platform and the principles and processes in designing the Travel Quality Survey (TQ) extension.

Chapter 5 places Singapore and Boston – sites of the two FMS-TQ pilots – side-by-side. It compares and contrasts the two regions' urban and public transportation contexts, as well as institutional arrangements and priorities. These factors have significantly shaped the technological and implementation designs of the pilots.

Chapter 6 details the proof-of-concept pilot conducted in Singapore in summer 2015. We put the platform to test for the first time, which proves some of the key capabilities of interests and exposes a number of areas needing significant re-design.

The chapter covers the trial design, pilot implementation, findings, and lessons learned from the pilot.

Chapter 7 presents the Boston pilot conducted in spring 2016. The pilot assumes a design considerably different from that employed in the Singapore pilot. The chapter details the pilot's design, implementation and findings, similar to Chapter 6. It concludes by discussing lessons learned from both pilots.

Chapter 8 reflects on this journey during the last one and a half years – building consensus among diverse stakeholders, designing for academic inquiry and for practical user experience, working with two different public transport agencies, etc. It accentuates the importance of institutional arrangement in determining success of public-facing projects.

Chapter 9 discusses the implications of this research, identifies area for future research, and concludes.

Chapter 2

"So, how are we doing?"

"How satisfied are you with our service?" This question constitutes the timeless yardstick in the service sector. It is not hard to find words such as "service quality" and "customer satisfaction" in surveys, reports, and initiatives of private and public sectors providers alike, though the search for clear definitions or best measuring practices – particularly for public services – is much more arduous. This chapter first defines and distinguishes quality of service and customer satisfaction – two concepts on which this thesis is founded. While the definitions described in this chapter are rooted in the public transport context, the discussion also takes a bigger-picture look at these concepts in the service sector in general. It then synthesizes prior academic undertakings in understanding public transport service quality and customer satisfaction, and presents current practices of measuring these metrics in the public transport industry. The chapter concludes by discussing areas in which the industry should, and could, significantly benefit from new tools, systems, and mentalities in assessing how well they are serving their customers.

2.1   Definitions

"Quality of service"

Public transport literature commonly defines "quality of service" as "perceived performance from passengers' point of view" (Hensher, Stopher, & Bullock, 2003; Kittelson & Associates, National Research Council (U.S.), Transit Cooperative Research Program, United States, & Transit Development Corporation, 2013; Tyrinopoulos & Antoniou, 2008). This customer-oriented emphasis distinguishes itself from traditional performance measurements that focus on system throughput and efficiency – annual passenger trips, vehicle miles, etc. Hensher et al. (2003)

points out that passengers evaluate services in many ways beyond the quantity of service consumed, taking into account also qualitative factors and individual preferences. As the business world often considers customers as the best judges of service (Berry, Zeithaml, and Parasuraman 1990), the term "service quality" in the public transport context has essentially come to mean service aspects that directly influence how passengers perceive their trip and of the provider overall.

Enumerating the defining aspects of service quality is much more difficult, as there is little consensus among scholars and agencies on the best combination of attributes. Studies have measured these concepts with as few as six (Kittelson & Associates et al. 2013) and as many as 31 components (Habib, Kattan, & Islam, 2011); most fall between 8-22 factors (J. de Oña, de Oña, Eboli, & Mazzulla, 2013; Eboli & Mazzulla, 2007, 2010, 2011; European Committee for Standardization, 2002; Friman & Fellesson, 2009; Hensher et al., 2003; Tyrinopoulos & Antoniou, 2008). Of course, this diversity comes partially from the varying levels of specificity and ways to categorize attributes, so overlaps are common.

The most frequently included attributes are (in alphabetical order):

- Accessibility (distance and ease of access to stops);
- Connectivity (whether route is direct, transfer facilities);
- Driver's service;
- Fare;
- Information availability;
- Onboard comfort;
- Reliability;
- Safety;
- Service frequency;
- Service of non-driver staff;
- Stop and station conditions;
- Travel time;
- Vehicle cleanliness;
- Vehicle crowding level; and
- Wait time

17

It should be noted, however, that many literatures blur the line between quality of service and customer satisfaction, so most of the above attributes are discussed in the former context in some literature but the latter context in other works. Indeed these two concepts are tightly connected, and I will further discuss their relationship later in the chapter.

Of all these service quality elements, which ones are the most fundamental? Taking a step back to examine the general concept of service quality, one would discover a wealth of scholarship and debate, stemming largely from the marketing discipline (Babakus & Boller, 1992; Brown, Churchill, & Peter, 1993; Cronin & Taylor, 1992) The hallmark constructs of service quality is the SERVQUAL model, conceptualized by Parasuraman, Zeithaml, and Berry (1988). The authors conceptualize "service quality" as a comparative function between consumer expectations and actual service performance (the "Gap Model"). Through stated preference studies of multiple service industries, they identify five pillars of service quality, under the acronym RATER:

- Tangibles: Physical facilities, equipment, and appearance of personnel;
- Reliability: Ability to perform the promised service dependably and accurately;
- Assurance: Knowledge and courtesy of employees and their ability to inspire trust and confidence;
- Responsiveness: Willingness to help customers and provide prompt service; and
- Empathy: Caring, individualized attention the firm provides its customers.

With a few adjustments, this five-dimension framework can be adapted to fit the public transport context (Elmore-Yalch, 1998; Tripathi, Kumar, & Gunjan, 2012). Table 2.1 displays my own adaptation of the RATER model. The "Empathy" dimension is perhaps the trickiest, since mass transit service is, arguably, by nature at odds with delivering intimate, personalized services to customers. Some scholars hence propose excluding the "Empathy" category (Barabino, Deiana, & Tilocca, 2012; Too & Earl,

2009) when assessing public transport quality. I believe, however, that empathy is still possible through relationship-building: understanding the customers, seeking feedback on their experiences and input for improvement. While enabling this capacity for empathy in public transport services resembles a paradigm shift, it is an increasingly relevant priority. The so-called Transportation Networking Companies (TNCs) – providers of app-based, on-demand ride services – have been rapidly gaining mode share, and one of the key reasons for their popularity is the more amicable interactions between customers and drivers (Li & Zhao, 2016). This thesis is largely inspired and motivated by the potential to drive this exact paradigm shift in the public transport sector through technological innovations, which I will discuss further in the next chapter.

Table 2.1 - Public transport quality attributes under the SERVQUAL model

Service quality aspects	
Tangibles	Accessibility to stop/station
	Stop/station facilities
	Wait time
	Travel speed/time
	Seating & personal space
	On-board comfort
	Directness of route/ need for transfers
Reliability	Reliability of wait time
	Reliability of being able to board bus/train
	Reliability of travel time
Assurance	Availability & accuracy of information
	Bus driver's skills
	Attitude and quality of customer service
	Sense of safety & security
Responsiveness	Assistance to customers when needed
Empathy	Understanding of customer experiences
	Personalizing customer relationships

"Customer satisfaction"

The term "customer satisfaction" has become synonymous with "service quality" in the public transport sector. Given that service quality is by definition customer-centric, assessing it typically involves customer satisfaction measurements (Eboli and Mazzulla 2009). Likewise, results of customer satisfaction surveys have become the de facto service quality indicators for many public transport agencies – I will reference multiple agencies' approaches later in this chapter. Despite its prevalence, the concept of customer satisfaction has scarcely been defined clearly in the public transport literature. Most surveys simply ask passengers to rate their satisfaction with the overall service and individual attributes on a Likert scale (Elmore-Yalch, 1998), and interpret the numeric value as the level of satisfaction.

Is equating "satisfaction" with "service quality" justified? As with "service quality," most of the prior scholarship on customer satisfaction comes from the market research world. The "Disconfirmation Paradigm," which highly resembles the "Gap Model" established by the service quality literature, views satisfaction as an emotion resulting from confirmation or disconfirmation of expectations (Cronin, Brady, & Hult, 2000; Cronin & Taylor, 1992; Oliver, 1981; Parasuraman et al., 1988; Rust & Zahorik, 1993). Namely, a consumer is satisfied when 1) the service experience is as expected and her expectations are confirmed, or 2) the experience is better than expected and thus positively disconfirms the expectations. On the flip side, dissatisfaction arises when service is below expectation (Churchill & Surprenant, 1982).

Given these similarities, numerous researches have sought to disentangle the relationship between "service quality" and "customer satisfaction." There is no clear consensus among the debate. Some theoretical papers distinguish satisfaction and quality by their timeframe – specific, short-term evaluation vs. more general, long-term judgements – though even scholars who prescribe to this distinction have conflicting views on the corresponding timeframe for each concept (Bitner & Hubbert, 1994; Oliver, 1993; Parasuraman, Zeithaml, & Berry, 1985). Subsequent

empirical studies find that people do not clearly distinguish the two concepts from each other – their assessment of service in respect to both "quality" and "satisfaction" posit a relative judgement of experience versus expectation and result in future purchase intentions as a consequence (Iacobucci, Amy, & Kent, 1995). Other scholars, assuming distinction between satisfaction and service quality, have looked at how the two concepts inter-relate. Among their research, there is a partial consensus that "customer satisfaction is the result of a customer's perception of the value received, where value equals perceived service quality relative to price"(Cronin, Brady, and Hult 2000). This is to say, in case a customer perceives a service experience mediocre in quality, she might still be rather satisfied if she feels she has gotten a good "bang for the buck."

Synthesizing the above literatures in the public transport context, I have decided to conform to the industry practice and use customer satisfaction as a measurable indicator for service quality. There is no strong evidence that these two terms evoke significantly different connotations, especially for the average rider. Even if we posit that there are granular differences at a theoretical level, the constant and low price of a bus or subway ride (between $1 - $3 in the U.S.) would only lessen the distinction between satisfaction – which has an associated value assessment -- and service quality. So in the public transport context, asking passengers the question "how satisfied are you with our service?" may in fact well answer the question "how are we doing in servicing our customers?"

The importance in studying and improving ways to assess service quality and customer satisfaction are two-fold. First, public transportation, being a service industry, has an inarguable responsibility to ensure that (most of) those whom it serves are happy with the service. Second and more practically, service quality and customer satisfaction have been proven to shape behaviors – this could affect ridership and the health of the system. Numerous studies find that perceived service quality is one of -- if not the most – important determinants of customer satisfaction (Cronin et al., 2000; Eboli & Mazzulla, 2011; Fornell, Johnson, Anderson, Cha, & Bryant, 1996), and these two indicators are affirmed, by an even larger body of

literature (Table 2.2), to affect the customers' intention for future patronage. As many new urban mobility services have emerged over the recent years, people now have more choices than ever when making a trip. For public transport to improve its appeal and maintain its relevance in urban mobility, it would have to accord higher priority to better understanding and improving users' experiences. The next section summarizes the efforts in understanding public transport service quality and customer satisfaction to date.

Table 2.2 - Literature linking quality, value, and satisfaction to customer behavioral outcomes

Source	Relevant Constructs	Link(s) to Outcomes	Empirically Tested?
Parasuraman, Zeithaml, and Berry (1988)	SQ, BI	SQ	Yes
Parasuraman, Berry, and Zeithaml (1991)	SQ, BI	SQ	Yes
Anderson and Sullivan (1993)	SQ, SAT, BI	SQ, SAT	Yes
Boulding et al. (1993)	SQ, BI	SQ	Yes
Taylor and Baker (1994)	SQ, SAT, BI	SQ	Yes
Zeithaml, Berry, and Parasuraman (1996)	SQ, BI	SQ	Yes
Taylor (1997)	SQ, SAT, BI	SQ, SAT	Yes
Athanassopoulos (2000)	SAC, SQ, SAT, BI	SQ	Yes
Cronin and Taylor (1992)	SQ, SAT, BI	SAT	Yes
Anderson and Fornell (1994)	SQ, SAT	SAT	No
Gotlieb, Grewal, and Brown (1994)	SQ, SAT, BI	SAT	Yes
Ostrom and Iacobucci (1995)	SAC, SQ, SAT, VAL, BI	SAT	Yes
Fornell et al. (1996)	SQ, SAT, SV, BI	SAT	Yes
Patterson and Spreng (1997)	SAT, SV, BI	SAT	Yes
Hallowell (1996)	SAT, BI	SAT	Yes
Andreassen (1998)	SQ, SAT, SV, BI	SAT	Yes
Bolton (1998)	SAT, BI	SAT	Yes
Chenet, Tynan, and Money (1999)	SQ, SV, SAT, BI	SAT	No
Oliver (1999)	SAT, BI	SAT	No
Garbarino and Johnson (1999)	SAT, BI	SAT	Yes
Bolton and Lemon (1999)	SAT, BI	SAT	Yes
Bernhardt, Donthu, and Kennett (2000)	SAT, BI	SAT	Yes
Ennew and Binks (1999)	SQ, SV, SAT, BI	SAT, SV	Yes
Zeithaml (1988)	SAC, SQ, SV, BI	SV	No
Bolton and Drew (1991)	SQ, SAT, SV, BI	SV	No
Gale (1994)	SQ, SV, BI	SV	No
Chang and Wildt (1994)	SAC, SQ, SV, BI	SV	Yes
Hartline and Jones (1996)	SQ, SV, BI	SV	Yes
Wakefield and Barnes (1996)	SQ, SV, BI	SV	Yes
Cronin et al. (1997)	SAC, SQ, VAL, BI	SV	Yes
Sirohi, McLaughlin, and Wittink (1998)	SAC, SQ, SV, BI	SV	Yes
Sweeney, Soutar, and Johnson (1999)	SAC, SQ, SV, BI	SV	Yes

Source: Cronin, Brady and Hult (2000).
SQ = service quality; BI = behavioral intention, SAT = satisfaction, SAC = sacrifice, SV/VAL = service value

2.2 Understanding public transport service quality and customer satisfaction

Academic research on these topics has overwhelmingly focused on identifying attributes that most strongly influence riders' satisfaction of local public transport services. Most of these efforts aim at testing a particular model and devising a new service quality indicator ever-so-slightly different from the existent types. The studies are best differentiated along two dimensions: data collection and analysis method. As the results are heavily tied to the studies' context – the status quo of the public transport system, the local culture, etc. – there is little convergence among the findings.

One of the most common approaches involves asking respondents to rate the importance of various service attributes (e.g. wait time, comfort), their level of satisfaction with those attributes, and their satisfaction with the service overall (Efthymiou, Kaziales, Antoniou, & Tyrinopoulos, 2014; Tyrinopoulos & Antoniou, 2008). The main advantage of these so-called "stated importance" methods is ease of interpretation – it is simpler to obtain the attribute's importance, and results are more intuitive (Weinstein 2000). But it has also some critical disadvantages. First, it is susceptible to the so-called "Top Box Problem." Most customers, when asked about the importance of an attribute, tend to rate it as either "very important" or "important" – ticking off one of the top boxes on the answer sheet (Morpace International et al. 1999). As Weinstein (2000) notes, passengers may rate police presence on buses as "very important", but it may have little to do with their overall satisfaction with the service. This phenomenon mutes differentiation in importance ratings among the factors of interest, potentially masking truly significant drivers of satisfaction from being identified. Furthermore, stated importance methods add to the length of questionnaires – respondents would have to state the importance *in addition* to rating the satisfaction level/perceived performance for each service attribute (Weinstein, 2000).

A related approach – derived importance measures – purports to overcome the above mentioned disadvantages. It asks respondents to rate only their satisfaction

with individual service attributes (e.g. wait time, comfort) and the overall service (Budiono, n.d.; Celik, Aydin, & Gumus, 2014; J. de Oña et al., 2013; Del Castillo & Benitez, 2012; Eboli & Mazzulla, 2007, 2010; Habib et al., 2011). Researchers then determine the importance of each attribute from statistical association between individual attribute ratings and the overall satisfaction rating (Morpace International et al. 1999). This method mitigates the "Top Box Problem" and involves shorter questionnaires, and hence is often preferred by researchers and academics (Weinstein, 2000). The downside, however, is the analytical complexity involved – the results are not always intuitive or easy to explain to stakeholders (R. de Oña, Eboli, & Mazzulla, 2014; Weinstein, 2000).

The last major approach attempts to quantify importance of attributes through stated preference experiments. Instead of asking passengers for ratings of individual attributes, this approach examines passengers' expressed choices among hypothetical sets of alternatives with varying service attributes (Cantwell, Caulfield, & O'Mahony, 2009; Eboli & Mazzulla, 2008, 2010; Prioni & Hensher, 2000). The most seminal work on public transport quality using this approach is by Hensher et al (2003), which presents to respondents 27 "bus packages" – service scenarios with varying levels of performance for 13 attributes. The average number of passengers choosing a given bus package is used to approximate the overall perceived satisfaction with the service scenario; the contribution of each attribute to the overall satisfaction can also be computed through discrete choice models.

The data analysis that ensues is largely divided into two camps. The majority of studies seek to quantify the relationships between reported satisfactions of service overall and of individual attributes. For this research question, multiple regression analysis is the most common analytical technique (Kim & Lee, 2011; Weinstein, 2000) The overall reported satisfaction level is regressed on satisfaction rating for individual attributes to tease out their significance. (Kim & Lee, 2011; Weinstein, 2000)Some studies, particularly those that include a large number of service attributes in their explanatory set, first perform structural equation modelling to identify potential latent variables, reducing the many service attributes tested into a

smaller numbers of dimensions that contribute to satisfaction measures (J. de Oña et al., 2013; Eboli & Mazzulla, 2007; Karlaftis, Golias, & Papadimitriou, 2001; Shaaban & Khalil, 2013; Stuart, Mednick, & Bockman, 2000). Studies that collect data through stated preference surveys tend to employ discrete choice models in analyzing service attributes' importance for customer satisfaction (Eboli & Mazzulla, 2010, 2011; Hensher et al., 2003; Prioni & Hensher, 2000). The second, smaller group of studies attempt to discern potential heterogeneities in customers' assessment of public transport services. Wallin and Andreassen (1995) find that high- and low- frequency public transport users have different preferences; Eboli and Mazzulla (2009) propose a Heterogeneous Customer Satisfaction Index that improves upon the Customer Satisfaction Index by taking riders' heterogeneities in perception into account.

Given the diversity of service attributes examined by various studies, their findings offer little defining insights altogether. A wide range of service factors has been found to be important in customer evaluations of public transport service quality (Stathopoulos & Marcucci, 2014). A partial list includes: reliability (Friman, Edvardsson, and Gärling 2001); frequency, travel time and fare level (Hensher et al., 2003); comfort and cleanliness (dell'Olio, Ibeas, & Cecin, 2011; Eboli & Mazzulla, 2007); stops and waiting environment (Iseki & Taylor, 2010); safety issues (Friman and Fellesson 2009); network coverage/distance to bus stops (Tyrinopoulos & Antoniou, 2008), etc. If an agency were to look to these literatures for insights on service quality or customer satisfaction, it may be somewhat disappointed. Given the local nature of the results, one cannot simply generalize and reference the factors deemed significant in any study as the key characteristics to monitor within a particular public transport agency. In terms of methodology, since rarely does a study apply and compare two approaches on the same data set, there is little insights on which method works best – and under what circumstances. Furthermore, much of the scholarly attention is placed on experimenting with new modelling and score-building techniques, rather than exploring new methods to collect customer satisfaction data in more efficient and meaningful ways. This latter pursuit is arguably more relevant to public transport agencies.

25

2.3 Current industry practices

In terms of measuring perceived service quality and customer satisfaction, the state of the art within the public transport sector may not be emblematic of the 21st century. In the U.S., the most prevalent channels are intercept surveys (onboard or at stops/stations), telephone surveys, and web-based surveys. Many agencies use multiple, complementary channels to capture a wider respondent pool. Intercept supplemented with telephone survey has been the most popular method, although an increasing number of agencies have adopted web-based tools in recent years. In Singapore, the country's Land and Transport Authority (which plans, regulates, and oversees public transport services), conducts intercept surveys of riders at stops/stations. There is no known case of agencies routinely leveraging newer technologies, such as smartphones[1], to conduct customer satisfaction surveys. Surveying frequency varies among agencies. In Spitz et al (2006)'s examination of survey practices among U.S. transit agencies, about 45% conduct customer satisfaction assessment one or more times a year, 30% less than once a year, and 20% have never done one.

Typically, public transport agencies carry out repeated cross-sectional sampling of customers. Each cross-section measurement thus resembles the one-time academic exercises discussed earlier. Akin to in academic studies, these questionnaires also commonly ask respondents to rate their satisfaction with overall service and its individual attributes – wait time, reliability comfort, etc. The number of service attributes presented and the level of detail vary widely. In a TCRP review, the number of attributes measured by agencies ranges from as few as five and as many as 48 (Morpace International et al. 1999). Agencies may also solicit passengers' reported importance of various service attributes. Evaluations are collected periodically, and ratings are averaged to generate scores for historical comparisons among cross-sections. This is known as the "scorecard approach." Agencies may

---

[1] Smartphone: a cellular phone that performs many of the functions of a computer, typically having a touchscreen interface, Internet access, and an operating system capable of running downloaded applications. (source: Oxford Dictionaries)

publish these scores to demonstrate the progress (or regress) made over the recent periods; some agencies, such as NJTransit, even call it "The Scorecard" (Figure 2.1). The idea is that public transport services can be aggregately measured and tracked from time to time, as, for example, with a student's academic performance.
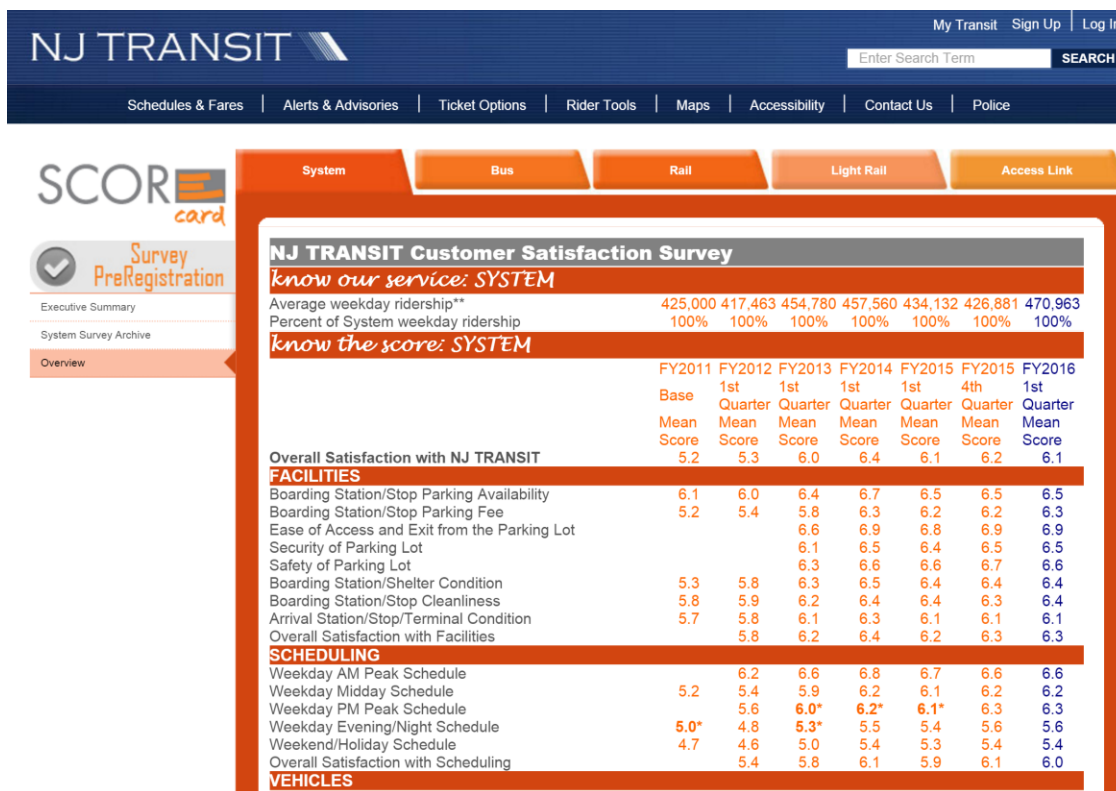
Figure 2.1 – NJ Transit's Scorecard

The primary limitation of this scorecard approach is the incapability for ubiquitous, detailed assessment and feedback (Dunlop, Casello, & Doherty, 2015; Elmore-Yalch, 1998). Considering that public transport customer satisfaction surveys are often done at an annual (or lower) frequency, questionnaires often ask respondents for general assessments of the service (Carrel, Mishalani, Sengupta, & Walker, 2015). This puts regulators and system administrators at a disadvantage in obtaining high-resolution data – information that can reveal performance variations by driver, route, and time of day, as well as the precise areas for commendation and targeted improvement (Elmore-Yalch, 1998). For example, according to NJTransit's

Scorecard, customers' overall satisfaction decreases from 6.4 in 2014 (first quarter) to 6.1 a year later, and satisfaction scores have dropped for virtually all service attributes (NJ Transit 2015). But what does this mean for the agency? Are the differences significant? If so, to which areas should the resource-strapped agency devote its capital and personnel for improvement? Even if a few attributes – say, cleanliness, service frequency – have been identified as priority areas, it still reveals little information as to which routes, vehicles, and/or stops should be improved. Of course, traditional in-person questionnaires, administered during or after the ride, can solicit trip-specific assessment, but their scopes are limited without a major commitment to staffing. Public transport experiences can vary from trip to trip, but it would not be realistic to deploy survey teams to every bus, stop, and station every day.

Some agencies have attempted to go beyond the general service assessment by asking their customers about specific trips in gauging service quality. For example, the Massachusetts Bay Transportation Authority (MBTA) asks their Customer Opinion Panel members to choose a particular recent trip to base their reflections on (Boroyan, 2014). Singapore's Land Transport Authority (LTA) asks respondents to reflect on their most recent bus or subway trip in answering the customer satisfaction surveys (Land Transport Authority, 2016c). However, such recalled assessment may be inaccurate, as people's actual and recalled experiences often differ due to psychological heuristics or unobserved events (Fredrickson & Kahneman, 1993). Pedersen et al (2011), by recording 62 volunteers' predicted, experienced, and remembered satisfaction of bus and subway trips for a month, reveal that recalled satisfaction is significantly lower than experienced satisfaction. Abou-Zeid et al (2012) observe a similar bias, noting that subjects report lower satisfaction with public transport after experiencing a commute by automobile. This phenomenon implies that, the sooner that a rider is given the chance to report her level of satisfaction, the more accurate she is able to report the extent of disconfirmation. Recently, and as detailed further below, some public transport agencies have partnered with academics to attempt to shorten the lag between the ride and survey

solicitation with new technological tools. Our project, FMS-TQ, is a further attempt at capturing more accurate, high-resolution, and actionable customer feedback on public transport experience.

Another approach uses panel-based instead of cross-sectional surveys. Panel surveys are advantageous for several reasons. They reduce the chance of confounding differences in measured satisfaction across years with differences between cross-sectional samples (Spitz et al., 2006). Since changes in behavior of each individual are directly observed, the sample sizes required to measure differences in customer satisfaction can be much lower than the cross-sectional approach. Panel surveys also allow for the costs of recruiting individuals to complete surveys to be spread out over multiple survey periods (Chow, 2014). A number of agencies – NJTransit, Metrolink (Los Angeles), and GO Transit (Greater Toronto and Hamilton, Canada) – tracks customer satisfaction with panels through the use of web-based surveys (Chow, 2014; Spitz et al., 2006).

The prevalence, merits, and shortcomings of the common customer satisfaction survey channels -- intercept, telephone, web, and hybrid methods – are described below. The discussion will reference practices in Singapore by the Land and Transport Authority (LTA), as well as at the top 10 U.S. transit systems (as measured by unlinked passenger trips): New York City Transit (NYCT), Chicago Transit Authority (CTA), Los Angeles County Metropolitan Transportation Authority (LACMTA), Washington Metropolitan Area Transit Authority (WMATA), Massachusetts Bay Transportation Authority (MBTA), Southeastern Pennsylvania Transportation Authority (SEPTA), New Jersey Transit Corporation (NJ TRANSIT), San Francisco Municipal Railway (MUNI), Metropolitan Atlanta Rapid Transit Authority (MARTA), and MTA Bus Company (MTABUS). Table 2.3 summarizes the practice, which shows that even among the top agencies, data collection approaches are divided somewhat equally among intercept, web, and telephone channel.

Table 2.3 - Service quality/customer satisfaction survey methods used in Singapore and by top 10 U.S. transit systems

Agency	Most recent survey (publically available)	Intercept	Web-based	Phone	Other
Singapore					
LTA	2015 Public Transport Customer Satisfaction Survey	X			
United States					
NYCT	2014 Customer Satisfaction Survey			X	
CTA	2013-2014 RTA Customer Satisfaction studies	X	X		
LACMTA	Spring 2015: Metro Rail/Bus/System-wide Customer Satisfaction Survey	X		x	Focus groups
WMATA	Voice of the Customer survey (1st Quarter, FY 2013)			X	Mystery rider
MBTA	Customer Opinion Panel		X		
SEPTA	2012 Customer Satisfaction Survey	X		X	
NJTransit	ScoreCard (2nd Quarter, FY2016)		X		
MUNI	2013 MUNI On-board Customer Survey	X	X		
MARTA	2014 Quality of Service Survey	X			
MTABUS	Unknown				

2.3.1 Intercept survey

The most common method for measuring service quality and customer satisfaction is intercept survey, either onboard or at stops/stations. The LTA has employed this method annually for Singapore's public transport services since 2006 (Land Transport Authority, 2016c). Of the 27 U.S. transit agencies and Metropolitan Planning Organizations (MPO) surveyed in 2006, three-quarters report using intercept surveys (Spitz et al., 2006). Under this approach, questionnaires are distributed to passengers onboard vehicles or waiting at stops or stations. Questionnaires can be in paper format, which respondents may complete and return either in-person or mail back later. Onboard surveys can also be conducted as personal interviews, in which case a surveyor asks riders a short series of questions and records the answers. With the aid of technology,

interviewers may also administer surveys on tablets, digitally recording the results directly (Agrawal, Granger-Bevan, Newmark, & Nixon, 2015; Ching, 2012; Schaller et al., 2005). In addition to customer satisfaction, intercept surveys are also used by agencies to collect information on rider demographics, origin and destination of the trip, and reasons for taking public transport (Schaller et al., 2005). The exact types and wording of questionnaire, as well as survey frequency, vary among agencies. Large agencies typically conduct five or more on-board/intercept surveys annually, primarily focused on specific routes or geographic areas. Smaller agencies typically conduct surveys every one to three years, often covering the entire network (Schaller et al., 2005).

Singapore's Public Transport Customer Satisfaction Survey (PTCSS), conducted annually by the LTA, interviews close to 4,000 regular bus and subway commuters at stops and stations. Respondents are asked to rate their level of satisfaction with, and importance of, eight service attributes for their recalled last public transport trip. The satisfaction ratings are weighed by their respective importance measures to produce an overall satisfaction rating for the bus and subway services (Land Transport Authority, 2016c).

Table 2.4 - Results of the 2015 Singapore PTCSS

PT Service Attributes	Satisfaction Ratings		Satisfied (%)	
	2014	2015	2014	2015
Waiting Time	6.3	6.7	69.1	76.8
Reliability	6.9	7.1	79.7	84.0
Service Information	7.3	7.3	85.2	86.9
Bus Interchange/ Bus Stop/ MRT Station Accessibility	7.5	7.4	89.6	90.2
Comfort	7.0	7.1	82.2	84.0
Travel Time	6.9	7.1	82.4	85.4
Customer Service	7.2	7.3	84.0	86.0
Safety and Security	7.5	7.7	88.9	90.0
Overall Satisfaction	**7.1**	**7.2**	**91.3**	**91.8**

Source: Land Transport Authority, 2016

Most of the top 10 U.S. transit systems also use intercept survey to gauge service quality and customer satisfaction, though many of them use it in parallel with telephone- or web-based questionnaires as well. The Los Angeles County Metropolitan Transportation Authority (LACMT) stands out as the only Top 10 agency that uses onboard surveys as the primary method for gauging customer satisfaction. Since 2001, the LACMT has distributed paper surveys once a year on buses and trains. The questionnaire focuses on service quality indicators such as on-time performance, operator courtesy, cleanliness, and overall satisfaction with service. It boasts a high response rate of over 50%, with 15,000 to 20,000 surveys usually completed every year (LACMT 2016). The onboard survey effort is supplemented by focus groups, which provide more in-depth qualitative information about one or two specific topics – ticketing vending machine redesign, using real-time information, etc. (LACMT 2016). Likewise, the Metropolitan Atlanta Rapid Transit Authority (MARTA) conducts its Quality of Service Survey on an almost annual basis by randomly sampling from patrons on rail platforms and on fixed-route buses. During fiscal year 2014, the agency collected 6,512 responses in total (MARTA, n.d.).

The prevalence of intercept surveys stems largely from their advantages in gaining direct access to customers and obtaining relatively representative samples. Surveys can be conducted on particular lines, or at specific locations or times, to

examine the service quality and customer satisfaction in detail. For instance, the King County Metro of Seattle, Washington has surveyed only riders in the downtown Ride Free Area (Schaller et al., 2005). The Chicago Transit Authority (CTA) has conducted one survey of riders on all bus and rail routes on the West Side and another on the Douglas Line segment of the Blue Line (Schaller et al., 2005). The stratified samples, combined together, can render a highly representative cross-sectional view of the entire system. In fact, for municipalities where only a small percentage of the population uses public transport, these advantages make intercept surveys the most cost-effective way to gather information (Schaller et al., 2005).

On the other hand, conducting intercept surveys for an extensive system can be resource intensive. Staff is needed to administer and retrieve every individual survey, and then enter data into a database. Surveying a whole network, or even a large number of lines, may take several months (Chow, 2014). Given the personnel involvement required, the frequency of surveying – and hence the comprehensive and continuity of information collected – may be limited. Public transport experiences can well vary from trip to trip, but it would not be realistic to deploy survey teams to every bus, stop, and station every day. Of course, this last shortcoming is relevant to all conventional public transport survey methods, rather than being unique to intercept surveys.

2.3.2   Telephone-based survey

Telephone-based methods have been popular among public transport agencies in collecting customer feedback, and often used as a supplement to intercept surveys. Respondents are recruited in-person, by telephone, or via other channels (such as mailing), and once indicating interest in participating, an interviewer contact the respondents by phone. Since the 1990s this is usually conducted with the use of computer-assisted telephone interviewing, CATI. The interviewer reads questions from a computer screen to the respondent, and type the responses into the computer.

Depending on the logical flow of the questionnaire, CATI can skip parts of the questionnaire and bring up the intended question based on the response entered. Since responses are immediately entered and recorded into a computer, CATI eliminates manual data transfers and reduces transcription errors. Its weaknesses, on the other hand, lie in the difficulty in recruiting a representative or targeted sample, especially in municipalities where a small percentage of the general population are public transport users. Given these weaknesses, some guides discourages smaller public transport agencies from using CATI for surveying customers (Baltes, 2002).

Of the Top 10 U.S. transit agencies, three use telephone surveys as the main service quality assessment tool: the New York Metropolitan Transportation Authority (NYMTA), the Washington Metropolitan Area Transit Authority (WMATA), and the Southern Pennsylvania Transportation Authority (SEPTA). The 2014 NYMTA Customer Satisfaction Survey, for instance, interviewed over 1,800 New York City residents using a random sample of landline and cell phone numbers. In-depth interviews were conducted with 1,200 customers who took at least one ride in the past 30 days. The interviews were conducted in English, Spanish, and Chinese, and lasted on average 27 minutes (New York City Transit 2014). The WMATA Voice of the Customer survey administers a 17-minute phone survey to approximately 770 customers each quarter. Similar to the NYMTA survey, it randomly samples from landline and cellphone numbers, and respondents need to have taken the local public transport service in the past 30 days in order to be eligible. WMATA also conducts a separate phone survey of its paratransit users on a semi-annual basis (Washington Metropolitan Area Transit Authority 2014). SEPTA's Customer Satisfaction Survey samples households in the service region by phone, and asks respondents about their public transport experiences in the previous 7 days (SEPTA 2013).

2.3.3   Web-based survey

Web-based method is a late entrant into the public transport survey realm. In 2006, when the Transit Cooperative Research Program (TCRP) published the most

recent report on this matter in the sector, just 25% of the 36 agencies interviewed had used web surveys to collect customer satisfaction information (Spitz et al., 2006). Schaller (2005)'s sampling of 52 U.S. public transport agencies found 44% of organizations using web surveys for any purpose. However, most of these practices still relied largely on in-person or telephone recruitment – only 6% of the agencies surveyed recruited respondents by e-mail or with a web link (Spitz et al., 2006).

The primary advantage of web-based surveys is time and cost efficiency. In Spitz et al (2006)'s report, 70% of respondents who were using web-based surveys then cited "fast turn-around" and "cost effectiveness" as their motivations for employing such tools. Several studies have found that online surveys have the ability to collect a large sample more quickly at a lower cost compared to traditional surveys (Chow, 2014). The cost saving stems not only from the reduced need for in-the-street fieldwork, but also from the faster turn-around and less effort for data entering and cleaning (Schaller et al., 2005). Web-based method have other valuable advantages, allowing for presentation of complicated subject matter, question design, and graphics, and for strict ordering of the questions to be completed (Chow, 2014; Evans & Mathur, 2005; Schaller et al., 2005). Web-based methods also provide for convenience in tracking respondents over time for panel studies, as people's email addresses tend to change infrequently.

The main concern for using web-based technology for surveys has been biases against population with no, or limited, access to the Internet. In Schaller's 2005 report, nearly all agencies that used web-based surveys reported worrying about not being able to reach a reliable cross-section of their audience (Schaller et al., 2005). The saliency of this disadvantage has since decreased, as Internet access has much proliferated over the past decade. Web-based channel has become the dominant or sole survey method for several of the largest transit agencies in the U.S. Since 2011, NJTransit has been conducting system-wide online customer satisfaction surveys, with the most recent 18th tracking period completed in November and December 2015 (NJ Transit 2015). The Massachusetts Bay Transportation Authority (MBTA) gathers ongoing feedback through a web survey, which is sent to a rotating third of

their customer panel every month (Boroyan, 2014). These initiatives within leading public transport agencies have shown web-based surveys as a viable method, and we are likely to see more adoption of this approach in the future.

2.3.4   Other methods

Aside from the techniques above, agencies may combine multiple methods as complements. For example, when the Regional Transportation Authority last conducted a customer satisfaction survey for three agencies in the Chicago region, it leveraged the web via email invites as well as onboard recruitment efforts (RSG 2015). The San Francisco Municipal Railway (MUNI) conducts quarterly online customer satisfaction surveys, sampling from an opt-in online panel; in addition, it also conducts a multilingual onboard survey (last completed in 2013) of over 22,000 customers (SFMTA 2014).

Intercept interviews are also often used to gather names, telephone numbers and/or email addresses for phone or web-based surveys later. For example, for its Customer Satisfaction Survey, SEPTA intercepts riders at various locations to request their contacts, but also supplements the pool with a random telephone sampling of households in the region (SEPTA 2013). Such a combination is quite suitable where incidence of public transport user is too low for digital dial telephone interviewing, yet the survey is too long or complex for onboard and intercept interviews. As Schaller points out, the combination is particularly cost-effective for commuter rail, for which a large percentage of riders pass through a downtown terminal during specific time periods (Schaller et al., 2005).

Another once-common method is mailing. Questionnaires are randomly sent to potential respondents to complete and return by post. With the proliferation of telephone and Internet access, as well as the emergence of new technologies, however, there remains essentially no advantage to mail surveying. Turnaround is longer, and surveyors likely need to send a follow-up mailing to increase the overall

response rate (Baltes, 2002). As a result, none of the leading public transport agencies examined in this thesis still use this method to collect customer satisfaction data.

One final approach to assessing service quality is a Passenger Environment Survey (PES), conducted using the "secret shopper" technique. Developed by the New York City Transit[2] in 1983 and since majorly revised, PES is an internal performance audit of the passenger experience in vehicles and stations. Different from the methods described earlier, where ordinary riders constitute the source of feedback, secret shoppers are usually trained personnel who use the service as a normal customer, and rate the attributes from the perspective of a regular passenger (Eboli and Mazzulla 2012). The advantage is that these trained "customers" have common standards against which to rate the service. This is a tradeoff against representativeness of responses, however, and it technically defies the very customer-centric nature of service quality surveys.
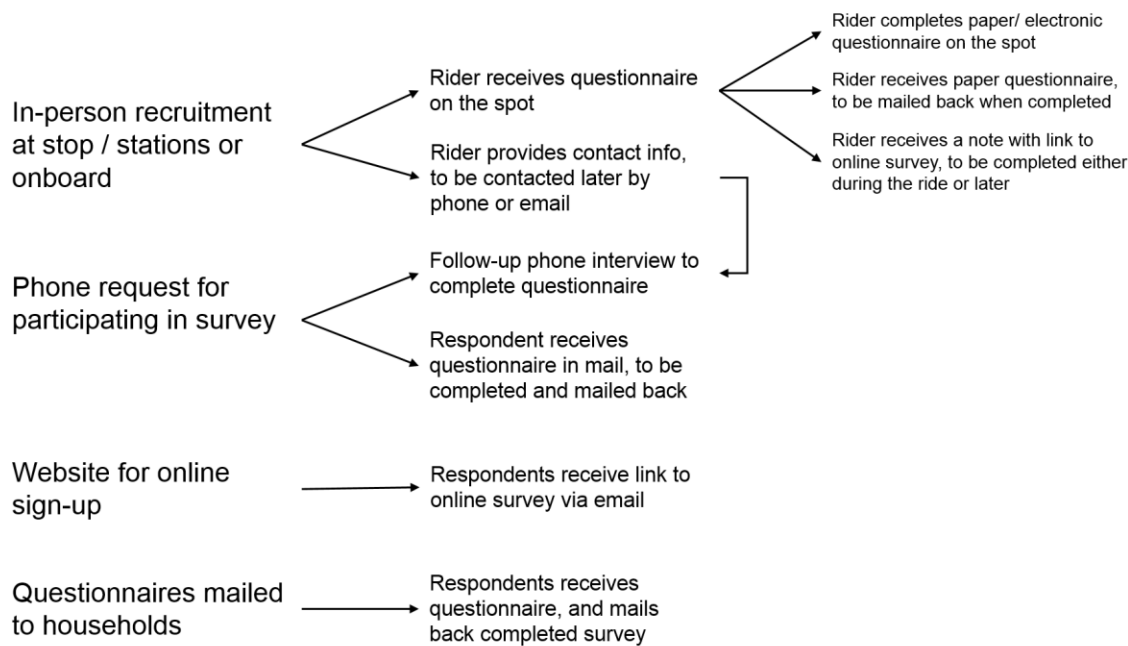
Figure 2.2 - Flows of conducting various methods of customer surveys

---

[2] The New York City Transit (NYCT) is one of the six agencies of the New York Metropolitan Transportation Authority (NYMTA), operating the New York City Subway, Staten Island Railway, and New York City Bus systems.

The review of industry practices reveals that service quality and customer satisfaction assessment remains at an aggregate level. Public transport operators and regulators largely collect and treat customer satisfaction measurements as a scorecard, aggregating ratings provided by respondents on the service in general. This does not fulfill the full value of customer satisfaction inquiry, as much pertinent information on the day-to-day public transport experience may be lost in the surveying process. Recognizably, this inefficiency is a result of the limitations of the tools available to operators. Although intercept, telephone, and web-surveys can obtain a fairly representative sample and approximate general public sentiments, none of these methods are cost-effective or practical in obtaining high-resolution information.

2.4 Role of smartphones in advancing public transport service quality assessment

In recent years, the attention has turned to smartphones as a potential new medium for data collection. The comparative advantages of smartphone-based surveys include information richness, real-time speed, cost efficiency, and ubiquity. Surveyors can obtain details about the trip experiences from smartphones and their sensors – precise time and location information (GPS, Wi-Fi, cell tower), vehicle acceleration profile (accelerometer), and ambient temperature (thermometer) and noise (microphone). The geospatial and temporal information can help associate the collected data to specific service runs or stations/stops. As people tend to carry their phones with them most of the time, information can be collected and transmitted to the server in real-time. Such surveys may reduce the need for mobilizing field surveyors, reducing the time and cost of survey administration. They can also be deployed over extended periods of time and space, enabling intra-day, inter-day and inter-seasonal assessments for numerous routes, stops, segments, etc. Last but not least, well-designed smartphone apps can enhance the user-friendliness and

interactivity, potentially offering a better respondent experience than traditional surveys.

No public transport agency has yet adopted the smartphone approach in their routine service quality and customer satisfaction assessment, but several recent academic studies have explored using the technology to collect more detailed data on rider experience. In 2012, an MIT research team launched the Flocktracker, a smartphone app for trained surveyors. The team initially conceptualized the platform to capture bus riders' experiences in Dhaka, Bangladesh. Surveyors would board specific bus routes, use the smartphone app to record bus positions, record crowding conditions, survey riders about their satisfaction of the ride, and input these answers into the app (Ching, 2012). Two subsequent MIT studies adapted the Flocktracker to collect bus riders' opinions regarding perceived security along routes in Mexico City, Mexico and St. Louis, Missouri, respectively (Butts, 2014; P. C. Zegras, Butts, Cadena, & Palencia, 2015). Despite the differences in local settings and survey topics, these three projects shared one key similarity -- their data carried detailed geolocation and time stamps, hence allowing researchers to analyze and visualize spatial and temporal patterns.

As smartphone penetration rate increases, data collection can be crowdsourced by empowering individual riders as sensors. Tiramisu, by Zimmerman et al of Carnegie Mellon (2011), uses passenger-generated GPS traces and reports to make real-time bus arrival predictions. It also solicits intelligence on bus crowding level from passengers waiting at stops or already travelling onboard. Carrel et al (2015) devise an Android app (San Francisco Travel Quality Study) to examine the relationship between objectively measured service quality (e.g. travel and wait times) and riders' satisfaction, emotions, and modal choice. During the month-long study, participants are asked to take public transport on at least five days and fill out the corresponding daily in-app surveys (for which they received a reminder every day). The system generates valuable multi-day data for understanding customers' ride experiences, even though the resolution of feedback is at the daily level rather than being trip-specific. For users who take more than one public transport trip on a given

day, they are allowed to submit only one set of ratings, even though their experience could vary widely from trip to trip. In contrast, Dunlop et al (2015) use a BlackBerry OS-based survey app (TOES) to measures riders' emotional state before, during, and after each bus trip. The trip stage is determined by users' specification, manually inputted into the app. Dunlap et al's study finds that passengers' anxiety and discomfort perception are highest while on a crowded bus, and new bus users are more sensitive to crowding noises, smells, route/schedule uncertainty, and self-consciousness than experienced public transport riders who are used to the environment. Both Carrel et al's and Dunlop et al's studies demonstrate the feasibility of leveraging emerging technologies to yield more granular insights on people's public transport experiences.

For smartphones to become a more practical and common tool for customer surveys, however, there remain several major challenges to overcome. There exists a flipside for almost every advantage of smartphones mentioned earlier. While the smartphone can collect rich info on one's travels through its sensors, it can cause substantial battery drainage (Love, 2013). Incentivizing and sustaining participation is also a significant challenge. Though people tend to carry their phones with them and can generate intelligence ubiquitously, they do not necessarily want to keep additional apps running on their phone or respond to survey solicitations in the middle of a game or even Facebook browsing. Carrel et al (2015) found it hard to sustain participation after the initial weeks, even with an enticing incentive upfront (a free monthly pass). Another challenge highlighted by both Carrel's and Dunlop's studies, is the trade-off between high information resolution and user burden. If one wants to sense service quality metrics specific to a trip, Tiramisu and TOES would require users to manually signal to the app when they are taking public transport. Lastly, while a smartphone-based system is poised to lessen the staffing required to conduct surveys, its penetration rate among the general population is still too low to make it the sole channel. As of mid-2015, 28% of Americans did not own a

smartphone (Poushter, 2016); nearly half of the smartphone-dependent population[3] (19% of Americans) have limited data plans – having had to cancel or suspend their cellphone service due to the financial constraints, or occasionally maxed out their phone plan's data limit (Pew Research Center 2015). Hence, in order to collect more representative samples, agencies that use smartphone tools will need to, at least for now, supplement it with more conventional survey mechanisms.

But these current shortcomings should not deter researchers and agencies from improving and leveraging smartphones tools. Smartphones components are getting gradually more energy efficient (although the gain is partially offset by bigger screens and the propensity to leave more apps running) (Newman, 2013), so can apps and algorithms be designed to be less draining for the batteries. One can lessen user burden by enhancing the platform's automation and inference capabilities (an area where FMS-TQ innovates in); participation is likely to be sustained if we integrate the survey mechanism into a comprehensive app that people are constantly using to check real-time information and plan and pay for their trips. The issue of unrepresentativeness will be mitigated by the rapid increase in smartphone adoption -- projected to grow by another 60 million by 2019 (eMarketer 2015). In fact, smartphones might soon edge over web-based methods in terms of accessibility. Recall that nearly one in five Americans relies on their smartphones for staying connected to the Internet world, and ownership of mobile devices is still rapidly growing but has plateaued for desktop computers .

---

[3] Smartphone owners who have no or limited alternatives to access the Internet.
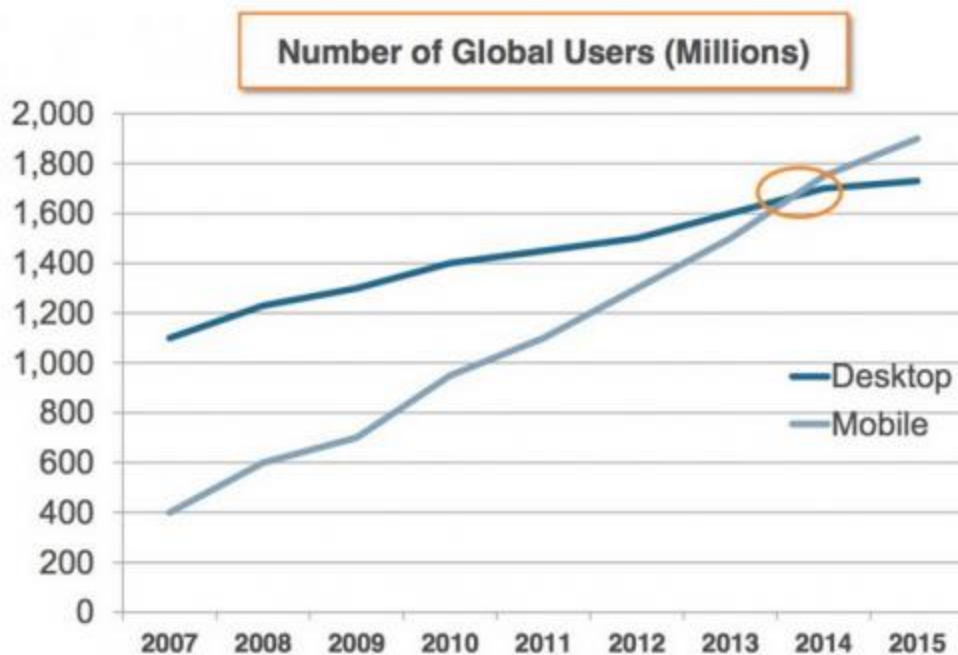
Figure 2.3 - Ownership of desktop vs. mobile computers

Source: comScore (2014)

The comparative advantages of smartphone systems are poised to improve precisely the weaknesses of existing survey methods. Automatic data solicitation can increase survey frequency from annually or quarterly to multiple times a day, and the real-time transmission would enables passengers to provide feedback on individual trips. This increases both information resolution and accuracy – surveyor may ask respondents about specific attributes on a specific public transport experience soon after, or even during, the trip. This capability is reminiscent of the iconic rating mechanism that is part of the app-based mobility services, which has opened up new frontiers in soliciting individualized feedback for quality control and cultivating relationships with riders. Companies such as Uber and Lyft have built meticulous quality control on the mandatory mutual ratings between drivers and passengers at the end of every trip (Cook, 2015). As a result, not only do they have detailed service quality data for analysis and monitoring, these new mobility service providers have also become much more self-aware and self-regulating of their quality standards, as I

will elaborate in the next chapter. Although mass transit may not require feedback on every passenger trip, the ability to examine performance by driver, route, and/or time of day could be very valuable for quality monitoring and identifying areas for improvement.

With high-resolution sensing comes great opportunities to collect more actionable information. One of the biggest inadequacies of existing pubic transport survey methods, as earlier discussed, is difficulty in interpreting aggregate scores in terms of further action. A satisfaction score of, say, 6.5 out of 10 for bus stop conditions does not reveal which of the thousands of stops in the city needs improvement. With trip-specific surveys, agencies can now ask respondents to report the condition of their boarding and alighting stops, which can be clearly identified by either GPS information or self-report.

Though most public transport agencies recognizably face many pressing short-term funding and operations constraints, it is my hope that these issues will not prevent them from investigating into the next generation of data gathering platforms. Strategic leveraging of new technology brings not only new data but also – and more importantly – new ways to engage customers and govern service. The following chapter will discuss this new approach to monitoring and managing public services. In these tumultuous, changing times, technology should not be an end in itself, but instead a means to enable new organizational mindsets and capabilities.

Chapter 3

Customer Feedback: Data Collection, Institutional Learning, and Relationship-Building

"We must use all available technologies and methods to open up the federal government, creating a new level of transparency to change the way business is conducted in Washington and giving Americans the chance to participate in government deliberations and decision-making in ways that were not possible only a few years ago."

--- Obama-Biden campaign, May 2008

In 2008, when smartphones were only in the technological cradle and Big Data was a few years away from entering the public lexicon, President Obama exhorted the nation to rise to a challenge: using new technology to make governance and public services more open, participatory, and collaborative. In the years since, leaps in digital connectivity and data management capacity have drastically improved information flow between governments and the people. Social media channels such as Facebook and Twitter now enable agencies to efficiently disseminate information and respond to constituency feedback; the Open Data movement has greatly enriched the public's knowledge of our world by releasing much once-inaccessible data.

This new connectivity brings deep implications to the relationship between governments and constituents in public service provision. In the traditional view, people are recipients of services provided by the government that they have elected. Given the new communication channels and ever-increasing amount of information generated by our digital lives, the public can become more involved in monitoring and providing feedback on service performance. This prospect has given rise to the concept of "co-monitoring" – agencies using public feedback to supplement the official monitoring and regulation (Kaufman, 2014). The promise is that service providers can gather details from the users'

perspective in a more dynamic, low-cost, and real-time manner; providers' responsiveness and working partnerships with their customers could also be enhanced. The customers would consequentially benefit from the improved quality of service and strengthened sense of citizenship and empowerment.

This transformative process highlights three inter-related and somewhat sequential aims of engaging customer feedback – data collection, institutional learning, and relationship-building between consumers and providers. The transit sector can much benefit from each and all of these objectives, and our FMS-TQ project seeks to strengthen agencies' capacities on all three fronts. This chapter details each of the three dimensions of leveraging customer feedback, drawing examples from public and private service sectors. It also illustrates particular opportunities, challenges and strategies for transit agencies to innovate in this sphere.

3.1   Feedback as data collection

At the most basic level, soliciting customer feedback is about intelligence-gathering: obtaining customer opinion about a business, product, or service. These opinions have been widely-accepted in modern management theory as important factors to business success. Feedback sheds light on customer's perception of the quality of offerings, and this perception influences customer satisfaction, which in turn exerts powerful impact on a company's financial and reputation outcomes (Walter, Steyrer, & Wiesel, 2010). Common channels for data collection include focus groups, individual interviews, follow-up surveys, social media, and hotlines. According to a 2014 benchmarking effort involving 218 large organizations, 100% of them have some form of a "Voice of the Customer" program (Temkin, 2014). The public sector is also increasingly striving to identify customer needs and monitor customer perceptions of the services provided. Customer satisfaction surveys, as detailed in Chapter 2, as well as hotlines for customer suggestions and complaints, are quite common among government agencies.

Given the widely differing nature among public services, conventional channels of feedback solicitation work to various degrees of effectiveness. For services that one would

only seek occasionally and where the service provider already has customers' phone numbers or email addresses, the agency can actively solicit feedback after the transaction via text or email. Some examples include the Department of Motor Vehicles asking permit applicants' about their service experience (California DMV, 2012), or the "I Paid a Bribe" initiative that spans across 29 countries, asking people by text to report any bribes that they have paid for their governmental service (I Paid A Bribe, n.d.). This active solicitation very much mirrors private sector practice, such as hotels contacting guests after their stay to seek their opinion on service quality.

On the other hand, every-day services, such as water and public transport, tend to invite customer feedback by publicizing hotline numbers or contact emails. Inevitably, feedback that comes in through such reactive channels are generally more negative in nature (Gigler & Bailur, 2014). Due to the effort involved in making a call or writing an email, people tend to not give feedback unless they have had an unpleasant experience; and even then, not all of those who are dissatisfied may overcome the burden to call or write to authorities. Despite this bias, the complaint-based system is actually an effective data collection tool for sectors that normally maintain good quality but occasionally deviate from standards. Many water utilities rely on customer complaints on water quality and pressure to detect changes in source water quality, measure the effectiveness of hydrant maintenance, and determine where pipe breaks have occurred (Whelton, Dietrich, Gallagher, & Roberson, 2007). Many disease outbreaks have been preceded by customer complaints about aesthetic water quality problems (Whelton et al., 2007). But this orientation towards negative feedback would not be nearly as effective for public transit agencies. Transit service performance can vary considerably by time, route, and operator; passengers also have heterogeneous judgement standards. This gives rise to the issue mentioned in this thesis' opening – performance shortfalls are disproportionally represented in the projected voice of the customer. For this reason, large-scale surveys remain the cornerstone of customer feedback solicitation in the transit sector, which, as discussed in Chapter 2, have the benefit of sample representativeness but disadvantages of low-resolution and high costs.

In recent years, the proliferation of mobile Internet connectivity and social media has opened up a new realm of customer feedback in both private and public sectors. Mobile

connectivity gives people the ability to send information on the go, potentially providing more accurate pictures of on-the-ground realities (Gigler & Bailur, 2014). Websites that aggregate service providers onto common platforms further reduce the burden of initiating feedback, and incentivize customers through virtual or in-kind rewards. The website Yelp tellingly illustrates this unprecedented amount of crowdsourced data. Whereas one might previously not bother to call or write to individual businesses to give feedback, she can now easily give her review on the establishment's page on Yelp. In fact, it's so easy to review anything on Yelp that one can find reviews for some unusual listings on the platform – the City of Boston (yes, the entire city)[4], local prisons ("Only 3 basic cable channels and no nets for the basketball hoops are the least of the problems here. But the staff is decent...")[5], or the C Train of the New York City subway (2 out of 5 stars)[6]. While the ease of giving feedback does not entirely eliminate the negativity bias inherent in unsolicited reviews systems, Yelp's self-posit as a transparent review platform (as opposed to a complaint hotline) has brought in quite a lot of diversity in the spectrum of reviews. Out of a Yelp-released dataset of 42,153 businesses, the reviews for many of the top 30 categories of businesses actually assume a bell-curved shape (minimaxir, 2014), meaning that most of the ratings hover around 3 stars, and the extremely positive and negative reviews are somewhat balanced.

Twitter and Facebook also join the ranks for ground-up data gathering channels. According to an analysis of four European countries[7] by IBM's Social Sentiment Index, 67% of commuters talk about their commute on social media (IBM, n.d.). Unlike Yelp's value proposition as a neutral archive of reviews, Twitter and Facebook accounts are managed by the organizations themselves and serve as a dynamic communication channel between the account owners and their followers. As a result, while they are effective means for public agencies to quickly broadcast announcements to their constituents (Bregman, 2012), communications originated from customers tend to concern service shortfalls and anger sentiments. When researchers at Purdue University analyze a sample of Twitter posts

---

[4] https://www.yelp.com/biz/city-of-boston-boston-3
[5] https://www.yelp.com/biz/hagerstown-prison-hagerstown
[6] https://www.yelp.com/biz/the-c-train-new-york
[7] Germany, Spain, France, and the Netherlands

regarding the Chicago Transit Authority, they find that "transit riders are more inclined to assert negative sentiments to a situation than a positive sentiment" (Collins, Hasan, & Ukkusuri, 2013). Figure 3.1 and Figure 3.2 show examples of such messages that customers have directed at transit agencies through Twitter and other social media. It is thus not surprising that fear of online criticism is one of the major barriers for transit agencies to using social media. According to a survey of U.S. and Canadian transit operators in 2012, 60% of responding agencies considered the issue of public criticism "important" or "very important" (Bregman, 2012).

Survey Subject	Agencies Measuring	Tweet
Overall Satisfaction		"The @cta has to be the most hated public transit system in the country. Never on time, dirty, outdated, slow, dangerous, etc. #ctafail" - @KyleDeGiulio
Speed/On-Time Performance		"the m train super slow, everyone told to get off at myrtle-bway, http://mta.info has no updates to explain this @MTA" - @MJalonschi
Reliability		"Hey @MTA! My friend @Kirznyc waited 3 hrs & 3 buses b/c lifts not working & drivers not trained pic.twitter.com/cjvkGgMjSX" - @gemaree
Safety from Accidents		"Either there's something seriously wrong with this bus, or the driver doesn't understand how an accelerator works. #sfmuni" - @nonsoccermom
Personal Security		"filing complaint on @wmata website to report attempted assault by bus driver—which drop-down menu choice applies? pic.twitter.com/HCYdBllFAy" - @IMGoph
Value for Money		"Did @wmata really put out an ad bragging about how much of our money they're spending? pic.twitter.com/HGqd5BOxvh" - @kctimpf
Cleanliness		"getting complaints about garbage @cta red line. personally, i just saw actual human waste on that train. u? (cc: @LittleBirdBill) #DreadLine" - @tracyswartz
Temperature		"Hey #BART it's 70 degrees and your train is packed. Heat is really really unnecessary" - @KCalder3
Communications		"Started tracking the 9 minute bus when it was 18 minutes away ... over 40 minutes ago. @CTA #FAIL pic.twitter.com/2o0FmnREcj" - @tankboy

Figure 3.1 – Sample Twitter posts concerning public transport experiences

Source: Kaufman, 2014

48

Figure 3.2 – Examples of public transport-related content from various social media

Source: Kaufman, 2014

Despite the prospect of receiving negative public comments, public agencies who view open review platforms as opportunities to learn from their customers and to correct misinformation, can benefit much from these customer voices. Whereas a city might not have the human resources to send employees to constantly inspect its services, citizens can serve as ubiquitous, free sensors on the street. As Clark and Rokakis discuss, this network of individuals is "essentially employed to bridge organizational information gaps and asymmetries" (Clark & Rokakis, 2014). For example, during adverse weather events that affect travel throughout a region, messages from customers have helped the local transit agencies identify trouble spots (Bregman, 2012). These customer-generated data all give rich intelligence on realities on the ground, and can potentially fill monitoring gaps for geographies or time periods that are previously under-monitored by the agency.

3.2   Feedback as institutional learning

It is one thing to collect better data, but another to put them to good use. After organizations set up channels to listen to customers' voices, they need to embed the

information into their institutional learning process. This concept is advocated by numerous experts in business management (Berry & Parasuraman, 1997; Wirtz & Tomlin, 2000) and international development realms (Andrews, Pritchett, & Woolcock, 2012). In order to ensure that customer feedback translates into better services, the data need to be systematically-collected and built into the organization's performance metrics and decision-making processes.

One of the most telling instances of feedback institutionalization comes from the transportation sector itself: the transportation networking companies (TNCs), as first discussed in Chapter 2. With such large, decentralized networks of service providers – over one million drivers for Uber alone (Carson, 2015) – and no uniform vetting or training, the TNCs cannot ensure or monitor the quality of every individual driver through aggregate surveys or secret shopper tests. So the TNCs, as earlier mentioned, solicit passengers' feedback through the app at the end of every ride. The trip-specific feedback, which consists of ratings out of five stars and/or comments, becomes an effective alternative. Having interned with Uber during one of my graduate school summers, I have seen first-handedly the power of such customer-generated intelligence as a service quality monitoring backbone. Every week, Uber and Lyft convey riders' comments anonymously to the corresponding drivers, encouraging them to keep up the good work and identifying areas of improvement, if any. In case a rider reports a serious issue, the incident is investigated. Though customers' rating and comments are far from objective – sometimes inflated or over-critical – they nonetheless provide a telling picture of each driver's performance when compared to the average performance of all drivers in the same city. If any individual driver's average rating falls below a certain threshold, the company offers him or her advice and recommends courses to improve the service. The driver's privilege to use the app may be suspended if his or her low rating persists (Cook, 2015). Customer feedback is supplemented with objective performance data, such the speed profile of a ride to detect speeding (McGoogan, 2016).

The public sector has also gained increasing capability to institutionalize citizen feedback. Since 1997, more than 200 cities around the United States have established traditional 3-1-1 services, allowing citizens to report non-emergency issues or inquire about particular government services (Goodyear, n.d.; Holzer, Schwester, McGuire, & Kloby,

2006). The most common types of calls involve reporting of debris on the road or broken street lights, as well as questions about citizen services such as trash pick-up (Schaeffer, n.d.). The 3-1-1 system aims to ease the longstanding difficulty in efficiently directing citizen questions and concerns to the right department. Municipal departments are often not well-integrated, and their differing procedures and policies often delay the processing of citizen requests. In contrast, when a citizen calls the 3-1-1 number, it goes to a single call center open 24/7. The center staff record the request and route it to the right department; the systems also include search and automation tools that help civil servants to respond to citizen inquiries expeditiously (Schaeffer, n.d.). Data collected via 3-1-1 may also be fed into customer relationship management software systems, hence allowing officials to assess service delivery performance and make informed decisions on how city resources could be better deployed and managed (Holzer et al., 2006).

The success stories of 3-1-1 systems are too numerous to list in full, so I will illustrate only a few cases here. The City of Minneapolis has leveraged the call system to fix its long-standing issue of graffiti removal delays. Prior to that, residents often complained that the city took too much time to clean up graffiti after it was reported to the police. The police had to investigate the report and take photos, before turning over the requests to the Public Works Department. This procedure was often held up as the police had more pressing criminal activities to handle. As a result, the average cleanup would take two to three weeks. The 3-1-1 system shortens the process by directing citizens' reports to the Public Works department immediately. A clean-up crew would photograph the graffiti and share pictures with their police counterparts (Barkin, 2009). This story demonstrates that when governments muster the political will for reform and leverage technology to further this purpose, they can remove bottlenecks in their institutional processes.

Institutionalizing feedback does not always mean reacting to reported concerns – it can also take on a preemptive flavor. During the Recession of 2008 and its aftermath, the City of Buffalo, N.Y. used 3-1-1 system data to identify neighborhoods with the worst decay. It culled data from complaints and requests for services, and conducted weekly sweeps through the hardest-hit areas in an attempt to staunch the deterioration – providing information to residents about employment and healthcare services, sealing vacant houses, mowing empty lots, and trimming overgrown trees. The targeted sweeps represented an

180-degree turn from the random efforts in the past to fix blighted neighborhoods, allowing governmental departments to target limited municipal resources on neighborhoods with the greatest need (Newcombe, 2014). Instead of simply responding to complaints, cities are proactively using constituent feedback to tackle problems before they get too big. The City of Chicago uses analytical software to sift through 3-1-1 calls to try spotting rat infestations before they become an actual problem (Newcombe, 2014). Well-managed citizen reports can provide much insight towards economic development, public health, and other urban policy decisions.

The platforms are continuously becoming more encompassing and convenient, too. For example, New York City now allows residents to submit requests to its 3-1-1 call centers through Skype and text messages (StateScoop, n.d.). Third-party platforms are also adding to the diversity of channels, with SeeClickFix (SCF) being the most prominent example. SCF is a web- and mobile-based platform that allows citizens to report non-emergency neighborhood issues. With smartphone-enabled GPS tracking capacity, SCF app users can file geo-tagged reports on the go. In addition to facilitating individuals and community groups to follow reports regarding a certain watch area or specific issue types, SCF has also integrated itself with the local government and official workflows in many municipalities. Elected representatives and civil servants assume the roles of local SeeClickFix Watchers (SeeClickFix, 2010), meaning that user-submitted reports are directed to their attention and they are accountable to resolving the tickets. Cities are also gradually exploring more systematic integration of crowdsourced reports with official work orders. For example, the New Haven Traffic and Parking Department has integrated SFC with their emergency work order management system, ensuring that SFC reports get managed as systematically as officially-prescribed orders (SeeClickFix, 2015). Such high degree of collaboration with decision-makers increases the speed and likelihood of issues being resolved, which boosts the public's propensity to contribute on the SCF platform. As of 2015, 220 municipalities are paying to integrate SeeClickFix, with many more relying on the free version of the tool (Goodyear, n.d.).

With great volume of feedback comes great responsibilities to process and respond to them. One pitfall of opening governments up for feedback is insufficient human resource to respond in time and lack of financial resources to fix a reported issue. The institutional

apprehension is reflected in a survey of 130 transit agencies across the U.S. in 2014. When asked about drawbacks to adopting web-based feedback tools, 64% of respondents selected "lack of staff to respond to comments in a timely manner" – more than any other reason (Sager, 2014). Given the resource-strapped nature of public agencies, especially transit systems, the concern about the institutional capability to integrate feedback is a significant one.

The resource constraints however should only be an additional motivation for efficient, innovative feedback systems and greater data transparency. The SeeClickFix team explains the rationale in one of its blog posts:

> "For the pothole complaints, the answer might be to fill the pothole with asphalt and fix the problem. But what if there were no resources? The answer might be "you can see how many potholes that need to be fixed as well as we can. Now you the citizen understand what we're up against."

> "What if the problem is graffiti, littering or a broken park bench and tax dollars have dried up? For that situation we provide a tool that not only allows distribution of communication to the traditional "fixing" channels, but to the rest of the community as a whole (anyone can create a watch area). When you open up the fixing channels to community groups, parks groups, private business and individual citizens, you distribute the responsibility as well as the communication." (SeeClickFix, 2010)

While transparency and distributed responsibilities are not the panacea for the growing pains of institutionalizing feedback, an explicit display of willingness to hear constituents' voices nonetheless builds bridges and fosters citizens' respect for their governments. This ties well to the discussion of the third objective of soliciting constituent feedback – to further the broader ideals of a responsive government and empowered society.

3.3   Feedback as relationship-building

For all levels of government, strengthening relationship with their constituents remains a high priority. Government transparency, accountability, and a politically engaged

constituency are highly-valued ideals of democratic societies. Kelling and Wilson, in their famous "broken window" study (1982), find that government's accessibility to its citizens improves people's satisfaction and perception with the government, even if the outcomes of public services remain unchanged. Conversely, satisfaction with government services declines when people feel hopeless to affect change in their communities. They observe that citizens stop calling the unresponsive police to report crimes after a while. Recent studies confirm the highly positive correlations between citizens' ability to interact with the government regarding questions or concerns and their satisfaction with public services (Dudley, Lin, Mancini, & Ng, 2015). As with the broken window that went unfixed (Wilson & Kelling, 1982), the potential for a pothole unfilled, graffiti not cleaned up, or dead animals left on the road are signs that a government is not responsive to community needs.

The benefits of responsive governance, combined with increasing mobile connectivity and municipal budget constraints, catalyze the thinking on "co-production" as a new model of service delivery (Clark, Brudney, and Jang 2013, 687). The concept of co-production, coined by Nobel Laureate Elinor Ostrom in the late 1970s in context of public services, is defined as "a process through which inputs from individuals who are not 'in' the same organization [being] transformed into goods and services" (Ostrom, 1996; Ostrom, Parks, Whitaker, & Percy, 1978a; Parks et al., 1981). As Bovaird and Loeffler (2012) synthesize, co-production of public services can take many forms, such as:

- Co-commissioning of services – involves constituents in thinking about what needs to be delivered, to whom, and to achieve what outcomes. It includes:
    - Co-planning of policy -- e.g. community planning workshops;
    - Co-prioritization of services – e.g. participatory budgeting; and
    - Co-financing of services – e.g. fundraising, agreement to tax increases;
- Co-design of services – brings in the experience of users and their communities to the design of public services. Examples include user consultation and customer journey mapping;
- Co-delivery of services – citizens and the public sector performing the services together, building on each other's assets and expertise. It includes:

- o Co-management of services – e.g. community management of public assets, school governors; and

- o Co-performing of services – e.g. peer support groups (such as expert patients), meals-on-wheels, and Neighborhood Watch; and

- Co-assessment of services -- citizens working alongside professional staff and managers to help organizations to better understand how they feel about the services. It includes:

  - o Co-monitoring – e.g. user feedback and ratings; and

  - o Co-evaluation – e.g. tenant inspectors, participatory village appraisals.

At the heart of public service co-production is the idea that government-citizen relationships change from paternalistic, provider-customer dynamics to more collaborative interactions (Levine & Fisher, 1984; Ostrom, 1996; Whitaker, 1980). For instance, in context of public safety – where Ostrom's work on co-production is founded – citizens are not simply "clients" of the police; they are active participants in the production as well as consumption of community security (Ostrom, Parks, Whitaker, & Percy, 1978b; Percy, 1978). In other words, co-production is to make citizens more connected to the government and engaged in their community affairs. One type of co-production, service co-monitoring -- soliciting for feedback -- has long been understood and practiced by the private sector counterparts. Responsiveness is one of the five main dimensions of perceived service quality in the hallmark SERVQUAL model (Parasuraman, Zeithaml, & Berry, 1988). Numerous researches have concluded that the effectiveness and responsiveness of a firm's communication with customers are the primary drivers of relationship strength (Moore & Moore, 2004; Sharma & Patterson, 1999; Strauss & Hill, 2001). Though I have not found official statistics on the prevalence of customer follow-ups in the private sector, I trust that I am not the only person who constantly receives emails, texts, or phone calls from businesses, thanking me for my visit and asking for feedback.

Initiatives seeking constituent feedback would positively improve government-citizen relationships, especially when the government effectively communicates its good will to the public and provides tools that make feedback-giving easier. Let us re-examine the 3-1-1 system for its relationship-building benefits. Prior to the introduction of 3-1-1,

making an inquiry with the local government was a frustrating experience for most people. Every agency had its own contact line, many of which were open only during business hours; people's calls got transferred from one department to the next. In an oft-referenced anecdote, Michael Bloomberg, while campaigning to become mayor of New York City in 2001, noticed a leaking fire hydrant. After discovering, in astonishment, that the issue actually fell under the jurisdiction of the Department of Environment Protection (DEP), he waded through 14 pages of city listings in the NYC phone book to find the DEP's phone number (Oracle, 2006). After Mayor Bloomberg brought the 3-1-1 system to New York City, citizens would only need to remember one number: 3-1-1. The call center centralizes the point of contact between government and citizens -- a "governmental Wal-Mart" that provides citizens everything they need and is open 24-7 (Martin, 2014). 3-1-1's integration with social media channels, along with web-based platforms such as SeeClickFix, further eases the communications procedures for citizens. Aside from the earlier-discussed benefits of efficiency and effectiveness for service monitoring, these initiatives also enhance the government's pubic image. For instance, when Boston developed Citizens Connect, the city's mobile platform for citizen services, officials framed it as part of an effort "not only to provide more transparency around the City's performance but also to further establish Boston's commitment to providing the best possible City services to its residents, businesses and visitors" (City of Boston, n.d.).

In the urban mobility sector, companies such as Uber and Lyft again offer good examples of relationship-building with customers through soliciting and managing feedback. Placing strong emphasis on user experience at the center of their offerings, they are quite responsive to customers' concerns or issues, no matter the severity. From my personal experience as a passenger, when one complains about unsatisfying service – dirty vehicle, delayed arrival, etc. – he or she would receive a reply from the dedicated passenger support staff within several hours. The staff would apologize on behalf of the company for the service shortfall, convey that they are following up with the driver to ensure improvement, and lastly thank the customer for bringing the issue to their attention. If the complaint concerns something serious, such as safety or professionalism of the service, customer support managers would get involved to reach out to affected parties, seek reconciliation, and possibly offer compensation as a gesture of good will.

The TNCs' effort in relationship-building, along with their increasing share in the urban transportation market, accentuates the importance of relationship management for other transportation services. For public transport especially, the service anonymity and its orientation towards serving the masses give it scarce mechanisms to keep track of interactions between a given customer and service provider. Of course, this does not imply that customers expect their transit operators to interact with every passenger as closely as a personal taxi ride. Initiatives to better understand riders' actual trip experience would alone be a significant positive step, so would measures that enable more direct communication channels between riders and operators. The advance in technology is opening up an unprecedented window of opportunity. If the transit sector can muster the institutional impetus and start shifting its mentality, it has the potential to innovate and enhance its relationship-building with the customers.

3.4   Challenges and Ways Forward

Aspiring to enhance these benefits of customer's voices for the public transport sector, we develop the FMS-TQ project considering all three objectives of leveraging feedback. First and foremost, the platform should bring in new, meaningful data for the partner public transport agencies. This takes the form of the trip-specific, objective and subjective information on service experience, which we develop the exact parameters with our agencies partners (described in Chapter 4). As for the benefit of institutional learning, we ensure that the data collected can be matched to specific records in MBTA's official database, so they may incorporate the information into their operations decision-making, if the pilot proves the technology viable at scale. Lastly, we hope that the smartphone app will make it much easier for people to give feedback to the agency – the app would automatically detect bus trips and prompt customers with a quick survey. Riders would not need to look up the best number to reach the agency, or to remember all the details of the specific trip to provider to the hotline service representative.

We are also aware of the challenges associated with soliciting feedback. From the data collection perspective, the key challenge for service providers is obtaining a sufficient amount of constituent feedback to be representative of the reality on the ground. While

people might intuitively be eager to report service issues that inconvenience or frustrate them, would they bother to respond when things are going well? To understand people's motivations for giving feedback, I looked into the literature on online review platforms. According to existing studies, intrinsic motivations actually dominate among reasons to contribute feedback. Hennig-Thurau et al (2004) are the first scholars to propose a comprehensive, succinct list of potential motivations for engaging with opinion platforms: platform assistance, venting negative feelings, concern for other consumers, extraversion/positive self-enhancement, social benefits, economic incentives, helping the company, and advice seeking. Of the online reviewers surveyed, 65% report being primarily driven by non-economic motivations, especially concerns for other consumers and helping the company. Similarly, Yoo and Gretzel (2008)'s survey of a consumer panel on travel review site Tripadvisor finds that reviewers are mostly motivated by helping a travel service provider, helping other travelers make better decisions, and deriving personal enjoyment from making the contribution. Venting negative feelings through postings is not seen as an important motive. The conclusion is echoed in Parikh et al (2014)'s study of Yelp reviewers, as well as in Munar and Jacobsen (2014)'s survey of Norwegian vacationers. All these findings imply that appealing to people's sense of pride from helping others may be an effective strategy in encouraging feedback.

Some review and reporting platforms have leveraged this observation by bestowing intrinsic rewards to users to encourage participation. For example, SeeClickFix rewards its users with "Civic Points" that are commensurate with the importance of their actions – logging in, reporting an issue, generating discussion among users, etc. (Figure 3.3). The accumulated points qualify users for various feel-good titles within this virtual community. Yelp allows reviewers to develop their reputation by linking each review to a profile page, which contains summary data on review production; it also recognizes the most active users as "Yelp Elites", who receive a special badge next to their names on the website and invitations to local events. These points and titles make users' participation very salient – both to themselves and the rest of the community – and reinforce the good feelings from knowing that they have contributed.

	Civic Points quantity	Titles users earn based on Civic Points
• Signing up with SeeClickFix: 50 points	<100	"Street Smart"
• Logging in: 5 points per day	100-250	"Civic Pride"
• Commenting on an issue: 5 points	250-500	"Civic Crusader"
• Reporting an issue: 10 points	500-1000	"Municipal Avenger"
• Uploading an image: 20 points	1000-2000	"Digital Superhero"
• Uploading a Youtube video: 30 points	2000-5000	"City Fixer"
• Getting an issue you reported closed and archived: 30 points	5000-10000	"Heman"
• Getting at least one user to comment, vote, or follow your issue: 50 points	10000+	"Jane Jacobs"

Figure 3.3 - SeeClickFix's Civic Points reward scheme

Source: SeeClickFix

For the public transport sector, though, collecting valuable feedback might need more than just appealing to riders' intrinsic motivations. In an anecdote titled "Nobody Cheers About a Cell Phone Charger," Couzin and Grappone (2014) talk about the phenomenon that products and services of utilitarian nature tend to generate negative reviews. A reputation study that analyzed 10,000 e-commerce reviews of cell phone chargers scarcely found positive comments (Couzin & Grappone, 2014, p. 67). The reason is simple – "a phone charger either does its job or it lets you down. You're never going to be ecstatic about your phone charger, but you might be disappointed" (Couzin & Grappone, 2014, p. 67). This issue is very much relevant to public transport services, which tend to lack the emotional connection of a vacation or have little opportunity to actually exceed customers' expectations. Extra efforts are thus needed to proactively encourage reviews from a more representative cross-section of the customer base. Consumer psychologists recommend making the review process so easy that people will do it regardless of the mundaneness of their experiences (Couzin & Grappone, 2014). Similar to Amazon customers receiving an email solicitation for rating their recent purchase, public transport agencies can proactively ask riders to give quick feedback on their most recent ride.

Extrinsic incentives – money or prizes, etc. – may also encourage and help sustain participation.

Research on the incentive design points to the advantages of pre-paid rewards for survey. Studies have shown that fixed, prepaid schemes are significantly more effective at eliciting responses from respondents than actuarially-equivalent lotteries or donation to charities on participants' behalf (Halpern et al., 2011; G. M. Leung et al., 2004; Warriner, Goyder, Gjertsen, Hohner, & McSpurren, 1996). Furthermore, the amount of money offered in lotteries and the probability of winning have minute effects on enticing participation (Halpern et al., 2011). As Fehr and Falk (2002) point out, money in hand evokes duties of reciprocity, which by itself constitutes a powerful incentive. In fact, reciprocity is one of the key pillars of persuasion, as discussed in Robert Cialdini's book "Influence" (2006) – a hallmark work in marketing. Public transport agencies may thus consider offering free rides or passes to attract participants. However, for panel surveys that seek sustained participation over a longer time period, a combination of pre- and post-paid incentives (based on participation) may be even more effective (K. W. Axhausen, Löchl, Schlich, Buhl, & Widmer, 2007; K. Axhausen, Zimmermann, Schonfelder, Rindsfuser, & Haupt, 2002; Carrel, Sengupta, & Walker, 2015).

As for institutionalizing feedback and building relationships with constituents, the biggest potential challenge is mustering the political will to streamline organizational processes. Like any other organizational change initiatives, process and mentality changes often mean more work and unfamiliar terrain for employees used to the present routine. According to a 2012 survey, employees in the public transport sector are the sixth oldest among all industries in the U.S., with 66% of workers over the age of 65 (ICMA, 2013). The senior workforce would likely face a steeper learning curve towards new technologies, and this difficulty is only exacerbated by the risk of inadequate training and support resources, in face of tight budgets or budget shortfalls across U.S. public transit agencies.

Despite the challenges, leveraging constituent feedback to better monitor and manage public services represents a significant step towards better service delivery and a smarter city. However, unlike the IBM-esque Smart Cities – which is largely about sensors and Big Data analytics – this smarter city is about integrating human wisdom into urban governance. In an environment where constituent feedback is well solicited, managed, and

communicated back to the citizens, this wisdom of the crowd not only enriches the diversity and quality of data for the betterment of public service delivery, it also contributes to the making of a more human-centric and empowering city as a whole.

Chapter 4

Future Mobility Sensing – Transit Quality (FMS-TQ) Platform

The proliferation of smartphones has opened up a world of opportunities for transportation data gathering. Smartphones can constantly collect user location and movement information, as well as interact with users in displaying, soliciting, and transmitting information back to the server. This chapter details the development and features of the Future Mobility Sensing –Transit Quality (FMS-TQ) platform as one leading endeavor in leveraging smartphones to study urban mobility. It first describes the foundational platform Future Mobility Sensing, which has been in continuous development over the past five years, then the hardware and software adaptations involved to enable real-time transit service quality and customer satisfaction assessments.

4.1   Future Mobility Sensing (FMS)

The Future Mobility Sensing (FMS) platform is a smartphone- and web-based prompted recall survey system, originally developed to automatically infer users' daily travel behavior (trips, stops, modes, etc.). Developed at the Future Urban Mobility (FM) research group under the Singapore-MIT Alliance for Research and Technology (SMART), FMS consists of three core components as depicted in Figure 4.1: a smartphone app, a server, and a web interface for users. The app collects phone sensor data (Wi-Fi, GSM, GPS, and accelerometer) to capture the device's movement and location. The backend server receives the raw data, stores them in the database, and processes them using machine learning algorithms to infer the user's stops, travel modes, and non-travel activities. This synthesized information is then accessible to users either via the web interface or within the mobile app itself. Users can see their daily travel patterns traced over a map, and may be asked to validate the transport mode (e.g. car, bus, walking) and purpose (e.g. work, recreation) of each trip that FMS has detected and inferred (Figure 4.2). As the system's algorithms learn from the user's travel behavior, they become increasingly good at inferring trip mode and purpose over time, requiring less user validation for repeated trips.

Users may examine or validate this historical information through the web or app portals at any time (Zhao et al., 2015).
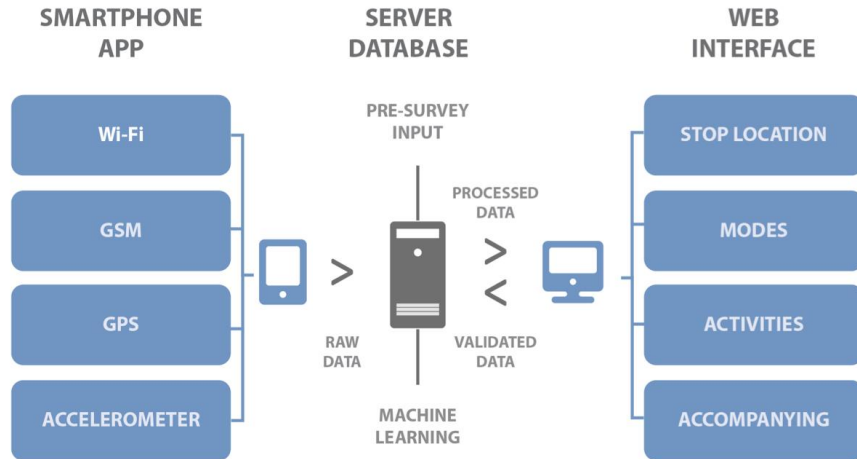
Figure 4.1 - FMS platform architecture

Source: Zhao et al., 2015
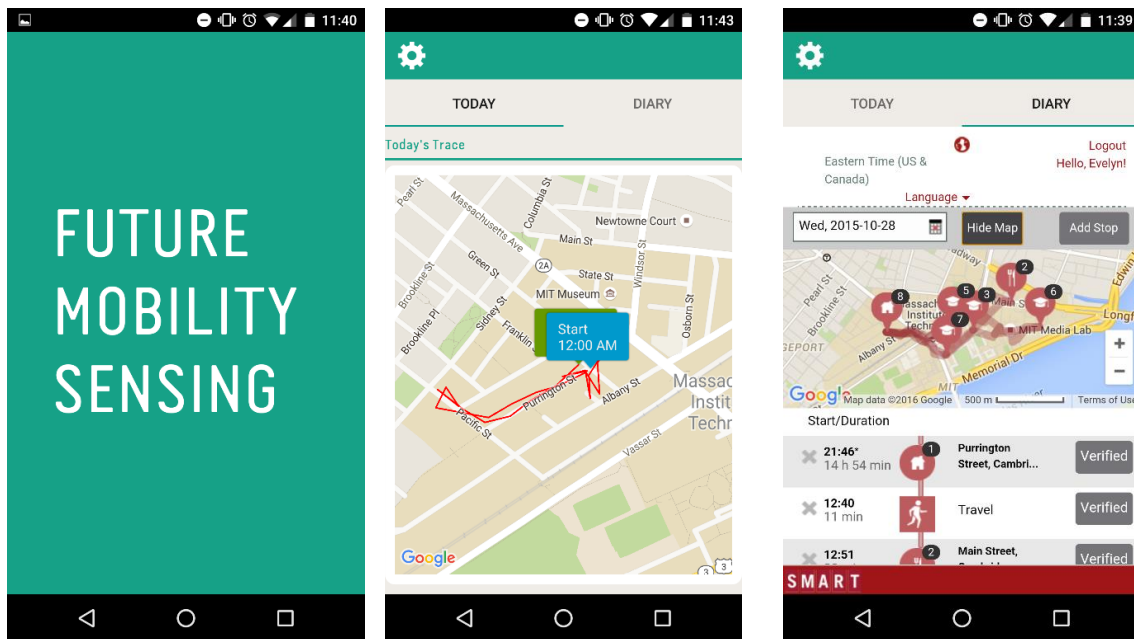
Figure 4.2 - FMS interface. Left: App launch screen; Middle: trace of user's travels of the day; Right: activity diary showing travel mode and stop inferred from machine learning algorithms

The smartphone app, available for Android and iOS phones, is designed to be non-intrusive. It runs in the background of the phone and collects data from sensors without user intervention. Only the recent extensions and experimentations of the FMS system have introduced some occasional exceptions. For example, the Happiness Survey project samples people's moods by pushing a short pop-up survey once a day in the app (Raveau et al., 2016). The TQ extension is the second project involving user's manual input. A major constraint to the FMS is battery consumption, which tends to be a major concern for location-based applications that traditionally use GPS for positioning. The FMS development team has applied numerous sampling methods to optimizing battery drainage while attempting to maintain location detection accuracy (Rudi Ball et al., 2014).

The backend algorithms translate raw data into trips and activities. First, the algorithms detect stops on a user's journey using location and point-of-interest (POI) data. Then, they use the collected GSM, Wi-Fi and accelerometer information to merge stops that would otherwise be interpreted as distinct stops. For example, if FMS detects multiple sequential stops within a mall and a walk-like speed profile, it would infer that the user is on an extended period of similar activities rather than making multiple short trips. Travel modes are detected based on GPS and accelerometer features, as well as public transport network information. Short duration stops that are insignificant from a data validation standpoint (such as stops in traffic or at bus stops or subway stations during their ride) are deleted for the purposes of presentation in the web interface. Non-travel activities (e.g. home, work, shopping, drop-off) are also detected based on previous validations by the user, POI data and other contextual information (Zhao et al., 2015).

The web interface provides a platform that enables users to review their processed data in the form of a daily activity diary and "validate" their data. It should be noted that though the web interface has traditionally been the sole portal for validation, FMS has recently enabled this functionality within the mobile app as well. Validation involves filling in missing information and updated incorrect inference of travel modes for particular trips or specific activities at the destination (Figure 4.3Figure 4.3). The validated data are uploaded and the algorithms learn to make better inferences as the user interacts with the interface. Supplemental data pertaining to a specific trip (e.g. whether the user traveled alone or with accompaniment) are also collected within the activity diary validation stage.
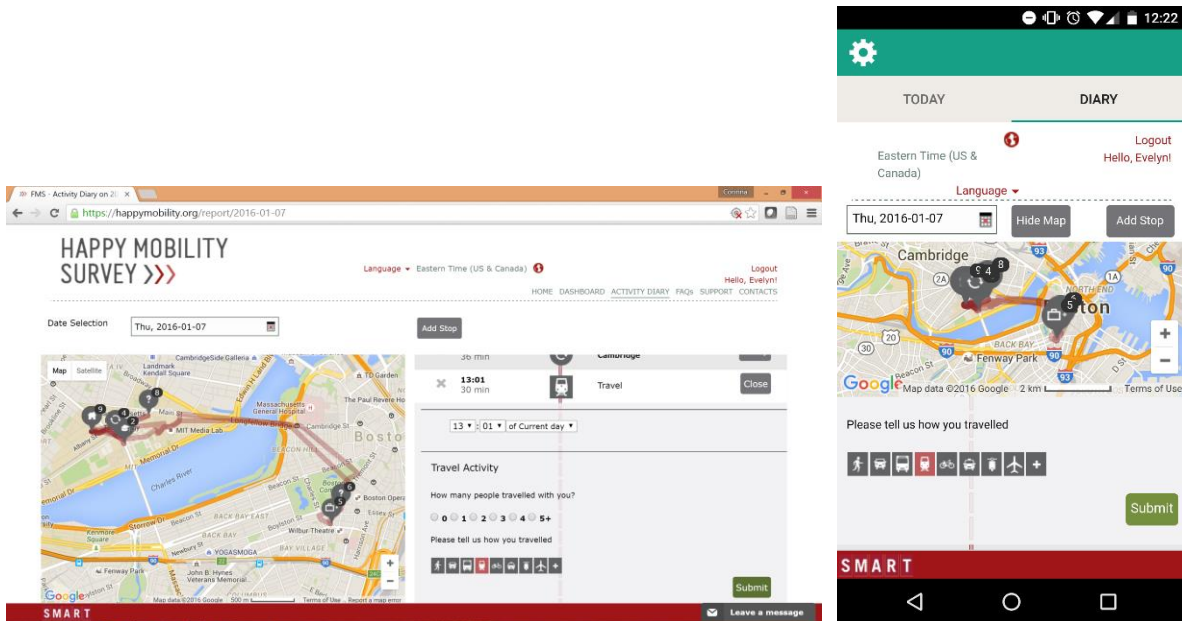
Figure 4.3 - User validation mechanism. Left: web interface; Right: mobile app interface

In empirical studies, the FMS system has demonstrated advantages over traditional travel survey methods (Zhao et al., 2015). FMS is capable of producing accurate, detailed, and rich data for travel surveys. By sensing how people travel rather than asking them to report their travels, it eliminates many problems that traditional self-reported surveys face: under-reporting of short trips, inaccuracy in location and times, and reporting of a typical day rather than the actual day. A further advantage of FMS is the small marginal cost of collecting additional days of data. While the participant may need to devote some effort in the beginning to familiarize himself with the app and provide validation, participation burden reduces significantly over time. The relative ease of longitudinal tracing over a period of time via FMS reveals large intra-user variations in the travel and activity patterns, and researchers have been able to typify these patterns to segment different kinds of users.

FMS forms the foundation of the FMS-TQ system. The rest of this chapter will describe the adaptations undertaken on the FMS platform, including the addition of additional external sensing capabilities (to detect external Bluetooth Low Energy transmitting devices, known as beacons), to enable event-driven real-time transit customer surveying and service quality monitoring.

4.2 Future Mobility Sensing – Travel Quality (TQ)

The TQ extension builds on the core FMS app and database – including all its travel sensing and learning algorithms – but involves two additional components (Figure 4.4). First, it adds the capability to detect signals transmitted by external devices (Bluetooth beacons); this allows the system to know with high precision when a user has come into an area of interest (in our case, arriving at a bus stop or boarding a bus), which enhances real-time transit trip detection without compromising phone battery performance. Second, whereas FMS only passively collects travel data through one's phone sensors, FMS-TQ also actively solicits interaction with users based on specific travel patterns detected.
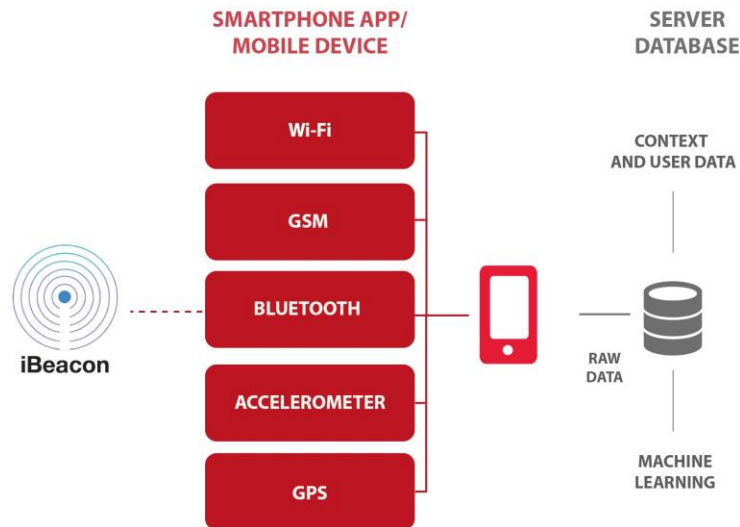
Figure 4.4 - FMS-TQ system architecture.

Note the addition of a beacon component unique to the TQS extension

Incorporating Bluetooth Low Energy (BLE) beacons into the platform is one of the most technologically interesting, challenging, and innovative aspects of this project. The beacons advertise (i.e. broadcast) small packets of data at a regular interval via radio waves. This allows mobile apps running on Bluetooth-enabled phones (both iOS and Android) to listen and react according to the signals. Given each beacon's unique identifier information, this technology essentially allows apps to understand the phone's physical position on a micro-local scale (iBeaconinsider, 2014). For this reason, BLE beacons are

increasingly used by retailers to push catered ads or offers to shoppers, according to shoppers' location within the shop or mall. Some large public facilities, such as museums and airports, have also begun using beacons to enhance wayfinding for customers and analyze people's movements within the venue. To the best of my knowledge, FMS-TQ is the first project that brings beacons onto a fleet of moving buses and outdoor bus stops. The BLE technology, as its name implies, requires drastically less battery consumption comparing to classic Bluetooth devices. Apple and Google have established two industry protocols – iBeacon and Eddystone, respectively – though both are compatible with Android and iOS devices. Beacons manufactured by different companies – of which there are many – vary in size, color, battery life, specifications and features (Figure 4.5).

Figure 4.5 - Select brands of BLE beacons.

Source: Trudel, 2014

We use this micro-local positioning function by installing beacons on buses and at bus stops. When a user of the FMS-TQ app enters or leaves a bus stop beacon's signal range, the app registers the arrival/departure times. Likewise, onboard beacons help the app determine when a user has boarded a bus, is travelling on a specific bus, or has alighted. This design works as beacons usually have a signal range of up to 70 meters (230 feet) (iBeaconinsider, 2014), sufficient to cover the length of a typical bus or a bus stop area. Given this ability to detect bus trips in real-time, FMS-TQ can be programmed to solicit users for feedback in a variety of formats at any point during or after the trip. For example, the beacon signal can trigger a survey asking a user about the bus stop condition while she

is waiting for the bus, about travel conditions while she is on board, or for a comprehensive evaluation of the trip experience after she alights. All this can be customized according to the research objective and scope of beacon installations. Each of the two FMS-TQ pilots conducted thus far – one in Singapore and another in Boston – has had customized algorithms, sampling rules, and survey questionnaires to cater to the local context. Table 4.1 exhibits the comparative design at-a-glance for the two pilots; the precise set-up of each pilot is described in Chapters 6 and 7.

Table 4.1 - Comparison of FMS-TQ setup between Singapore and Boston

	Singapore	Boston
Beacon installation	Beacon installed at 4 bus stops in downtown Singapore	Beacon installed on all Silver Line buses (approx. 45) and stops (50)

Available survey types	1. Entrance survey
	2. Trip-based surveys
	   A. At-stop survey
	   B. Onboard survey
	   C. After-alight survey
	3. End-of-day survey
		1. Entrance survey
		2. Trip-based surveys
		   A. Onboard survey
		   B. End-of-trip survey
		3. End-of-day survey

Sampling rule	• All users take the entrance survey • User receives maximum of one trip-based survey per day • User receives end-of-day survey at 8pm, if he has completed a trip-based survey, or has been inferred to have taken the bus, earlier that day	• All users take the entrance survey • User receives end-of-trip survey after every ride on the Silver Line • On select rides, user receives an onboard survey while she is still on the bus • Users who have been detected to have waited at a Silver Line stop during the day, but have not received a trip-based survey, receive an end-of-day survey at 8pm
Trip-based survey trigger	Phone enters and remains in beacon range for > 60 seconds	Phone enters and remains in bus beacon range for at least 60 or 90 seconds, depending on the survey

One might ask: "Shouldn't the FMS app itself be able to detect transit trips using a smartphone's sensors and the backend algorithms? Why do we need to supplement it with beacons?" The keywords are *in real-time*. Currently, FMS is capable of identifying some bus trips undertaken by the user when post-processing the travel data and acceleration profile

taking into account the local transit service information (GTFS[8]). But if we want the app to infer transit trips in real-time, it would need to frequently attempt to match the user's geolocation with the transit network on the server. This process is computationally- and battery-intensive, which would likely be counter-productive to sustaining participation among users who would quickly be discouraged by the effects on their phones' battery lives. In addition, phone-only detection infers speed of travel and looks for speed patterns resembling bus travel from changes in GPS locations over time. Any disruption in the strength or accuracy of GPS signals may thus compromise phone-only detection. Since beacons transmit close-range signals, they can help overcome FMS's two challenges associated with phone-only sensing of transit trips.

The biggest challenge in using beacons for trip detection is getting the signal strength just right. Just as GPS signal strength can be attenuated by physical objects like buildings, so can a beacon's signal strength (Figure 4.6). We ran a series of initial tests and found that the received signal strength (rssi) and proximity measurement accuracy were acutely affected by the number of people in the surrounding environment. Fortunately, attenuation can be mitigated by positioning the beacon to maximize direct line-of-sight with most users, and/or increasing the beacon's signal strength (an adjustable setting on most brands of beacons). But by setting the beacon signal too strong, the beacons would interfere with the app's detection accuracy if a phone is in proximity to many beacons – a bus-bunching scenario, for example. This would become a prominent issue in the Boston pilot due to the characteristics of the bus routes involved. I will later describe in further detail the particular designs and precautions taken in the Singapore and Boston pilots.
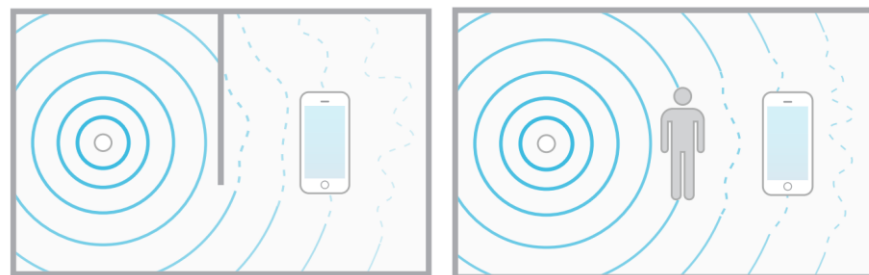
Figure 4.6 - Physical objects and human bodies can reduce BLE Beacon signal strength
Source: Apple, 2014

---

[8] The General Transit Feed Specification (GTFS) defines a common format for public transportation schedules and associated geographic information.

Upon evaluating various brands, the development team selected iBeacons manufactured by Estimote, a leading beacon manufacturer, for FMS-TQ. Each Estimote iBeacon contains a 32-bit ARM® Cortex M0 CPU, accelerometer and temperature sensors, a 2.4 GHz Bluetooth Low Energy radio, and a 1000 mAh CR2477 battery (Estimote, 2013) Key beacon settings – such as signal broadcasting strength, frequency, and identifier values – can be easily configured via the Estimote app & cloud. The manufacturer also offers a large software development kit (SDK) to help developers create their beacon-integrated apps. Physically, an Estimote beacon measures 54 mm x 36 mm x 17 mm and weighs only 28 grams, as heavy as five U.S. quarters. Battery life stands at about 21 months under Optimized iBeacon Settings[9], though its battery saving features can extend battery performance considerably. Its silicone casing renders the beacon water-and extreme temperature-proof – a major advantage for our design to install them outdoors at bus stops. The flat underside has a strong adhesive material for easy attachment to walls and other surfaces. Each beacon costs about $20.

Figure 4.7 - Estimote beacon size and construction

Source: Estimote, 2013; Harman, 2014

4.3　Design Framework and Principles

FMS-TQ is not merely a technology – it is a platform intended to cultivate a culture of rider engagement and empowerment in bettering transit services. This ambition places the principles of customer-centric design and meaningful data collection at the center of

---

[9] Used by Aislelab to test and compare battery performance across iBeacon brands. tx power of -12 dBmW and an advertising interval of 645 ms

our framework. Furthermore, these principles should be achieved in relative real-time, so information reaches the appropriate authority without much delay.

To facilitate the collection of meaningful and actionable service quality data, the platform design aims to ensure that feedback is solicited and recorded in relation to specific transit trips. We leverage the unique identifier values associated with each beacon to triangulate the ride. In the Singapore pilot, we were limited to installing beacons at only four downtown bus stops. In that case, we programmed the questionnaire solicitation mechanism conditional on the user's app registering signals from one of the four beacon-equipped stops. The questionnaires also began with a question confirming that the user was indeed taking the bus from that particular stop, ruling out potential false positives. This approach allowed us to identify the trip's origin stop; the app-based location tracing showed the route taken. In the Boston pilot, we were able to install beacons on all vehicles operating on the route of interest, as well as all bus stops along the route. This set up, in theory, would allow for triangulating the exact bus run by examining the beacon's unique identifier and by matching the geolocation timestamps with bus's real-time location (AVL) data.
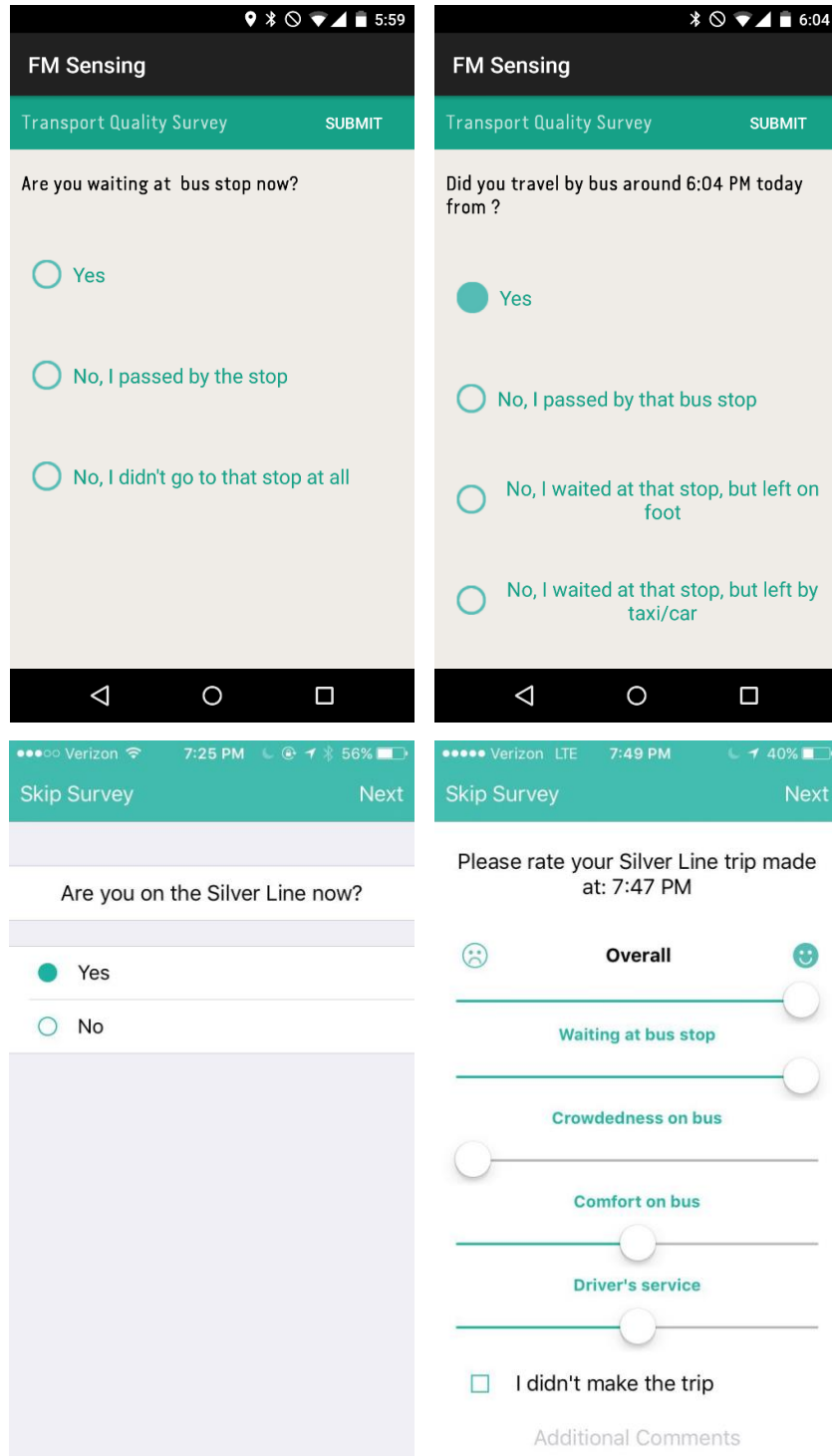
Figure 4.8 - FMS-TQ questionnaire trip confirmation screens

All FMS-TQ questionnaires contain a question confirming the user's bus trip, exhibiting the trip-specific nature of its data collection and intending to guard against false positives. Top row: screenshots from the Singapore pilot. Bottom row: screenshots from the Boston pilot.

A pre-requisite to receiving meaningful feedback is asking comprehensive questions relevant to the rider's transit experience. Passengers should have the opportunity to provide feedback on all aspects of their rides, and these attributes should be areas where the transit authority can realistically take action. Hence, for the two FMS-TQ deployments, the research team and the local agency partners invested much effort in refining the questionnaires. Table 4.2 displays the topics covered by the survey. The Singapore pilot covered virtually all aspects of the adapted RATER framework shown in Chapter 2; the only three attributes excluded were "safety" (not an issue in Singapore) and two "empathy"-related attributes (as earlier discussed, not measured on a trip-to-trip basis). The ongoing Boston pilot, striving to make the survey quick and easy to complete, keeps the list to only five attributes deemed the most critical by the MBTA. Aspiring to make feedback actionable, we ask participants of both pilots to provide specific reasons, if they express dissatisfaction with the service.

Table 4.2 - Service attributes surveyed by FMS-TQ in Singapore and Boston pilots

	Singapore	Boston
Topics covered by survey questionnaires	• Accessibility to bus stop • Stop conditions • Information at stop • The wait • Onboard crowding • Onboard comfort • Driver's service • Travel time • Convenience/connectivity • Reliability • Specific reasons of dissatisfaction, if any	• Bus stop condition • The wait • Onboard crowding • Onboard comfort • Driver's service • Specific reasons of dissatisfaction, if any
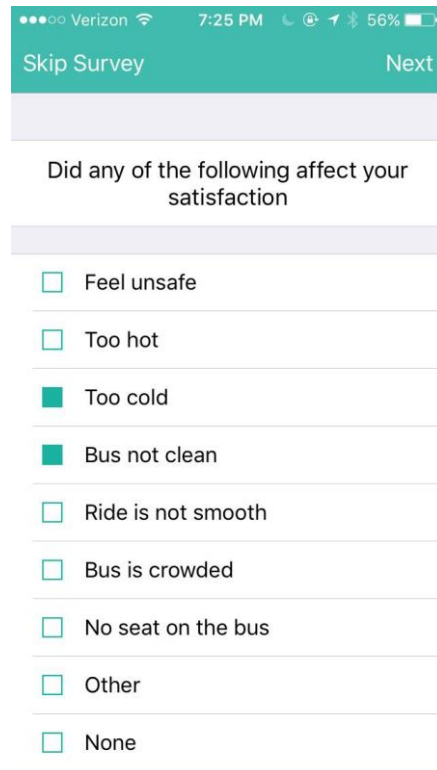
Figure 4.9 - Screenshot from FMS-TQ (Boston pilot), asking users to identify specific factors that affect their satisfaction

The aim to collect meaningful feedback should not compromise user experience. To this end, we strive to enhance the app's user-friendliness as well as minimize user burden. The in-app survey instruments contain only several questions, most of which can be answered with a single click on the smartphone. Visual aids help users easily indicate onboard crowding and their satisfaction levels (Figure 4.10). As for lowering user burden, we leverage the beacon-integrated FMS system to completely automate bus trip detection. Unlike precedents (Carrel, Sengupta, et al., 2015; Dunlop et al., 2015), FMS-TQ does not require any manual input to signal or record a bus trip. The system also limits the number of questionnaires sent to users per day through a sampling procedure. This is to preempt users from being annoyed with answering multiple questionnaires, in case they take many transit trips a day. In the Singapore pilot, the app randomly selected on each day whether to survey a user while she was waiting at the stop, travelling onboard, or just after alighting from the bus. In the Boston implementation, the algorithm stochastically decides which trip of the day to survey the user about. Regardless of what the algorithm selects, however,

users in Boston are provided with the opportunity to give feedback through a very short rating screen at the end of every trip detected, akin to the rating step at the conclusion of an Uber of Lyft ride. All these design elements are driven by the motivation to make the app easy and fun to use, with the intent of sustaining user participation.
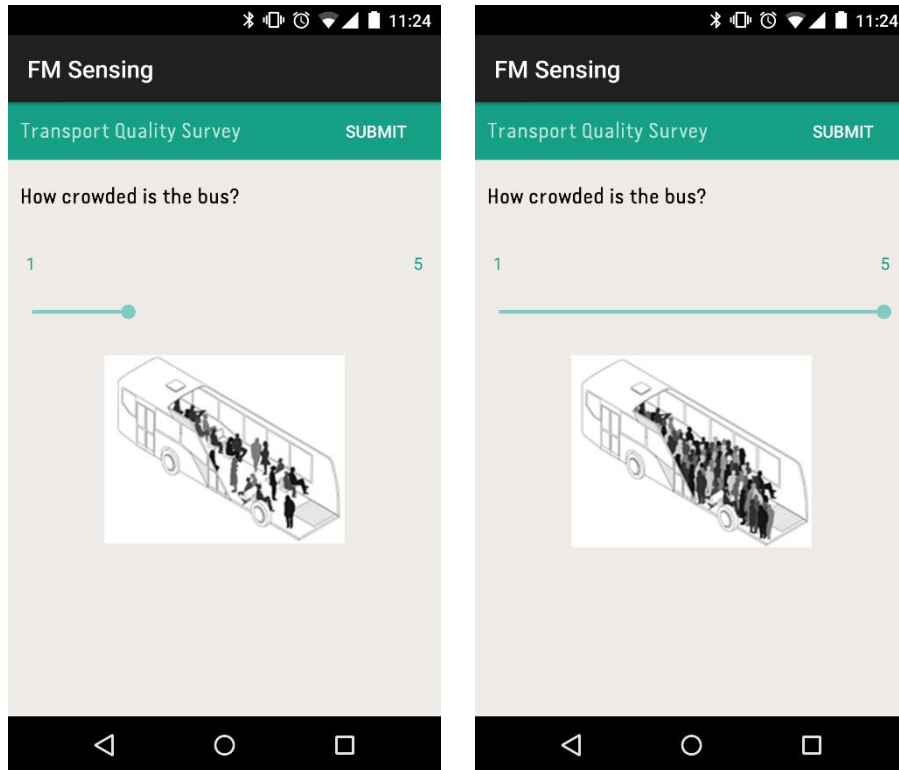
Figure 4.10 - Visuals help users answer questions that would otherwise lack a standard of judgement, such as crowding level

Data collection closer to 'real-time' not only helps passengers report their ride experience more accurately, but also offers value to transit operators in monitoring service quality. In both the Singapore and Boston implementations, FMS-TQ automatically records sensor-collected data, which can then be used to infer about the user's public transport experience – e.g., user's wait time at the stop, travel time on the bus, and acceleration profile of the ride. Feedback on transit service quality are solicited during or immediately after the ride, and uploaded along with sensor-collected data onto the FMS server at the

end of each day. Researchers can download the data and, with some post-processing and analysis, share insights with the transit agency.

Guided by these design principles, we adapt the platform to each pilot city's local context, described in the next chapter. Chapters 6 and 7 will discuss in detail the functionality, design and implementation of the FMS-TQ platform in Singapore and Boston, respectively.

Chapter 5

A tale of two cities: Singapore & Boston

Singapore and the Boston Metropolitan Area represent two very different urban transportation contexts, providing opportunities to examine the factors that may facilitate or hinder testing of innovative technologies for public transport. This chapter contrasts the two regions in terms of their urban contexts, public transportation systems, and the institutional structures and priorities underlying the two FMS-TQ pilots.

5.1   Two cities at a glance

Though Singapore and the Boston Metropolitan Area (from here on referred to as Boston) are only half a million apart in population and both boast a thriving economy, they differ in many other aspects relevant to this research. Table 5.1 compares these two regions. Singapore, an island city-state, is densely-populated (Department of Statistics, Singapore, 2015); Boston, situated on the easternmost part of the Massachusetts state, spans across a much larger area that includes some high-density cities but also many sparse suburbs (SAGE stats, n.d.). In Singapore, the public transportation mode share is three times higher than in Boston (Boston Region Metropolitan Planning Organization, n.d.; Land Transport Authority, 2011) -- more recent numbers show an even higher percentage of trips by public transport (63%) during peak periods (LTA, 2013). At the same time, due to a restrictive automobile licensing policy and very high auto ownership costs, Singapore's rate of private car ownership and driving mode share are much lower than those in Boston (Boston Region Metropolitan Planning Organization, n.d.; Land Transport Authority, 2016b).

Table 5.1 - FMS-TQ pilot sites' comparative context

Category	Characteristics	Singapore	Source	Boston-Cambridge-Newton Metropolitan Area	Source
Region	Area (km^2)	718	1	9,031	9
	Population	5,469,700 (2014)	2	4,812,658	10
	Population density (persons/km^2)	7,615	2	533	
	Income per capita (USD)	50,492 (2015)	2	64,311 (2014)	11
	Automobile ownership per 1,000 persons	95	3	600	12
Mode share (%)	Private transport	29	4	76	9
	All public transport	48	4	12	13
	Bus	25	4	5	13
	Rail (heavy rail, light rail, commuter rail)	19	4	8	13
	Walk/bike	23	4	6	13
	Other modes	0		1	13
Public transport demand	Unlinked daily boardings (millions)			1,297,650 (2014)*	14
	Unlinked subway boardings (2014)	2,899,000 (average day)	5	539,315 (typical weekday)	14
	Unlinked bus boardings (2014)	3,851,000 (average day)	5	387,815 (typical weekday)	14
Public transport network	# of municipalities served	1		175	14
	Total rail length (km) (2015)	183	6	102 subway, 620 commuter rail	14
	Number of rail routes (2015)	4 heavy rail, 3 light rail	7	3 heavy rail, 1 light rail (4 branches), 14 commuter rail	14
	# of rail stations (2015)	144	2	60 heavy rail, 53 light rail, 138 commuter rail	14
	Total bus route length (km)			1,213 (2014)	14
	Number of bus routes	357 (2013)	8	193 (2014)	14
	Number of bus vehicles	3,777 (2013)	8	991 (2014)	14
	Total bus kilometers travelled on typical weekday (2014)	901,700	2	over 137,600	14

Includes boardings for ferry, commuter rail, paratransit, and contract bus services, in addition to buses and heavy rail services.

Sources:
Singapore: (Department of Statistics, Singapore, 2015; Land Transport Authority, 2015a, 2015b, 2016b, LTA, 2014, 2015)
Boston: (Boston Region Metropolitan Planning Organization, n.d.; MBTA, 2014; ProximityOne, 2016; SAGE stats, n.d.; C. Zegras et al., 2016)

In addition to mode shares, the state of the public transport systems and their operating and governance structures also differ. Singapore's system has traditionally enjoyed a high reputation for efficiency (Credo Business Consulting, Siemens AG, 2014); its trains and buses deliver levels of service multiple times higher than Boston's (LTA, 2014; MBTA, 2014). Two private operators – SBS Transit and SMRT Corporation – run the bus and train services, regulated and overseen by the Land Transport Authority (LTA), a statutory board under the Singaporean Ministry of Transport. The LTA is also responsible for leasing operating licenses, constructing the rail lines, and planning and integrating the country's public transport services with the other land transportation modes. In contrast, the public transport services in Boston are primarily operated by a state agency, the Massachusetts Bay Transportation Authority (MBTA), which owns and maintains most of the transit assets.

In recent years, both public transportation systems have felt the increasing pain from capacity constraints and reliability issues, and have been open to collaborating with research institutions to test innovative ideas. The Singaporean government has been able to commit a great deal of resources into expanding and enhancing bus and rail services; the MBTA, though sharing the urgency to improve, often finds itself constrained by its dire financial situation.

5.2   Singapore

In the half-century since its independence in 1965, Singapore has rapidly transformed from a deteriorating, impoverished city into a modern metropolis with a highly competitive globalized economy. The government steered this resource-strapped city-state through several economic restructurings, gradually building up the manufacturing, finance, high technology, logistics and innovative technology pillars (Singapore Institute of Planners, n.d.). Conscious of its land scarcity, the country has adhered to transit-oriented development principles since the 1970s, building expansive road and mass transit networks across the island, coupled with its innovative housing policy that houses over 80% of Singapore's residents in high-density, mixed-use, affordable

housing communities (famously known as the HDBs)[10] (Housing & Development Board, 2016, p. 6). Most of the HDBs are situated along radial rail lines that bring people rapidly into downtown employment centers, and many public transport hubs are topped with large commercial complexes.

The Singaporean government has always placed a great emphasis on encouraging people to use public transportation over private motorized modes. In parallel with continued investments in the public transit network, detailed later in this Chapter, the government exerts strong effort to curb private car ownership and use with a hefty set of fees and taxes. In addition to a high import tariff on the vehicles, registering a new vehicle first requires obtaining a Certificate of Entitlement (COE), valid for 10 years, which was bidding at between S$42,000 ($30,000 US) and S$55,000 ($40,000 US) as of April 2016 (Land Transport Authority, 2016a). The total cost of buying and registering a Honda Accord in Singapore, for example, is about S$132,500 (or $96,600 US), compared to the $22,000 price tag in the US (Land Transport Authority, 2016d). The country further discourages automobile usage through congestion pricing, which has been in place since 1975 (Land Transport Authority, n.d.). Public transport is thus the dominant mode of travel in Singapore. According to the latest Household Interview Travel Survey (HITS) in 2012, public transport constitutes 63% of peak period trips (LTA, 2013). Indeed, for this densely-populated city-state, mass transit plays a critical role in the city state's viability.

5.2.1   The public transport system

Three modes make up Singapore's public transport system: rail services (heavy and light rail), buses, and taxis. Given this thesis's focus on public rail and bus services, as well as considering the comparability with the case of Boston, I will focus the discussion here on the first two modes.

Five heavy rail lines (called MRT, for Mass Rapid Transit) span the island roughly as they are named -- the North-South (red) Line, the North-East (purple) Line, the East-West

---

[10] The moniker comes from the public housing authority that develops and manages these affordable housing – the Housing & Development Board (HDB), a statutory board under the Ministry of National Development.

(green) Line, the Circle (orange) Line, and the Downtown (blue) Line (Figure 5.1). Three light rail lines (Light Rapid Transit, or LRT) span 43 stations, providing feeder services between the MRT stations and the public housing developments in the Bukit Panjang, Sengkang, and Punggol areas. Daily ridership is approximately 3 million trips for rail services (LTA, 2014). Major expansions of the rail network are currently underway. By 2030, the subway network is planned to expand to 360 km, from 183 km as of end 2014 (LTA, 2015), including a new 50 km Cross-Island Line and a 20 km Jurong Region Line, along with expansion of three existing lines (Land Transport Authority, 2013b). When completed, the new lines and extensions are projected to bring 8 in 10 Singapore's households within a 10-minute walk of a rail station (Land Transport Authority, 2013b).
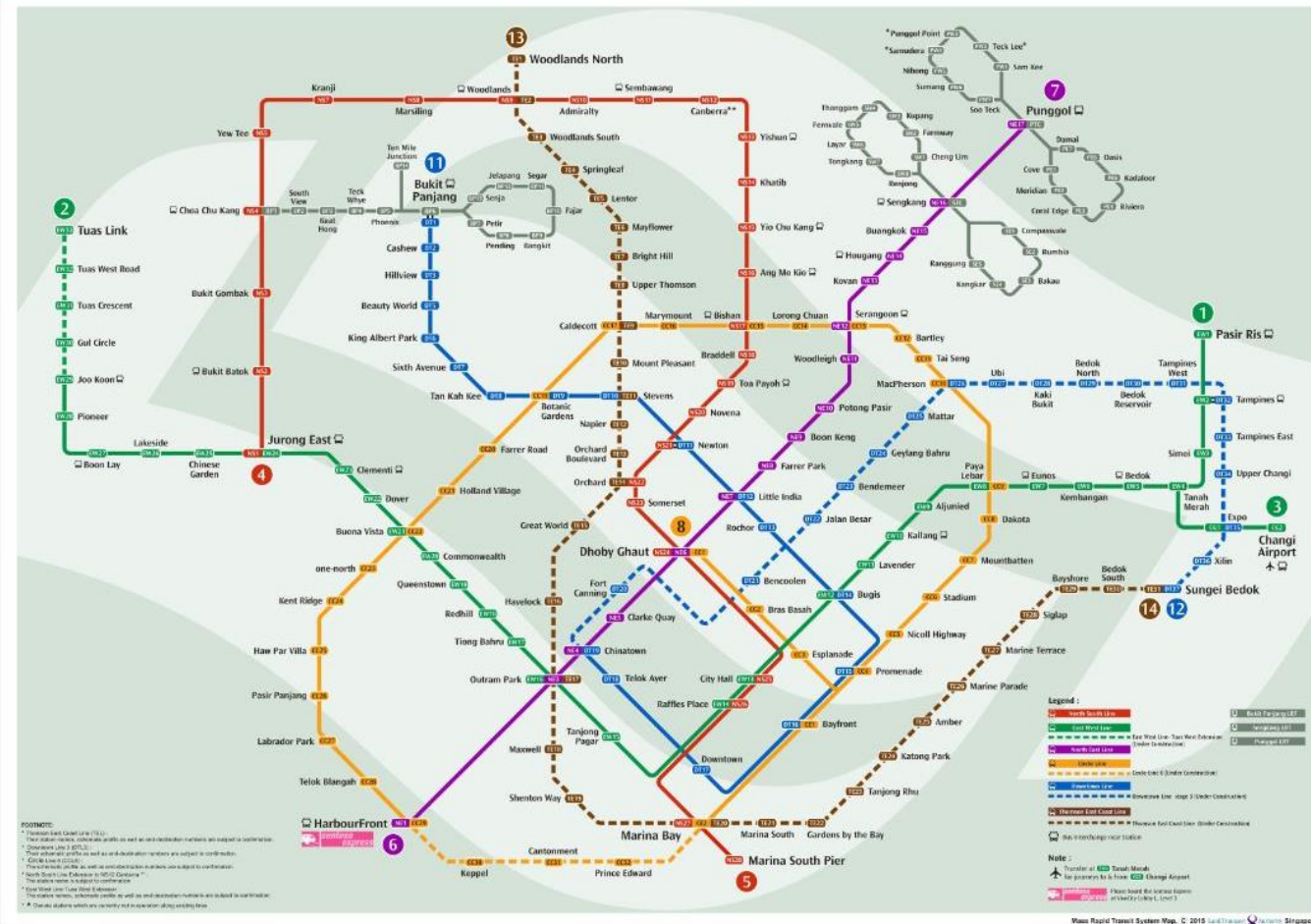
Figure 5.1 - Singapore's MRT and LRT system map

Source: Land Transport Authority

More than 3,700 buses operate more than 350 routes, serving 3.8 million trips daily (LTA, 2014). Bus services are classified into six categories (Land Transport Authority, n.d.):

- Trunk buses: traverse across the island;
- Feeder buses: offer transfers from rail stations and bus interchanges to surrounding housing estates and industrial areas. Some connect various neighborhoods within the same estate;
- Premium buses: serve between major housing estates and the Central Business District/business parks during the peak hours, at a cost premium;
- Shuttle buses: offer direct links to landmarks, amenities and places of interest, including housing estates, hospitals, business districts and tourist attractions;
- Night buses: from 11:30pm–2 am, the SMRT NightRider and SBST Nite Owl together serve 13 routes linking major nightspots and housing estates; and
- City Direct buses: add extra service capacity to/from the Central Business District during peak periods.
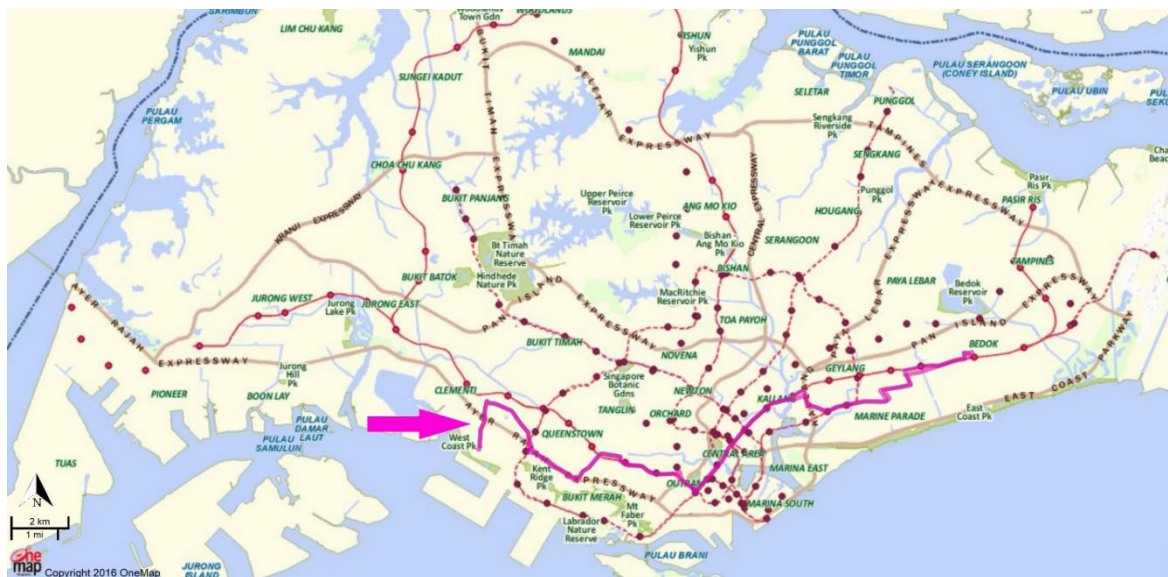
Figure 5.2 - Example of a trunk bus route (#33), route shown in magenta

Source: MyTransport.sg
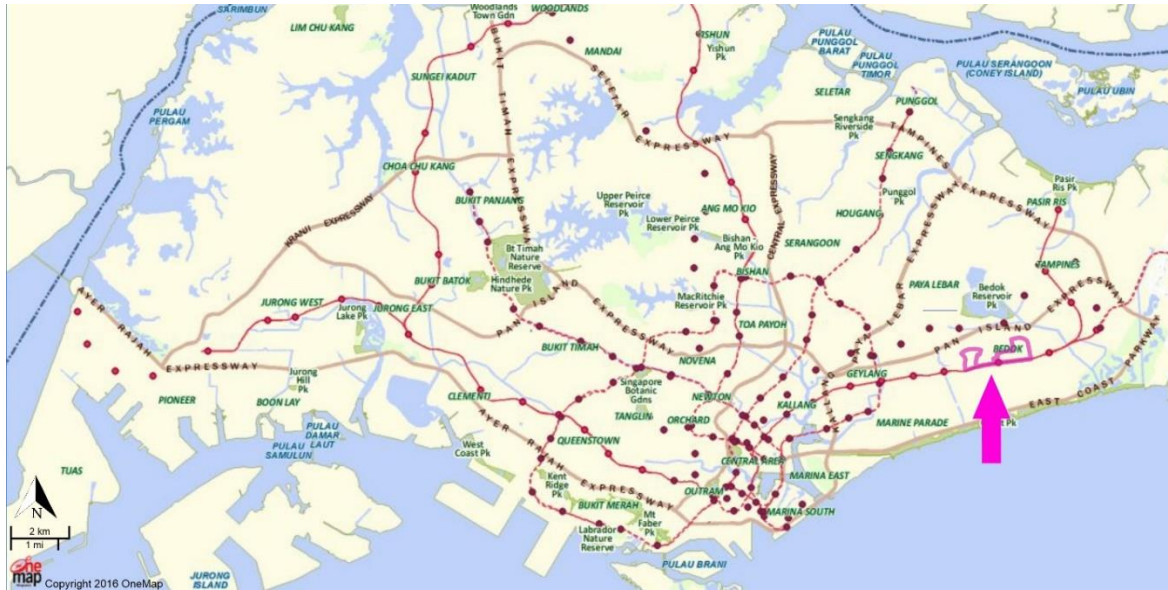
Figure 5.3 - Example of a feeder bus route (#222), route shown in magenta

Source: MyTransport.sg

Looking ahead to its future, the country is conscious of its major mobility-related challenges: increasing travel demand, limited land, and changing demographics and consumer expectations of transport (Land Transport Authority, 2013a). Singapore's population has already grown from 4.8 million in 2008 to 5.5 million in 2015, and generates 12.5 million trips each day (Land Transport Authority, 2013a). By 2030, the population is expected to grow by another 1.4 million and the daily travel demand to increase by an additional 50% (Howe-Teo, 2015; Land Transport Authority, 2013a). The island nation is scarce in land, and auto-oriented travel is resource-intensive; currently, roads occupy about 12% of Singapore's usable land, compared to 14% for housing (Howe-Teo, 2015). High-capacity public transport will thus have to be the backbone in serving ever-greater mobility demand. The public transport regulators and operators also have to uphold service performance to people's high expectations. As a result of both a growing user base and aging systems, Singapore has in recent years begun to experience problems that plague other transit agencies around the world: overcrowding and service disruptions. Train crowding was one of the sore points among voters during the 2011 General Election. Opposition parties jumped on how overburdened the transit system had become, as the Prime Minister apologized for the shortfall in infrastructure (Tan, 2015). In the first nine

months of 2015, the LRT had eight major breakdowns (defined as more than 30 minutes each) - double the combined number of the previous two full years (Tan, 2015). To a general pubic that holds the nation's transit to high standards, these issues are unacceptable.

Considering the importance of public transit to Singapore, keeping the system running smoothly and customers happy are high-priority items for the government. In addition to the earlier-mentioned rail network expansion, the Singapore Ministry of Transport has been rolling out an ambitious Bus Service Enhancement Programme (BSEP) to alleviate peak-time crowding – adding 1,000 vehicles and 41 new bus services by 2017 (Land Transport Authority, n.d.). The investment, albeit still in its early stage, has so far seen positive returns. Both public transport ridership and customer satisfaction of the service have been increasing each year since 2013 (Land Transport Authority, 2016c). The pressure and sense of urgency for further improvements, however, still loom large in the nation-state's agenda.

5.2.2   Land Transport Authority (LTA)

In midst of this political impetus for improvement and innovation, SMART FM/MIT entered into a collaboration with the LTA to test FMS-TQ. As earlier mentioned, the LTA is responsible for planning, operating, and maintaining Singapore's land transport infrastructure and systems. In terms of public transit, the LTA is responsible for developing and expanding the rail network, and serves as the central network planner for the city's bus services. While it does not operate any transit services, it works with the two operators in the city – SBS and SMRT – to identify areas of improvement and implement programs, such as the Bus Service Enhancement Programme mentioned earlier. At the time of the pilot, the LTA did not own any of the buses nor control most of the bus stops in the country, although it began restructuring the contracts with operators in late 2015, so that it would eventually own all operating assets and tender them out to operators (Singapore Government, 2015). This particular asset ownership detail ends up significantly influencing the FMS-TQ implementation design and our subsequent abilities to draw insights from the collected data.

85

The LTA's interest in exploring this new tool for assessing service quality and customer satisfaction was driven by two main factors. First, with the urgency to improve customer satisfaction and directives from the Ministry of Transport, the LTA was interested in initiatives that would make transit services more customer-centric, demonstrating to the public LTA's care for customer experiences. As mentioned in Chapter 2, the LTA had since 2006 conducted the annual Public Transport Customer Satisfaction Survey (PTCSS) – an intercept survey that solicited passengers' recalled satisfaction with their public transport experience. The ratings were aggregated into scorecards and published as a key performance metric (Land Transport Authority, 2016c). The FMS-TQ concept had the potential to lend new insights into customers' experience on public transport trips.

Second, the LTA was interested in exploring a number of research questions surrounding customer satisfaction. One question was the distinction between real-time versus recalled levels of satisfaction with transit services. Respondents' recall of an experience, as discussed in Chapter 2, would often differ from their actual experience due to psychological heuristics. The LTA was thus interested in taking advantage of FMS-TQ's capability to survey users during their transit trips, as well as retrospectively, to investigate potential correlations and distinctions between the two timeframes. The LTA was also interested in using customer happiness as a measure of satisfaction with transit services. Transit is supposed to improve people's quality of life by providing access to work, education, and recreation, as well as through improving the urban environment. So LTA contemplated whether gauging riders' well-being in surveys would better reflect service quality than satisfaction reports. As a preliminary exploration, we adapted FMS-TQ to incorporate happiness-oriented questions into the in-app questionnaires.

5.3   Boston

The Boston-Cambridge-Quincy Metropolitan Area (from here on referred to as the Greater Boston Area) has its development history rooted in public transportation. As early as the mid-1820s, hourly coaches crossed the Charles River, which had evolved by 1859 to over 40 miles of horse-drawn railway tracks throughout the metro region. This growth

accelerated in the last third of the 19th century and early 20th century, a period that saw the electrification of the streetcar network and the elevated and underground construction of the subway system (Block-Schachter, 2012, p. 69). As historian Sam Warner argues in *Streetcar Suburbs*, the street railway transformed Boston from a merchant town into a modern, divided metropolis -- an inner city of commerce and slums and an outer city of commuter suburbs (Warner, 1978). But with the changing of times and favorable economic and urban redevelopment, today's Greater Boston Area has become the scene of a thriving high-tech and innovation economy. These industries, along with the large number of higher-education institutions, draw many people to work, study, and live in Boston and the surrounding municipalities. Today, the Greater Boston Area is home to more than 4.7 million residents (U.S. Census, 2014). While the coaches and streetcars have transformed into subways, light rail, and bus lines, many of the public transport routes today still follow the trajectories of the legacy streetcar network.

5.3.1   The MBTA

The Massachusetts Bay Transportation Authority, commonly called the MBTA or the T, is the public transit operator for the Boston Metropolitan Area. It is the fifth largest transit system in the United States, serving on average 1.3 million trips each weekday (APTA, 2015, Table 30). Its service area covers 175 cities and towns, with a total population of over 4.8 million (MBTA, 2014, p. 2). It operates four urban rail lines (three heavy rail and one light rail line with four branches), 174 bus and trackless trolley routes, 14 commuter rail lines, and three ferry services (MBTA, 2014, p. 4). The MBTA is governed by a board of directors, an executive management team led by the General Manager, as well as an Advisory Board that represents the towns and cities in the MBTA's service district.

Figure 5.4 - MBTA service routes – subway, key bus routes, commuter rail, and ferry lines

Source: MBTA
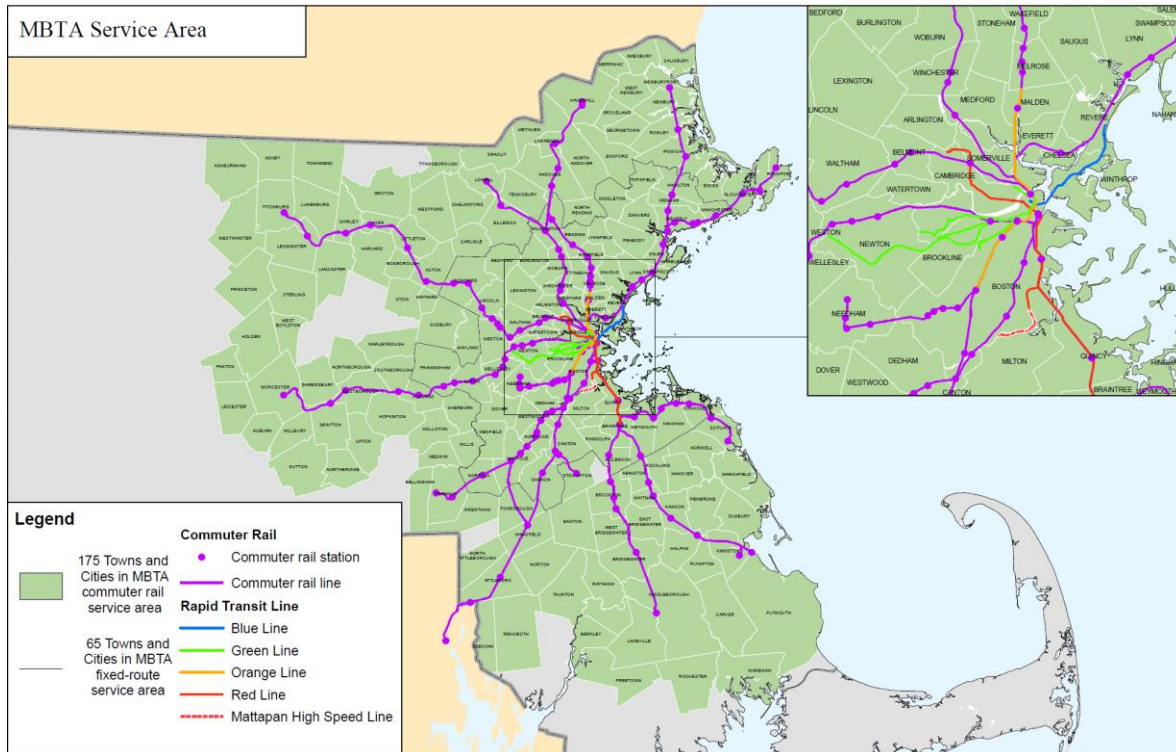
Figure 5.5 -MBTA service area

Source: MBTA

An aging system with inadequate funding, the MBTA has met with increasingly severe challenges over past three decades. Fast-growing expenses have far outpaced revenues – the projected budget shortfall is $200 million (10% of the $2 billion budget) for FY2016, $360 million by FY2019, and $800 million by FY2024 (Massachusetts Taxpayers Foundation, 2015). Debt service on the $9 billion debt – the largest among transit systems in the U.S. – adds to the expense and constrains the T's infrastructure investments. As with the operating budget, the MBTA also struggles for funding to maintain its aging infrastructure. Today, the State of Good Repair backlog sits at almost $7 billion (Massachusetts Taxpayers Foundation, 2015). This translates into service interruptions and delays for customers. The on-time performance rate for key bus routes and the bus rapid transit system is only about 70%. In 2015, passenger wait time metrics all trended in the unfavorable direction from prior years (MassDOT, 2015, p. 33). In addition to the financial woes, the T has also suffered from poor accountability and management (Governor's Special Panel, 2015). It is under this backdrop that the severe winter

89

conditions in 2015, as described at the very beginning of this thesis, caused the T to plunge into chaos. To many, the transit system became synonymous with incompetence.

This heightened urgency for reform helped plant the seed for our collaboration with the MBTA. As part of the official response to the winter crisis, a Fiscal and Management Control Board (FMCB) for the MBTA was established in July 2015. One of the key objectives of this board is to work with the MBTA to utilize performance metrics to improve the system's operations, transparency, and customer experience. With a new General Manager and a new MassDOT Secretary in place, the T aims to pursue more innovative, data-driven solutions to improve its operational performance and public image. Similar to our LTA partners, the Strategic Initiatives team at the MBTA is very interested in innovating the way they devise and use performance metrics, in addition to the monthly Customer Opinion Panel Survey discussed in Chapter 2, to better capture customers' experiences. Hence, when the team learned about our platform in September 2015, their strong interests and institutional support made the project come to being.

5.3.2 The Silver Line

For the scope of the pilot, we decided on the Silver Line. The Silver Line (SL) is Boston's sole bus rapid transit (BRT) line, with a total of four branches (Figure 5.6) and a daily ridership of over 33,000 on average weekdays (MBTA, 2014, p. 45). With only limited dedicated infrastructure and related elements, the Silver Line is not considered, by global standards, to be "true BRT" (Weinstock, Hook, Replogle, & Cruz, 2011). The Washington Street section has two routes from Dudley Square in Roxbury, mostly via Washington Street, to Boston's Downtown Crossing (SL5) and South Station (SL4). The Waterfront section runs from South Station to South Boston (SL2) and to Logan Airport (SL1). A bus tunnel linking these two corridors had been proposed, but the project is postponed indefinitely due to budget constraints and stakeholder opposition. A formerly named SL3 service ran from South Station to City Point via the Boston Marine Industrial Park; the service was discontinued in 2009 due to low ridership (MBTA, 2014, p. 42).
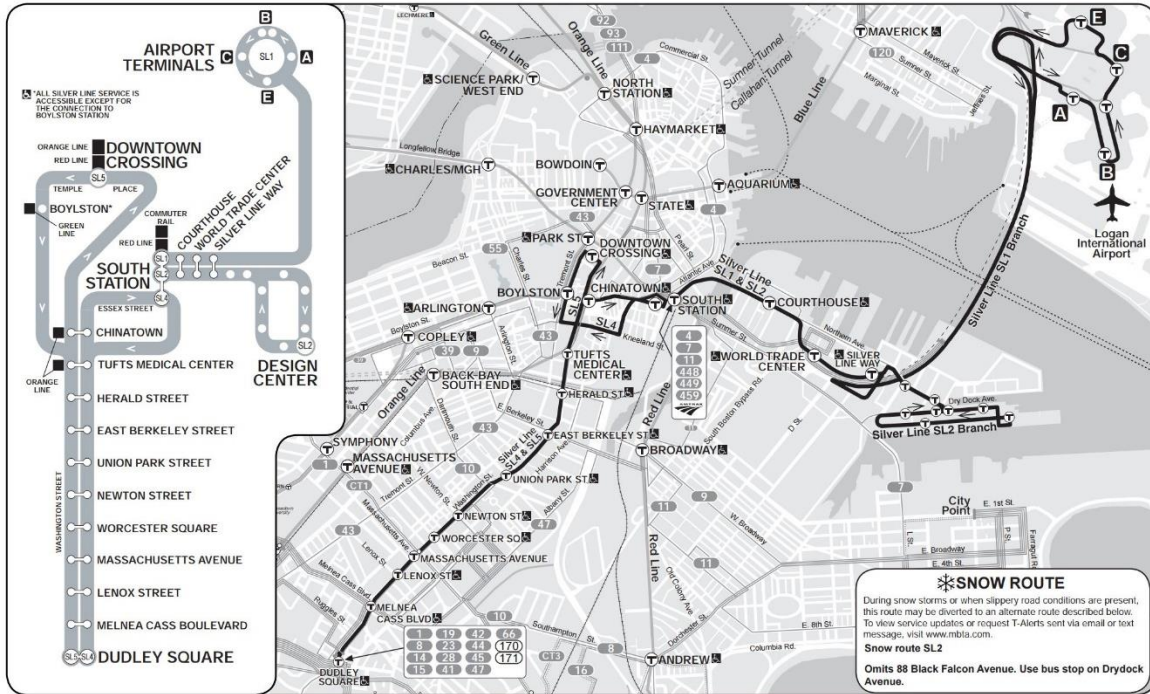
Figure 5.6 - Silver Line routes

Source: MBTA

Figure 5.7 - Silver Line vehicle and stop

Source: MBTA

We chose the Silver Line (SL) for two main reasons. First, it plays a significant role in Boston's urban transportation. All four branches of the Silver Line rank in the top 20 bus routes by typical weekday ridership (MBTA, 2014, p. 52). The system serves important corridors – between Downtown Boston and Logan Airport via the Seaport District, as well

as between South Station and Dudley via Downtown Boston, Chinatown, and Tufts Medical Center. Notably, it is the only mass transit service for the Seaport District and Logan Airport, two areas that generate large, ever-growing travel demand.

Equally, if not more, importantly, the Silver Line provides a close to ideal setting in terms of implementation logistics. The Singapore pilot illustrated clearly the disadvantage of installing beacons at bus stops only, as multiple bus lines can operate along the same route, hence weakening our ability to identify the precise bus route and run on which the user is reporting. Fortunately, as MBTA is the owner of its buses, they are able to quickly provide permission to install beacons on buses in Boston. The easiest way to enable trip identification is to limit the pilot to one or more bus routes in their entirety. In contrast to the rest of the MBTA's bus routes, the Silver Line operates special buses, 56 vehicles dedicated exclusively to SL services. In comparison, all regular MBTA bus services run on an interlining basis – vehicles pulling out of a given garage can be used on a number of different routes, sometimes switching routes between shifts on the same day. This would make identifying and installing beacons on vehicles on another specific MBTA bus route impossible. Furthermore, almost all SL stops have shelters, providing adequate infrastructure for installing the at-stop beacons. We had also briefly considered conducting a comparative study encompassing the Green Line, a light-rail service. Upon finding out the Green Line was served with a fleet of 205 vehicles which rotated through the different branches, we abandoned the idea – time and resource constraints would not have allowed us to install so many beacons.

The Waterfront services --SL1 and SL2 -- have nine and 12 stops/stations, respectively. From South Station, the services run for three stops in an exclusive busway tunnel, before continuing the rest of the way on surface streets. After the fourth stop – Silver Line Way – SL1 heads towards Logan Airport to service all four terminals there, while SL2 continues in the Seaport District. Both services loop back to re-enter the busway tunnel and return to South Station. Passengers enjoy free boarding at the airport on the SL1, courtesy of MassPort (the airport operator) in attempt to enhance customer convenience and reduce bus dwell time. A shorter SL2 circuit, designated the Silver Line Way (SLW) shuttle, runs between South Station and the Silver Line Way stop during peak periods only. Combined ridership on the three Waterfront services is approximately 14,000

on average weekdays (MBTA, 2014, p. 45). Most of the corridor's riders transfer to or from heavy rail at South Station. Given the rapid growth in residential and office real estate in the Seaport District, as well as air travel demand, ridership on the Waterfront routes will almost certainly continue to grow.

This growth will challenge Silver Line's service operations. City leaders and planners had originally forecast the number of jobs in the Seaport District to grow much more slowly, allowing a better mass transit system to be in place by the time the district fully developed. However, as success came sooner than expected, so did over-capacity and gridlock. The Waterfront lines encounter significant crowding during the peak periods, and traffic congestion in the district slows down vehicular travel times (S. Leung, 2013a). To make matters worse, a multi-year, mid-life overhaul effort for the dedicated Silver Line vehicles began in 2014, further reducing the fleet available for service (S. Leung, 2013b).

The Washington Street services – SL4 & SL5 -- run jointly between Dudley Square and Tufts Medical Center for 11 stops, before splitting eastward (SL4) towards South Station and westward (SL5) towards the Boston Common. Most of the Washington Street corridor features dedicated, though not segregated, bus lanes, although buses are still subject to congestion in the portion of the corridor without dedicated lanes in Downtown Boston (Stewart, 2014). This is one of the busiest corridors in MBTA's service area, with average weekday boarding increasing from 13,000 in the year of inauguration (2002) to almost 19,000 in 2012 (MBTA, 2014, p. 45; Stewart, 2014). While it does not have many of the features critical to a "true" BRT – stations with off-board fare collection or platforms level with the bus floor -- its infrastructure nonetheless represents an upgrade from that of regular bus services. The vast majority of SL4 & SL5 stops are equipped with real-time information signs, and all stops along Washington Street have a canopy and heating system, bike racks, artwork showcase, and call boxes for contacting police or customer assistance (Stewart, 2014, p. 44).

Figure 5.8 - Silver Line stop infrastructure on Washington Street

The historical context of the Washington Street branch adds significance to our initiative. The corridor used to be served by the elevated Orange Line (heavy rail) from 1901 until 1987. The Orange Line was then moved west towards Back Bay, running on a mix of surface and underground tracks to Forest Hills. Though the MBTA initially promised to replace the Orange Line service on Washington Street with light rail, it ultimately deployed the Silver Line BRT as a replacement. This resulted in much public criticism, especially considering Dudley Square is a historically African American community with relatively poor transit services. The Silver Line, with relatively poor levels of service is viewed by many as an inferior replacement to the original Orange Line service. The Silver Line project has thus been referred by some as "The Silver Lie" (Stewart, 2014). Given these issues and negative public perception of the Silver Line, the FMS-TQ project, which seeks to better understand passenger experience and improve service, is especially relevant.

Chapter 6

Proof-of-Concept Pilot in Singapore

This chapter details the first FMS-TQ pilot, conducted in Singapore. This initiative served as the proof-of-concept test, involving only a small number of participants from the local public transport regulating agency. The pilot proved the feasibility of some of our intended capabilities for the platform, but also highlighted areas of weakness. It provided the foundation and informed improvements for the subsequent implementation in Boston, which was much broader in scope (covering an entire bus line) and involved a larger number of participants recruited from the general public.

6.1   Trial design

FMS-TQ (Singapore) had five surveying instruments: an entrance questionnaire, three event-based questionnaires for the three bus trip stages – when the user was waiting at the stop, travelling onboard, and alighting from the bus – and an end-of-day questionnaire. To mitigate user burden, each questionnaire consisted of only a few multiple choice questions, and each user received at most only two questionnaires per day: one randomly-selected event-based questionnaire and the end-of-day questionnaire. When the app detected the beacon signals for the first time that day, the app would select which event-based questionnaire to generate using a sampling algorithm, as illustrated in Figure 6.1. For example, if the surveying system selected number 2 for a particular user that day, the user would receive an onboard questionnaire while she was riding the bus. Notice that algorithm, as shown in Figure 6.1, also contained another process, which ran at the end of each day depending on whether at least one bus trip was previously detected on that day. This end-of-day questionnaire intended to collect users' retrospective reflection on their bus experience earlier that day.
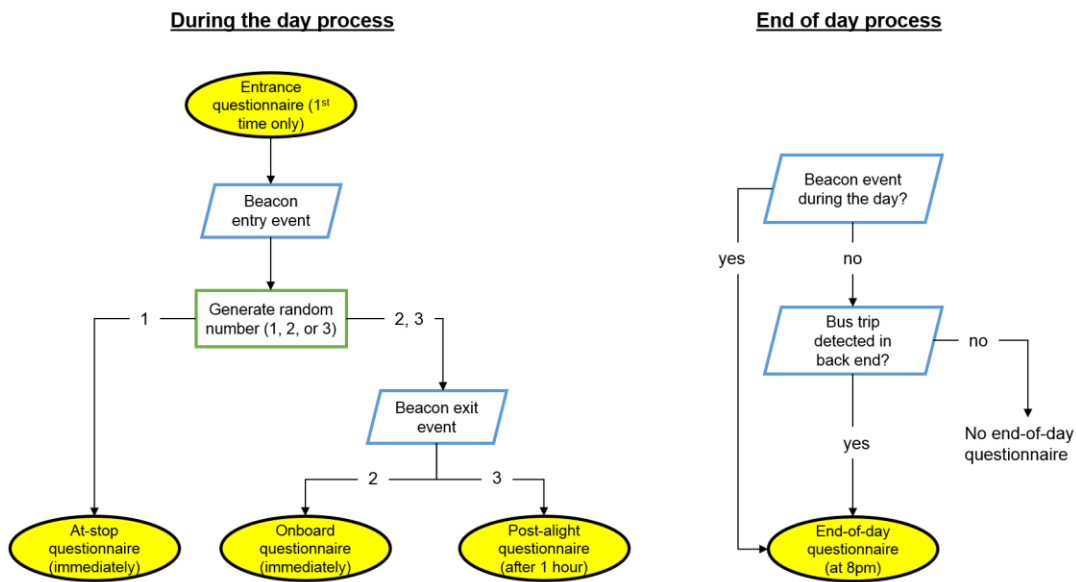
Figure 6.1 - Trip stage detection and survey solicitation logic

While the event-based questionnaires were designed to be answered in real-time, FMS- TQ was designed to accommodate late responses as well. In cases where the user failed to respond to any of the real-time questionnaires within the first hour, the question phrasing would change from the present to past tense. The system also monitored and analyzed people's response times and rates, to provide information for future refinement of the app and survey design.

The design of the survey logic was in large part influenced by institutional constraints. First, the buses were owned by the private operators, which limited the possibilities for installing beacons on the vehicles in a timely manner. Second, the LTA had limited authority over street bus stops. Given our target population for the pilot (LTA employees at one of the LTA offices), we chose stops near the LTA office for which the LTA had the rights to give us permission to affix beacons on (Figure 6.3). Weatherproof beacons were mounted in a strapped pouch near the top of the bus stops (Figure 6.4). This set-up aimed to maximize signal transmission and facilitate installation and removal without modifications to the stop infrastructure.
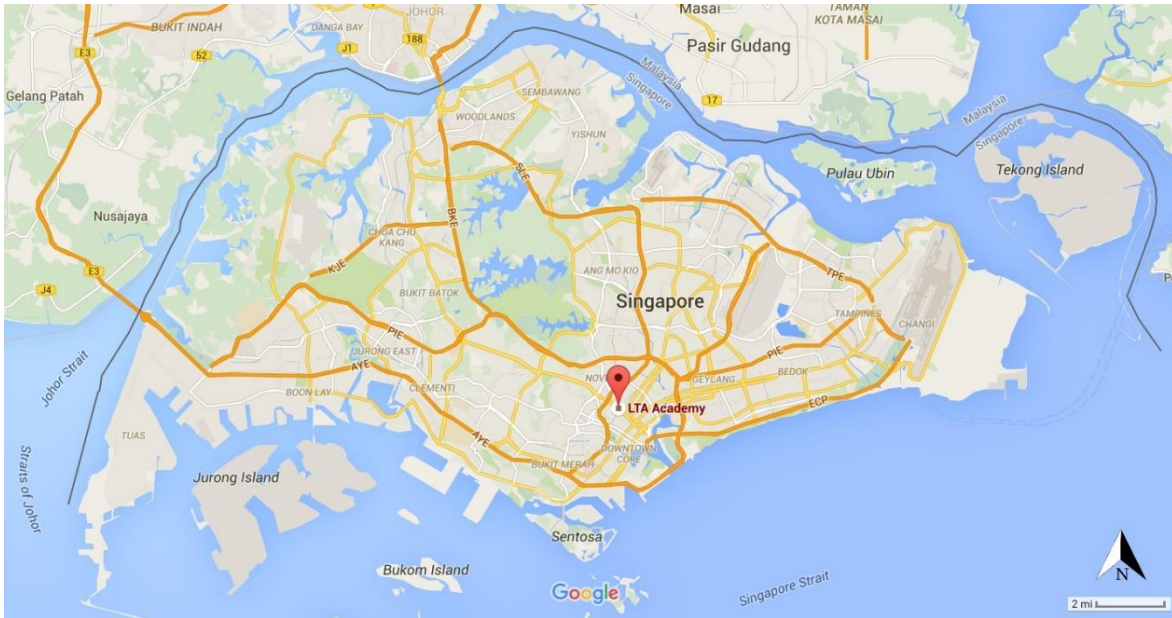
Figure 6.2 - Location of the pilot stops in context of Singapore

Source: Google Maps
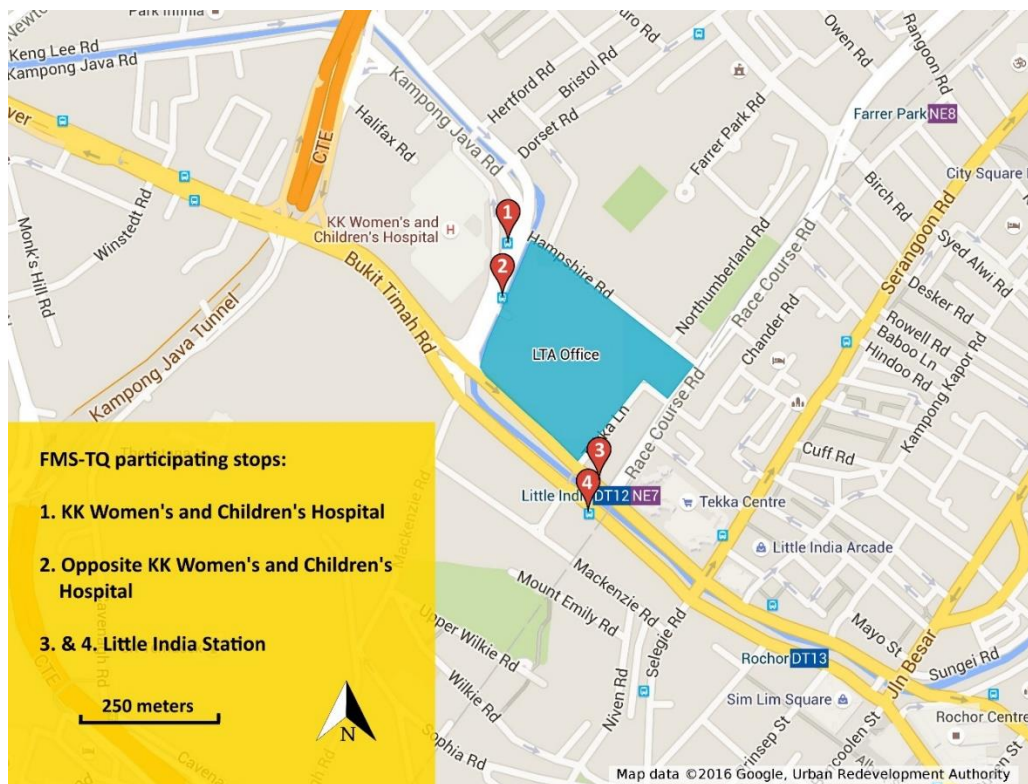
Figure 6.3 - Participating bus stops

Source: Google Maps

Figure 6.4 - Beacons installed at bus stops

The questionnaire design, undertaken in close collaboration with LTA, aimed to balance research and practical objectives, attempting to encompass the entirety of the bus trip experience while enabling the capture of responses in real-time. The 14 relevant service attributes, derived from the RATER framework, were to be assessed during the most relevant bus trip stage (Table 1). For example, passengers' satisfaction with wait time should be solicited in the onboard questionnaire, after riders boarded the bus; judgment on directness of route/convenience could be salient during all three stages of a bus trip. This categorization guided content of questions in the at-stop, onboard, and post-alight questionnaires.

Table 6.1 – Service attributes to be assessed

Bus service attributes	Trip Stage of Measurement		
	At stop	Onboard	Post-alight
Accessibility to bus stop	X		
Stop facilities	X		X
Wait time		X	
Travel speed/time		X	X
Seating & personal space	X	X	
Onboard comfort		X	
Directness of route/ need for transfers	X	X	X
Reliability of wait time		X	
Reliability of being able to board bus		X	
Reliability of travel time			X
Availability & accuracy of information	X		
Bus driver's skills		X	X
Attitude and quality of customer service		X	X
Assistance to customers when needed		X	X

Below summarizes the content and app logic involved in the overall survey and generating the five specific questionnaires to be implemented at different moments; a copy of the questionnaires can be found in Appendix A.

Entrance questionnaire (Figure 6.5): After installing the app, the user received an initial intake questionnaire to collect baseline information, including her demographic information, travel habits and attitudes, general satisfaction with public transport services, and general satisfaction with life.

At-stop questionnaire (Figure 6.6): Whenever a user arrived at a beacon-equipped bus stop with FMS-TQ installed and running on her smartphone, the phone would pick up the beacon signal. The beacon trigger, combined with subsequent GPS-based sensing, was intended to provide information which would enable subsequent estimation of the user's wait time at the bus stop via post-processing. If the app had randomly selected the at-stop questionnaire for the user and she had been detected to be the stop for at least 60 seconds, the instrument would appear on the user's phone screen. The questionnaire first asked for confirmation that the user was indeed waiting at the

specific bus stop. Upon confirmation, the user would be asked about the trip's purpose, whether any transfers were involved, satisfaction with the bus stop condition (accessibility, cleanliness, real-time information), and her physical and emotive state (e.g. tired, anxious, relaxed) while waiting.

Onboard questionnaire (Figure 6.7): The app would infer the user boarding a bus when the phone departed the beacon signal area. Similar to the at-stop procedure, if the app randomly selected to survey the user onboard, the questionnaire would first verify that the user was indeed travelling on the bus. If so, the app would then pose seven questions aimed at collecting three types of information: 1) subjective customer satisfaction of the wait and onboard experiences, 2) reasons for any dissatisfaction, and 3) observations on onboard crowding level. The latter two types of questions intended to provide feedback on service quality from a relatively objective perspective, covering bus stop condition, service information availability, crowding, comfort, and the driver's service.

Post-alight questionnaire (Figure 6.8): If selected, the post-alight questionnaire would begin one hour after the user exited the beacon area (bus stop). The questionnaire targeted the overall bus travel experience, gauging people's perception of their travel times, convenience, and overall satisfaction with the service. Since a rider's perceived bus experience could be much influenced by his/her activities during the trip, the questionnaire also asked about the user's onboard activities.

End-of-day questionnaire (Figure 6.9): At the end of each day, all users who had taken a bus trip that day were given an end-of-day questionnaire. We determined if a user had taken a bus trip in one of two ways: (1) she completed one of the bus trip stage questionnaires; or (2) we inferred a bus trip based on FMS's back-end analysis of her travel data. The end of day questionnaire was always sent to users in the 8pm to 10pm window. This questionnaire aimed to collect respondents' reflections on their bus experience and their evaluation of bus's impacts on their lives and happiness that day. It also asked users the degree to which their bus experience had met their expectations. Compared to the previously mentioned questionnaires, which focused more on specific service attributes, this retrospective questionnaire targeted riders' broader well-being

and travel choices. The pairing of responses to the trip-segment and end-of-day questionnaires intended to provide possibilities for comparing real-time and retrospective passenger satisfaction.
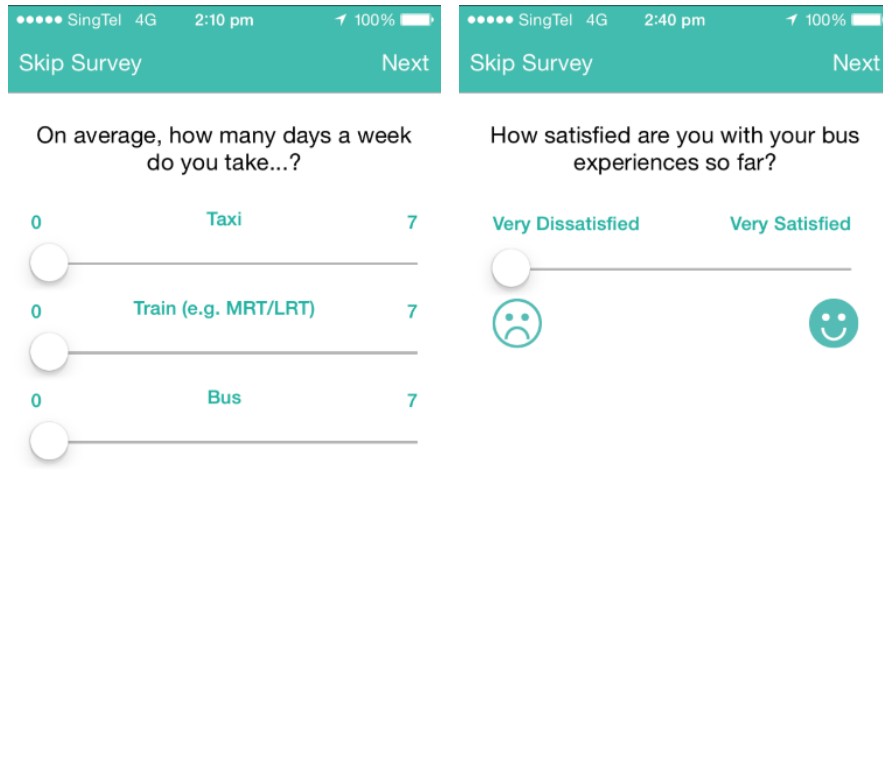
Figure 6.5 - Screenshots of select questions from the entrance questionnaire
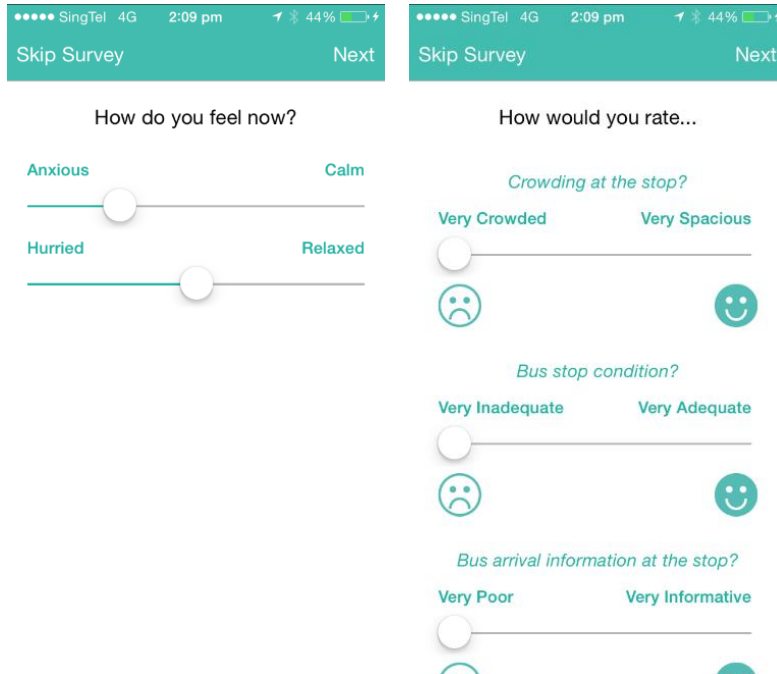
Figure 6.6 - Screenshots of select questions from the at-stop questionnaire

Figure 6.7 - Screenshots of select questions from the onboard questionnaire
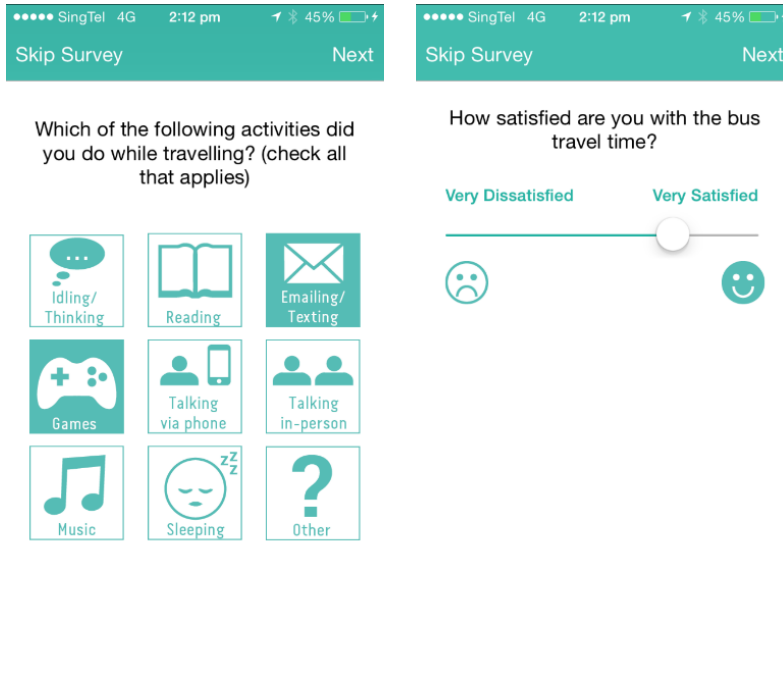
Figure 6.8 - Screenshots of select questions from the post-alight questionnaire
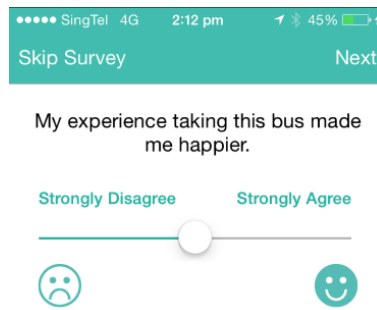
Figure 6.9 - Screenshots of select questions from the end-of-day questionnaire

In addition to passengers' responses, the app also automatically collected data from users' phone sensors, such as GPS coordinates, associated timestamps, and accelerometer data.

Since the Singapore pilot was a proof-of-concept test and open only to LTA employees, we opted for a small-scale incentive that largely appealed to intrinsic motivations. The recruiting material encouraged potential participants to "help innovate Singapore's bus service" by "transforming the way we sense public transport service quality and customer satisfaction." We also offered a prize draw of two SGD$100 gift cards, with chances of winning directly related to the number of questionnaires completed during the course of the pilot.

6.2   Implementation

The LTA disseminated an email invitation to its employees, appealing mostly to intrinsic motivations to help innovate research on public transport service quality. Android and iPhone users were invited to download the FMS-TQ app from Google Play and the iOS App Store, respectively, and then prompted to register for an account. The pilot began on June 18, 2015 and ended on July 4, 2015.

Despite the extensive app testing that the engineering team had conducted prior to the pilot, two issues did not manifest until later. First, only after the pilot began did we learn of a constraint for the Apple phones which prevented the app from automatically turning on the phone's Bluetooth. This meant that iPhone users needed to manually enable Bluetooth on their phones, prior to reaching the origin bus stop. Hence, we sent a follow-up email to all iPhone-using participants regarding this requirement. Another bug in the app affected the sampling generation process, resulting in post-alight questionnaires initially not being generated. The bug was fixed about halfway through the trial, but the ultimate number of responses to post-alight instruments was still lower than a proportional share.

Upon the trial's conclusion, we solicited feedback on participants' experience and suggestions via a web survey. We shared the responses with LTA colleagues.

6.3 Findings

In total, 32 users initially registered and 24 completed the entrance questionnaire. Unsurprisingly, the group was dominated by bus riders - 15 respondents reported of taking the bus every day, and five reported bus usage six days a week. The most common reason for choosing bus over other alternatives was convenience. Despite their employment affiliation with the LTA, less than half of the surveyed agreed with the statement "I like using public transport." In fact, the respondents may well be more critical of the local public transport services than the general population; only 46% reported being "satisfied" with their bus experience in Singapore - compared to 88% of those surveyed in the 2014 LTA Public Transport Customer Satisfaction Survey (Land Transport Authority (Singapore), 2015).

Over the course of the pilot, a total of 129 questionnaires were completed in addition to the 25 entrance responses: 22 at-stop, 23 onboard, eight post-alight, and 51 end-of-day (Table 6.2). The low number of post-alight responses reflects the previously mentioned bug in the software discovered after the pilot had begun.

Table 6.2 - Number of respondents and Responses

Questionnaire	Number of responses (Number of false positives)	Number of respondents
Entrance	25	25
At-Stop	22 (8)	10
Onboard	23 (9)	9
Post-alight	8 (5)	6
End-of-day	51	14

In order to enable real-time surveying, the system must be able to detect trips as they happen. Results show that FMS-TQ did successfully detect bus trips with essentially no user input, although it is less accurate than we had hoped. 26 bus trips detected through beacons were not caught by the location data-based mode detection (post-processing) in the backend, illustrating BLE beacon's advantage in signaling a waiting-for-bus event when location data are noisy or absent. However, there was also an unexpected number of false positives – waiting event detected by the app but invalidated by the user. Among the 53 beacon-triggered surveys, 22 were considered false positives – the user

verified that s/he was not taking a bus from the specific bus stop. Analysis of the questionnaire time and location stamps, along with exit survey responses, suggested that the false positive issue could be fixed by adjusting the app settings. For many of these false positive cases, when the user alighted at a beacon-equipped stop, the app picked up the signal and misinterpreted it as the person waiting or boarding at that stop. Though we had set a minimum questionnaire trigger threshold (user being in the beacon signal uninterrupted for 60 seconds) to prevent precisely such false positives, empirical results suggested that the threshold should have been set even more conservatively (i.e. greater than 60 seconds).

While the automatic trip detection was not 100% accurate, the user responses gathered still demonstrated FMS-TQ's capability to gather information on bus service quality. First and foremost, we were able to collect customer satisfactory ratings just as conventional survey methods could. Considering the small sample size and the lack of representativeness of the participant pool, the exact numbers or distributions of answers are not important – what is important is that every data point contained rich geo-location and time stamp information. I will discuss below the results on a few service attributes that exhibit interesting patterns that would not manifest in conventional surveys.
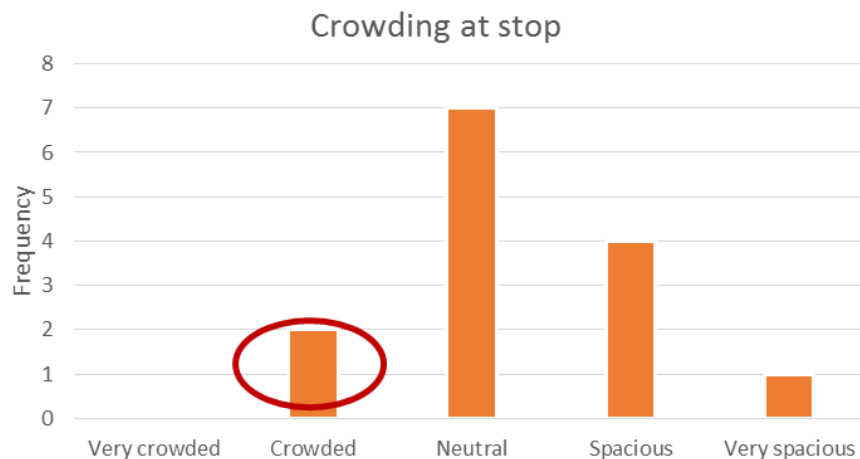
Figure 6.10 - Passenger feedback on stop crowding level

As Figure 6.10 shows, there were only two reports of bus stop crowding; both of

them took place in the evening rush hour. As shown in Figure 6.3, the participating bus stops were all located in a busy area in downtown Singapore, close to employment centers, commercial establishments, and a subway station. Intuitively, these stops tend be more heavily used by evening commuters, and should see relatively fewer passengers on reverse commute during the morning peak. This temporal pattern was mirrored in the onboard crowding levels (Figure 6.11) -- more crowding was reported on evening commutes.
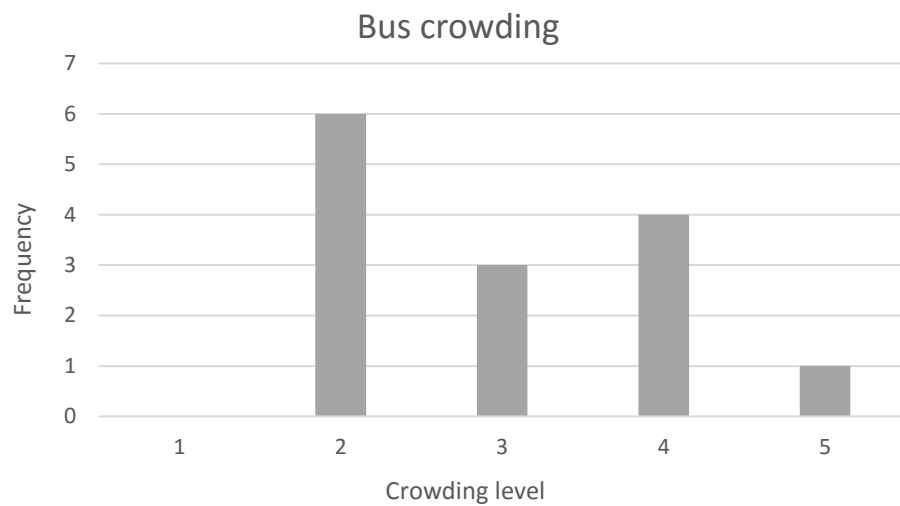
Figure 6.11 - Passenger feedback on bus crowding level

1.   2.   3.
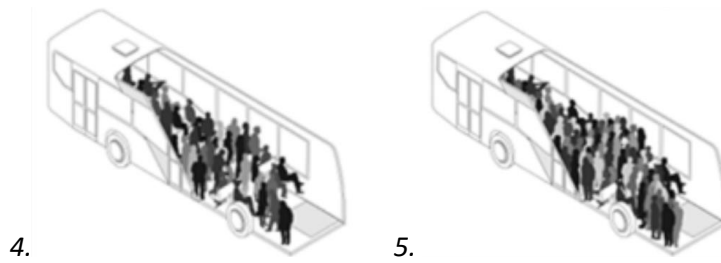
Figure 6.12 - In-app visual references for various crowding levels

Source: (Batarce et al., 2015)

The data also illustrated the intuitive connection between reported crowding levels and onboard comfort. On trips where crowding was more severe, reported onboard comfort levels tended to be lower. In fact, we can go one step further in drawing connections between passenger comfort and various other service attributes. As the onboard questionnaire prompted passengers for detailed reasons in case they reported being less than comfortable on their trip, we can associate these details to comfort and crowding attributes to gain a fuller picture of individual trip experiences (Figure 6.13).
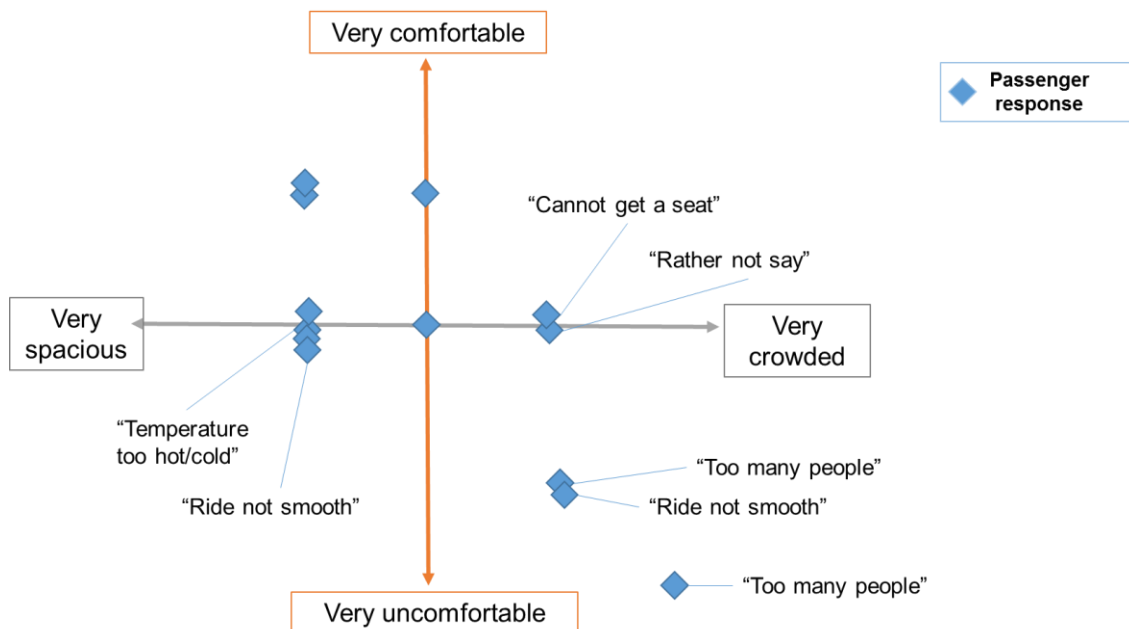
Figure 6.13 - Detailed qualitative feedback on onboard comfort and crowding conditions

Bus stop condition

Information at stop
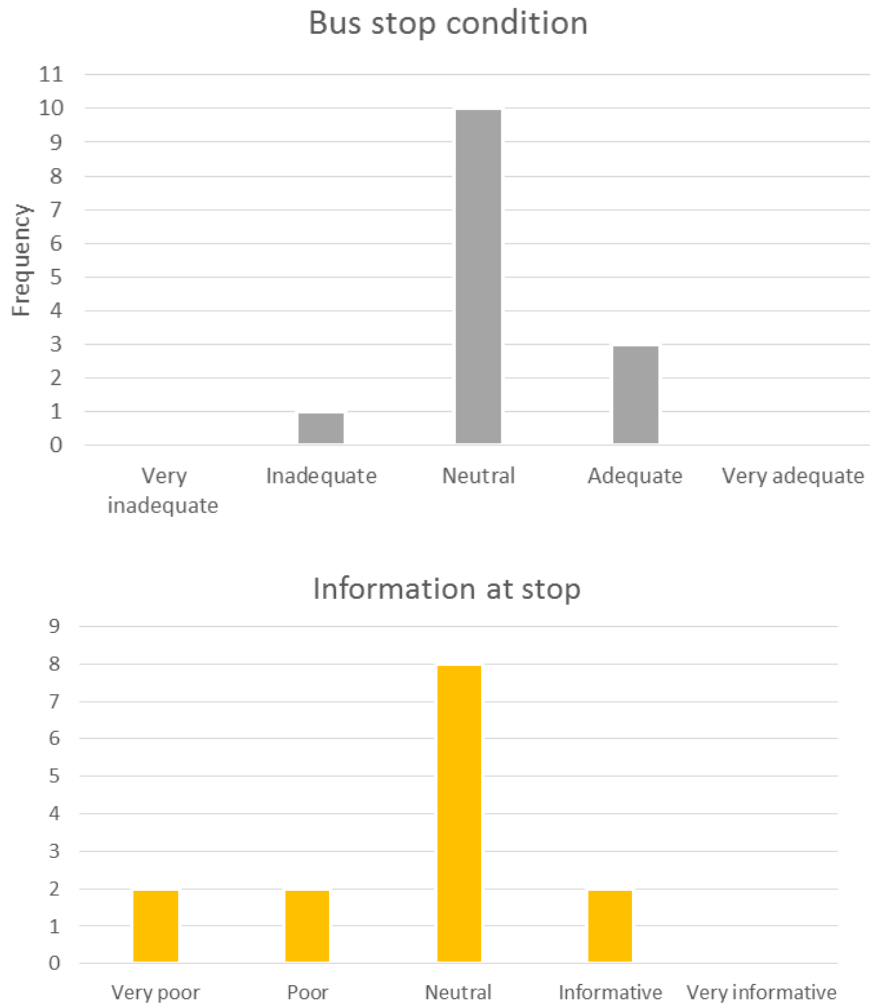
Figure 6.14 - Passenger assessment of bus stops. Top: bus stop conditions; Bottom: bus service information at bus stop

Participants reported largely neutral feelings towards condition of the bus stops and quality of service information available at them (Figure 6.14). This hinted at the possibilities that 1) the particular bus stops invoked neither positive or negative reactions from waiting passengers; 2) passengers lacked a clear personal rubric for evaluating bus stop conditions, distinguishing between neutral and unsatisfactory/satisfactory level of service; 3) people may not pay much attention to the bus schedules posted at stops given the access to real-time arrival information. The few "very poor" and "poor" ratings were

all regarding two of the five participating stops.

As earlier noted, these results were collected from a small, unrepresentative participant pool, aiming only to test and demonstrate capabilities of the FMS-TQS platform. However, if we were to have a larger sample size, the data can lend meaningful insights on user experience for public transport operators and regulators, including allowing for the examination of temporal and geographic variations in service quality and customer satisfaction.

FMS' possibilities to allow for longitudinal surveys at very low marginal cost (essentially the burden to the user to continue running the app and responding to questionnaires) provides a chance to observe customers' actual bus experience over time. For illustration, user #465 reported two very different bus experiences on the same commute on two consecutive days (Table 6.3). On June 22, the user reported satisfactory wait and onboard experiences, but on the next day s/he waited longer for the bus. According to the entrance questionnaire, this user would usually check real-time bus arrival information before the trip, which suggested that the bus might have encountered an abnormal delay. The user's responses also revealed onboard crowding and a somewhat rough ride. The differentiating details between these two bus trips would have almost certainly been lost in more traditional customer satisfaction surveys.

Table 6.3 - Onboard Questionnaire Response Summary

User ID	Date & time	Wait satisfaction*	Why less than satisfied with wait? (if wait rating < 4)	Bus crowdedness**	Comfort satisfaction*	Why less than satisfied with comfort? (if comfort rating < 4)	Driver service satisfaction*
442	6/24/2015 12:42	4	N/A	2	3	Ride is not smooth	3
448	6/26/2015 8:26	3	Other	3	4	N/A	4
453	6/22/2015 9:00	3	Wait time too long	2	3		"Don't know"
453	7/2/2015 9:00	3	Rather not say	4	3	Rather not say	"Don't' know"
453	7/8/2015 8:50	3	Rather not say	4	3	Cannot get a seat on the bus	"Don't know"
456	6/24/2015 7:37	3		2	3	Temperature is too hot/cold	3
456	6/30/2015 17:02	3	Bus too crowded to board	2	3	Too many people on the bus	3
456	7/2/2015 7:52	3		2	4		3
456	7/7/2015 7:54		4	2	4		3
458	6/19/2015 17:42	5	N/A	5	2	Too many people on the bus	5
458	6/24/2015 12:32	3	Wait time too long	4	2	Too many people on the bus	4
465	6/22/2015 18:14	4	N/A	3	4	N/A	"Don't know"
465	6/23/2015 18:15	3	Wait time too long	4	2	Ride is not smooth; Too many people on the bus	"Don't know"
473	6/26/2015 7:48	4	N/A	3	3	Cannot get a seat on the bus	4

Rate on scale of 1-5: 1 = very dissatisfied, 5 = very satisfied

*** Bus crowdedness: user shown a pictorial scale of 5 levels of crowding, with 1 = very empty, 5 = very crowded.*

For public transport operators and regulators, the most useful benefit of this smartphone-based system may be trip-specific feedback. The FMS-TQ system records the phone's geographical coordinates when questionnaires are solicited and returned, which, when combined with the automatically- collected GPS traces, enable us to infer the bus route taken. For example (Figure 6.15), at 18:14 on June 22, User #465 received an onboard questionnaire at location (1.3104978, 103.8477377). This matched the in-vehicle trip sensed by FMS via phone sensors between 18:12 (originating at 1.310086, 103.848) and 18:22 (terminating at 1.32541, 103.8419). We can thus infer that the user was traveling on one of the five bus routes serving that corridor at that time of day (lines 56, 57, 166, 851 and 980). If given access to automatic vehicle location (AVL) data, we could also know the exact bus trip by matching the user's coordinates in FMS with the AVL data. Alternatively, if we had asked for participants' smartcard number and obtained card's travel data, we could easily associate the response and smartcard travel timestamps. An even simpler approach, though at the cost of requiring additional user input, would be to include a question about the bus number in the onboard and alighting questionnaires. Had we the permission to install beacons on buses, the onboard beacon signals could also have helped identify the exact route and run that the user was on.
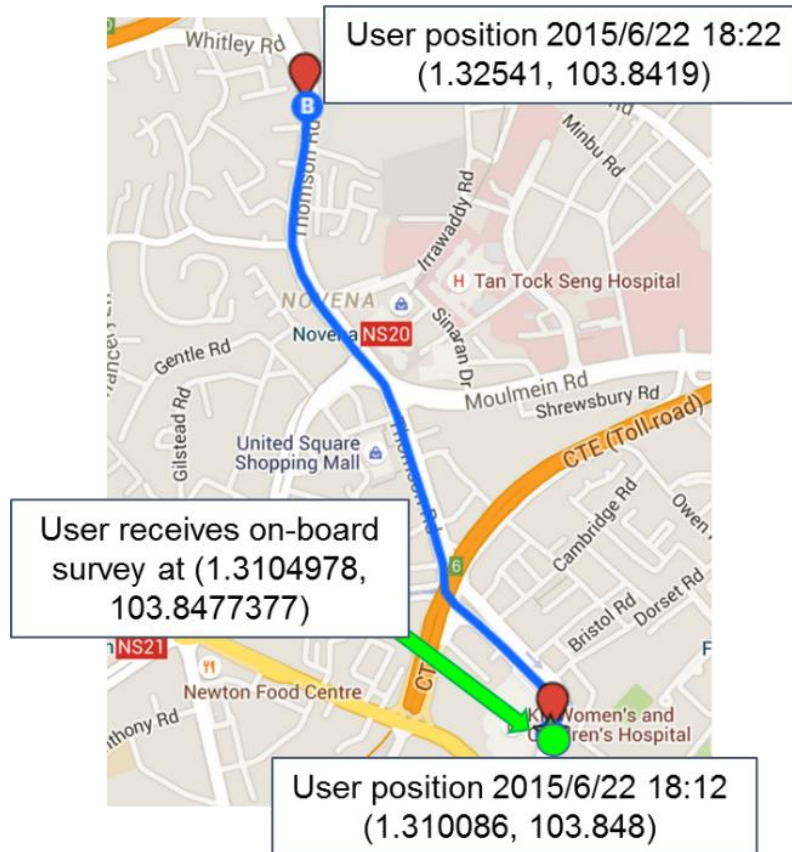
Figure 6.15 - Example of inferring bus trip from GPS coordinates collected by FMS-TQ

Overall, although the number of participants and completed questionnaires were far from sufficient to paint a complete picture of the local bus service quality, the data collected proved the feasibility of our concept – a mobile platform to automatically detect user's bus trips and solicit trip-specific customer feedback. Compared to traditional customer surveys, FMS-TQ can bring additional benefits of more real-time and ubiquitous data collection, allowing for more granular insights into individual bus experiences.

The exit survey revealed that participants largely embraced the app, though some aspects could use improvement. In terms of their overall experience, only one participant felt it was "very negative", while the rest said it was either "neutral", "positive", or "very positive." Most of the users were satisfied with the length of the questionnaires – accentuating the convenience of app-based surveys and our effort to minimize user burden – but suggested that the questionnaires were not always easy to understand. As we were unable to conduct detailed follow-up interviews of the participants, we

postulated that this lack of total clarity stemmed from the two happiness-related questions. As earlier mentioned, our LTA counterparts were interested in understanding public transport service's contribution to one's daily happiness and life satisfaction. We thus included two questions, drawn from the relevant literature (see Appendix A, Questions D7 and F4), to test their validity. According to the exit survey, these questions ultimately were confusing to many respondents. A few participants also voiced the desire to have more flexibility in providing feedback, such as textboxes to provide free-form response. Lastly, respondents were about evenly divided in terms of concerns about the app's effect on battery drainage and phone use, though most participants, even those who reported of some impact, said it was not too hindering.

Due to the pilot's small sample size, we were unable to answer some of the research questions that had motivated this endeavor. Recall that one objective was to discern the potential correlation or difference between real-time and retrospective evaluation of the bus service. In the end, we did not receive sufficient number of completed real-time and retrospective questionnaires that could be matched, thus we were unable to conduct the analysis. Another lessons learned was the weakness of asking about user experience in three separate, mutually-exclusive questionnaires (at-stop, onboard, and post-alight). While this design aimed to keep the questionnaires short and minimize user burden, it precluded us from obtaining a complete picture of one's experience on that particular trip, especially in a small trial. As a result, we also forwent trying to correlate objective operational data – wait time, travel time, etc. -- with riders' reported satisfaction with service.

6.4   Discussion

FMS-TQ went from conceptualization to deployment in less than five months. While this speed is impressive, it also meant that the platform piloted in Singapore was immature in some areas. Some were design imperfections. For example, intending to minimize user burden, we limited the algorithm to select only one bus trip to survey per day -- the first bus trip detected of the day. For most participants, however, this equaled to their morning commute to work trip, which was neither comprehensive nor representative of the trip

experiences that we intended to capture. A better sampling method, as we later adopted for the Boston pilot, would be to randomly select the nth trip of each day. Other problems stemmed from coding errors and testing oversights, which compromised the beacon detection accuracy and analytical capabilities. For instance, only after the pilot started did we realize that iPhone users needed to manually enable Bluetooth on their phones. Though it was corrected by a follow-up reminder, we had lost several valuable days to collect data from this group of users. Also, while we originally intended to derive users' wait time at the stop (from timestamps of the phone's entry and exit from the stop beacon's signal range), we overlooked this part in the engineering process. Unfortunately, the app was not programmed to capture and transmit the timestamp information to the server, so we were ultimately unable to calculate wait time. Lastly, institutional realities posed constraints that limited us from fulfilling some of our research intentions. One major limitation was our inability to install beacons on buses, which hindered the trip detection accuracy. As we were unable to obtain bus operational data, we could not match the collected responses to the exact bus routes and runs. These experiences allowed us to subsequently improve the FMS-TQ approach in the subsequent Boston pilot.

Before closing this chapter to discuss the Boston pilot, I want to acknowledge the importance of marketing, incentives, and institutional design. Given the internal nature of this proof-of-concept pilot, we did not carry out an extensive marketing campaign nor thoroughly considered incentives design. But for any public initiatives, a more targeted marketing campaign would be needed to recruit a bigger participant pool, and a more effective incentives program (such as with reward points or fare discounts) needed to sustain participation. These elements are considered in fuller extent in the Boston pilot, which I will discuss in detail in the next chapter.

Chapter 7

FMS-TQ in Boston

The Boston implementation builds on the foundation established by the Singapore proof-of-concept pilot. In collaboration with the MBTA, we conduct the public pilot the FMS-TQ on all four branches of Boston's Silver Line, the local Bus Rapid Transit system. The particular institutional arrangements have enabled an FMS adaptation more effective for bus trip detection and data inference than the Singapore pilot. The small number of responses from the first three weeks of the pilot suggest that customers report higher levels of satisfaction with the service than their recalled, overall impression. While I have been able to match some questionnaire responses to specific trips and infer further trip attributes, the reliability still has room to improve. The sample thus far is too small to statistically model the relationships between reported satisfaction and trip attributes; I propose a model structure for future data analysis.

7.1   Trial design

Being able to install beacons both on vehicles and at stops on an entire transit service line provides much benefit for the trial design. Signals from onboard beacons allow us to capture bus rides more accurately than the previous inference from stop beacon signals. The unique ID of each beacon further allows us to match the survey response to a particular bus. The app records the information of when the user is first detected at the stop, providing the possibility to calculate the passenger's wait time, by subtracting the at-stop timestamp from the boarding timestamp.

The Boston implementation of FMS-TQ comprises of a set of four surveying instruments: an entrance questionnaire, a short end-of-trip questionnaire, a more detailed onboard questionnaire solicited via a sampling protocol, and a brief questionnaire at the end of day for those who have been detected at a Silver Line stop but never boarded a bus.

Different from the Singapore pilot, which focused mainly on capturing a passenger's feelings in as real-time as possible, the Boston trial emphasizes understanding the customer's entire trip experience. Hence, the system forgoes the previous stop/ onboard/ alight questionnaire differentiation, and instead solicits information on the user's satisfaction with the trip at the end of every Silver Line ride. But recognizing the potential to taking advantage of people's idle time while travelling on the bus, we designed a slightly longer onboard questionnaire, which pops up on the app while the user is riding a Silver Line bus.

Figure 7.1 depicts the logic by which the app interacts with beacon signals and generates questionnaires. To mitigate user burden, the survey strategy will deliver at most one onboard questionnaire a day, even if the user takes multiple trips on the Silver Line. Fare card data show that 60% of regular Silver Line riders use the service once a day, 30% twice a day, 5% three times a day, and 5% four or more times. In other words, the vast majority of users take two or fewer trips on the Silver Line on a typical day. We used this distribution to guide the real-time sampling algorithm. When the app detects the first Silver Line trip for the user that day, it has a 50% chance of generating the onboard questionnaire for that trip. If the user does not receive a questionnaire on that first trip, then the app has a 90% likelihood of generating an onboard questionnaire upon detecting a second Silver Line trip on the same day by that user. If the user has been detected to take his third Silver Line trip of the day and has not yet received an onboard questionnaire on that day, the app has a 100% chance of sending him the questionnaire. This design improves upon the scheme used in Singapore, which resulted in a heavy over-sampling of users' first bus trip detected each day.

1st time app login

Entrance questionnaire

Each subsequent Silver Line trip

Stop beacon entry event

Bus beacon entry event

Stop beacon exit event

no bus beacon detected

End-of-day questionnaire scheduled for 8pm

Until 8 pm

no other onboard or end-of-trip questionnaire completed

End-of-day questionnaire

an onboard or end-of-trip questionnaire has been completed

End-of-day questionnaire cancelled

App generates random number (0 ≤ p ≤ 1)

phone in the same bus beacon range for >90 sec, with no persistent interfering signals

End-of-trip questionnaire scheduled

Bus beacon exit event

Next 4 minutes

app detects no other bus beacon signals for >90 sec

End-of-trip questionnaire

app detects another bus beacon signals for >90 sec

App starts monitoring new bus beacon

trip chosen for onboard survey

phone in the same bus beacon range for >90 sec, but persistent interfering signals detected

App monitors both beacons, until signal of one drops out

Onboard questionnaire (60 seconds later)

phone in the same bus beacon range for <90 sec

Disregarded as noise

Figure 7.1 - FMS-TQ app logic and survey generation flow

118

Below summarizes the content and app logic involved in generating each of the four questionnaires; a copy of the actual questionnaires is in Appendix B.

Entrance questionnaire: After installing the app on her phone, the user receives an intake survey to collect baseline information, including her mobility options, general satisfaction with transit, and demographic information.
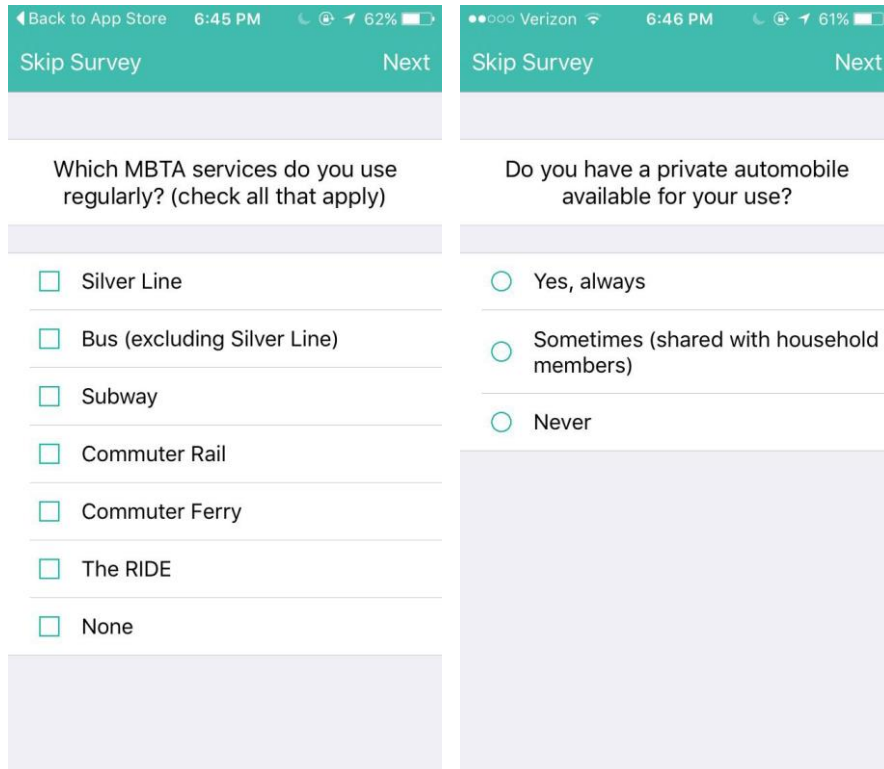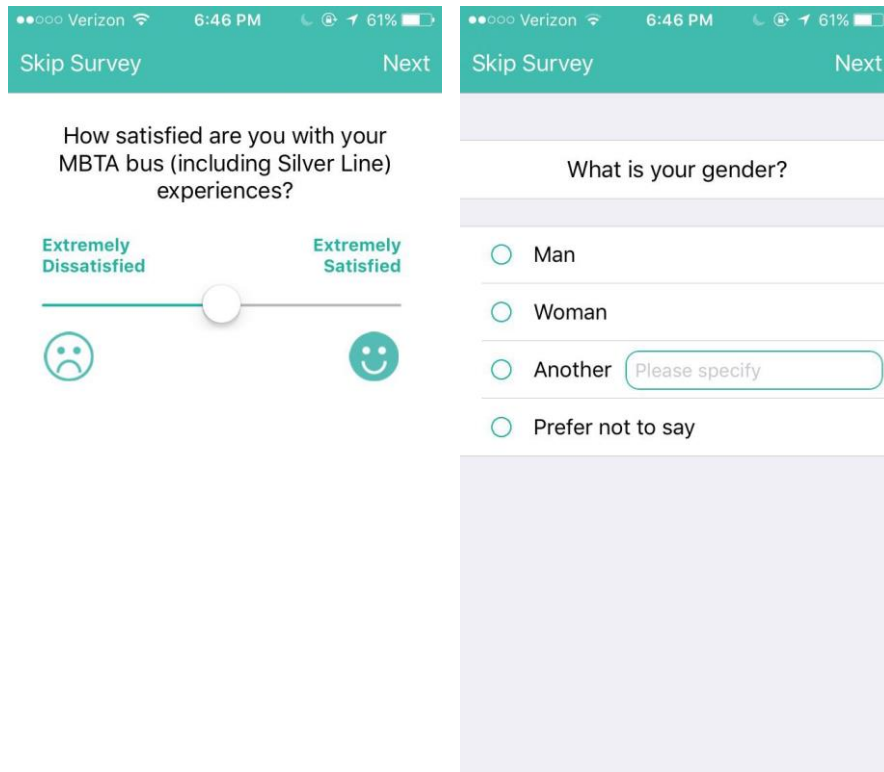
Figure 7.2 - Select screenshots of entrance questionnaire

End-of-trip questionnaire: When a user enters a bus beacon's signal range, followed by exiting from the same range after at least 90 seconds, the app assumes that the user has just completed a Silver Line trip. At this time, the app prompts the user for a quick rating on the trip experience. On a single screen, the user is asked to rate the service overall and four attributes: wait experience, travel time, comfort and cleanliness. Also on the screen is a text box for optional comments, as well as an option to report false positives – that the user did not actually ride on the Silver Line.

Figure 7.3 - End-of-trip questionnaire

Onboard questionnaire: On select occasions, as previously outlined, a user would receive an onboard questionnaire which first verifies that the user is indeed travelling on the bus. If so, it asks about 1) purpose of the trip, 2) customer's satisfaction with each of five service attributes – cleanliness of the stop, wait experience, onboard crowdedness, onboard comfort, and driver's service, and 3) reasons for dissatisfaction, if any.
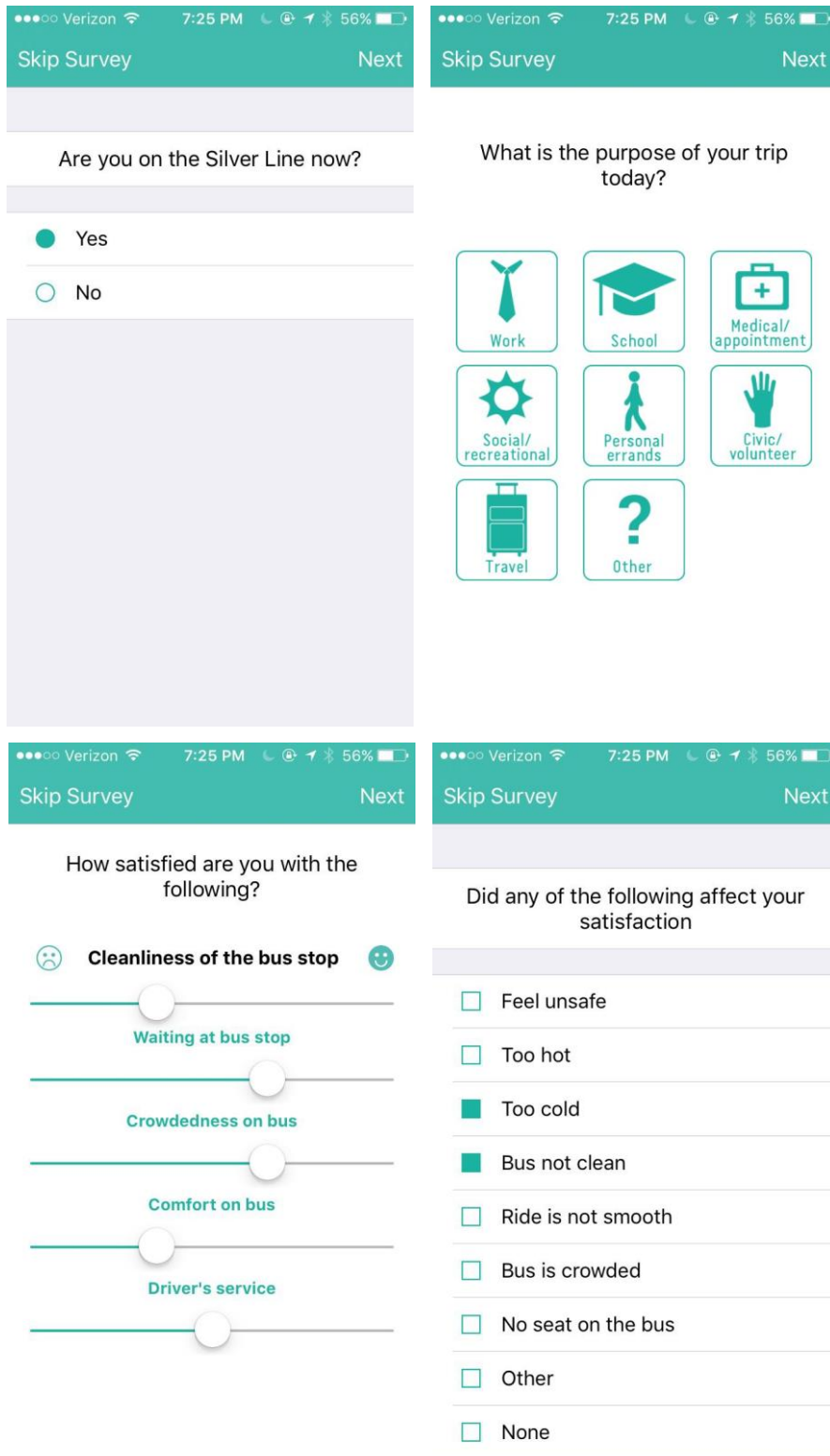
Figure 7.4 - Screenshots from onboard questionnaire

End-of-day questionnaire: This survey is solely solicited from users who have

been detected at a Silver Line stop but never boarded a vehicle. We incorporated this "safety net" to capture the experience of customers who might have waited at a stop and left without boarding a bus, as well as those who have actually ridden the bus but whose bus trip was missed by the detection algorithm for some reason. To generate this survey, the platform combs the backend data at 9pm each day, and pushes a questionnaire to the identified users. To minimize user burden, as long as one has received completed an end-of-day and/or onboard questionnaire that day, she would not receive the end-of-day questionnaire.

The beacon events, combined with subsequent FMS-based sensing, allow us to post-estimate users' wait times at the bus stop. The GPS data collected by the smartphone reveal the bus's travel speed and time, and the accelerometer data can lend insights into smoothness of the ride.

To remind users to enable Bluetooth on their phone, which is necessary for detecting beacon signals, the Boston-version of FMS-TQ has a reminder feature. At 8 am every day, the app would check whether the phone's Bluetooth is turned on; if not, the app would send an on-screen notification to the user, reminding him to "turn on Bluetooth to rate [his] Silver Line trip!" This reminder feature is an enhancement from the Singapore pilot, for which either the app forcefully turned on the Bluetooth on Android phones, and we relied on users with iOS phones to remember to enable Bluetooth.

7.2   Implementation

The implementation involved a number of preparatory steps, straddling hardware (beacon) installation, software development, and survey and instrument design, user experience design, and marketing and outreach strategies. Although it was the second FMS-TQ pilot, the Boston pilot entailed much work from scratch. The system configuration is entirely different from the Singapore pilot, necessitating much effort spent on devising and testing the system. Furthermore, the Boston pilot is public-facing, prompting a different approach to marketing and recruitment from the Singapore pilot with LTA officials.

Starting with the hardware component, we installed about 100 beacons on all Silver Line vehicles and stops. We decided to use Estimote iBeacons again, considering their durability to withstand outdoor climates and our engineers' experience in working with this brand in the Singapore pilot. We chose three colors – white, light blue, and dark blue – intended to help camouflage the beacons in the various environment of the different stops and stations. These limited color options offered by Estimote meant that we could not get a very close match for some of the installation sites. This imperfection was however mitigated by installing the stop beacons at a high position, out of general line of sight of the public. The bus beacons were installed behind a fiberglass panel and entirely hidden from passengers.

The beacons were first catalogued and configured. One of the first tasks was to enable the app to distinguish bus from stop beacon signals. Since every beacon had a programmable major and minor ID, we set all bus beacons to one identical major ID, and all stop beacons to another. Configuring the beacons to the right major value was hence crucial.

Figure 7.5 - Estimote beacons used in the pilot
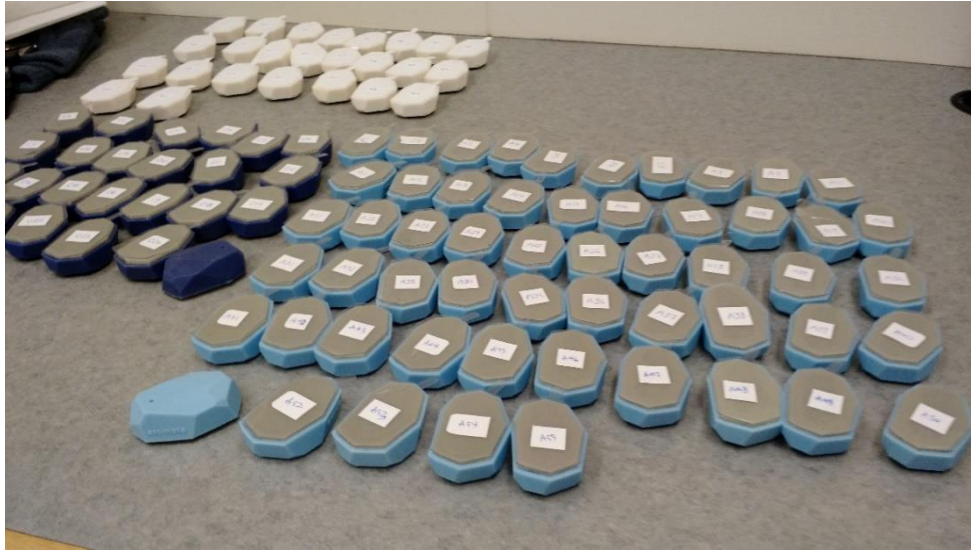
Figure 7.6 - Labeling and configuring the beacons

With MBTA's assistance, we were able to install the majority of beacons by mid-December. For buses, we found a spot inside one of the advertising panels near the ceiling of the bus. The fiberglass panel hid the beacon from view but did not totally block transmission of the beacon signal (as would a metal panel). It was also conveniently located at the center of the bus and in a high position – ideal for beacon signal transmission. As the beacons came with strong adhesive on the back, we were able to simply stick them onto the backside of the panel. We noted the identification number of each beacon and the corresponding bus number. We made several trips to the Southampton Garage in South Boston, where Silver Line vehicles are kept, to install on all the vehicles.

Figure 7.7 - Installing beacons on the buses

For vast majority of the 40 Silver Line stops and stations, we were able to find a discrete and effective spot to install the beacon. In the following months, we made periodic inspections to ensure that all beacons were still there (we lost one due to a stop renovation) and remained functional.
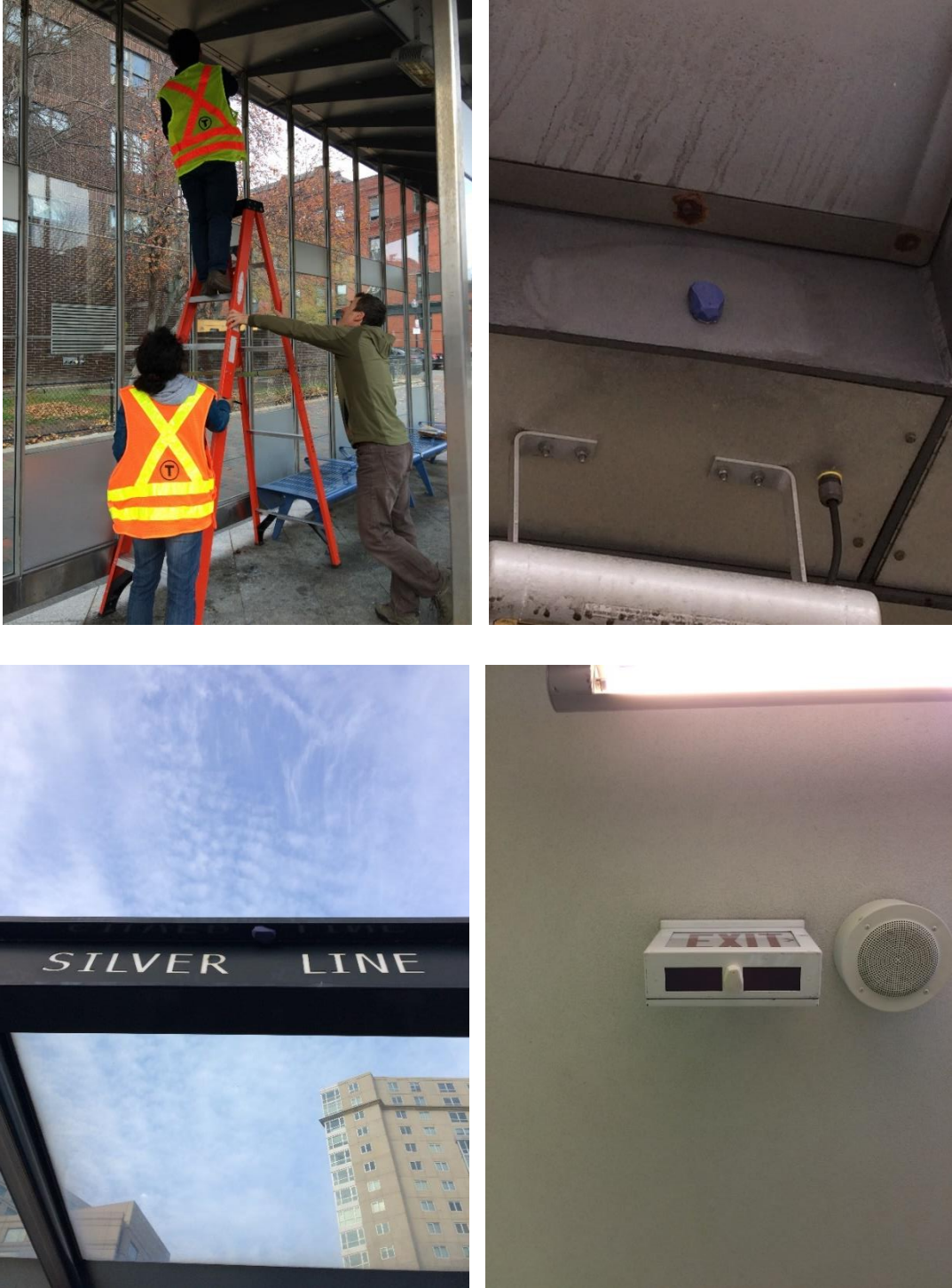
Figure 7.8 – Beacons installed on Silver Line stops and stations

The three-month pilot began on April 15, 2016. Prospective participants were directed to the initiative's website – **http://QualiT.mit.edu** – where they could learn more,

sign up for an account, and proceed to install the FM Sensing app from Google Play (Android) or AppStore (iPhone). Once logged into the app on their phone, participants receive the entrance questionnaire and would otherwise be surveyed automatically following the survey logic described above (Figure 7.1).

      We marketed the project through a wide range of channels, including the MBTA and MIT media, outreach to transportation management associations, employers and prominent organizations in areas served by the Silver Line. The messages appealed to people's intrinsic motivations to give feedback to help the MBTA improve the service, as well as to the fun and innovative nature of the act ("you can now rate your bus experience just like you can with Uber or Lyft rides!"). The outreach messages also highlighted the economic incentive: for every questionnaire completed, the user would be automatically entered into a monthly sweepstake to win one of three available MBTA monthly passes.

7.3   Findings

Though the pilot is scheduled to run through July 15, 2016, I cut off the data collection for this thesis on May 8, 2016. As of that date, the recruitment website had registered more than 900 unique visits; 60 people have signed up, 44 of which have logged into the app and 30 completed at least one survey. A look into the details by user shows that many of the inactive sign-ups joined early in the pilot and are MIT affiliates, reflecting the fact that most of our publicity effort in the first 10 days of the pilot was restricted to within MIT circles. None of the Silver Line branches serves the MIT campus, lowering the likelihood that MIT affiliates would be regular users, thus possibly explaining the low response rate among early adopters. In addition to the 30 entrance surveys, we have received 48 trip-based questionnaire responses: 25 end-of-trip, 18 onboard, and 5 end-of-day (Table 7.1). Recall that the end-of-day survey is only generated when participants are detected at a stop but never received an onboard of end-of-trip questionnaire on that day. The small number of end-of-trip surveys generated is thus a desirable outcome, meaning that most users who have waited at the bus stop are subsequently detected on a bus.

Table 7.1 - Number of respondents and Responses

Questionnaire	Number of responses
Entrance	30
End-of-trip	25
Onboard	18
End-of-day	5
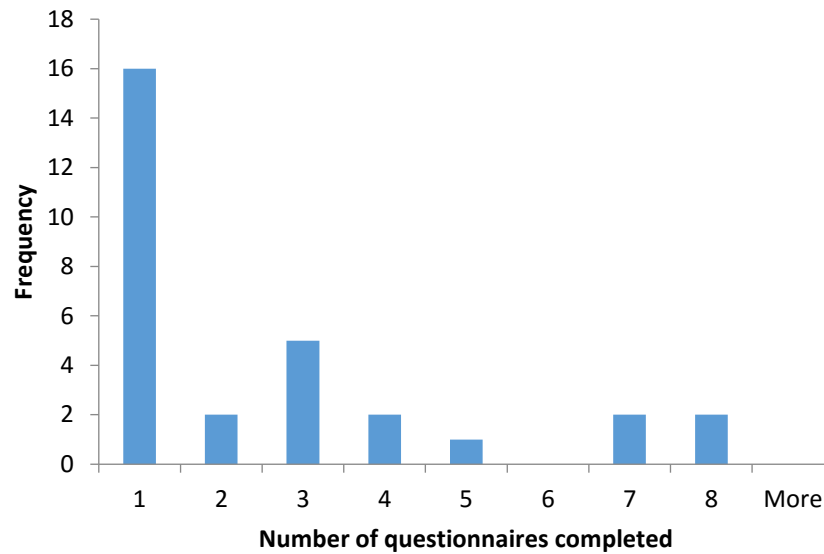
Figure 7.9 - Distribution of number of questionnaires completed per user

Among the 30 users who have completed the entrance questionnaire, four were disqualified from further participation as they did not report to be regular Silver Line users. The remaining 26 users can be grouped into three relatively even shares in frequency of Silver Line usage: 35% less than once a week, 27% one or two times a week, and 38% more than twice a week. The extent of available mobility alternatives among respondents varies: 54% of respondents report not having a private automobile available to use, while the remainder splits evenly between sharing with other members of the household and always having a car. Slightly over half of the respondents do not have a bike or bike-share program membership.

Demographically, the participants are significantly biased towards being male, younger, and from higher income households. The age and income bias align with

expectations, as the younger population are in general more smartphone- and app-savvy, and higher incomes correlate with smartphone and data plan ownership (Pew Research Center, 2015). However the significant gender imbalance is perplexing and unsupported by prior literature in the U.S., which reveals essentially no difference in smartphone ownership between men and women (Anderson, 2015), and that men are less likely to response to web- and phone-based surveys (Holbrook, Krosnick, & Pfent, 2008; Smith, 2008)

The entrance survey also reveal the respondents' relatively negative ratings of their overall experience with MBTA's bus services (including the Silver Line). The question presents a spectrum of satisfaction levels (1 = very dissatisfied, 7 = very satisfied), with a toggle initialized at 4, the "neutral" position; respondents can move the toggle to select their answer. Sixty-two percent of respondents rate it as worse than neutral, including 16% choosing a score of 1 or 2. Only 15% of respondents give a rating of 5, and none reports a score of 6 or 7. This result contrasts drastically with those from the latest MBTA customer satisfaction survey, in which only 24% of respondents report some level of dissatisfaction and 66% report some level of satisfaction (MBTA, 2016). This hints at yet another potential source of bias in our pilot – people unsatisfied with the bus service may be more likely to sign-up and install an app intended to rate it.

Table 7.2 - Summary statistics of entrance questionnaires

Weekly Silver Line usage		Gender	
Less than once	35%	Woman	27%
One or two times	27%	Man	65%
Three or more times	38%	Prefer not to say	8%

Access to private car		Age	
Yes, always	23%	< 18	0%
Sometimes	23%	18 to 21	4%
Never	54%	22 to 34	58%
		35 to 44	8%
Own bike or bike-share membership		45 to 64	27%
Yes	46%	65 +	0%
No	54%	Prefer not to say	4%

Satisfaction with MBTA bus experience		Race	
Very satisfied	0%	American Indian/Alaska Native	0%
Satisfied	0%	Asian	8%
Somewhat satisfied	15%	Black of African American	0%
Neutral	23%	Native Hawaiian/Pacific Island	0%
Somewhat dissatisfied	46%	White	73%
Dissatisfied	8%	Other	4%
Very dissatisfied	8%	Prefer not to say	15%

The trip-based satisfaction ratings are much higher than participants' general rating of satisfaction with MBTA's bus services, as reported in the entrance surveys. Figure 7.10 and Figure 7.11 show the distribution of the same ratings. At a glance, ratings for stop cleanliness and driver's service are positively-skewed, while reported satisfaction for the wait, onboard crowding and onboard comfort tend towards neutral (4) and fall off on either side. Examining all responses for every single user shows the pattern holds true for all except one user–reported satisfaction in the onboard and/or end-of-trip surveys is at least as high, and for many responses much higher, than that expressed in the entrance surveys. This comparison suggests that people may have a general dissatisfaction towards the local bus services, perhaps colored by an occasional particularly poor experience, while specific satisfaction with most bus trips may actually be just fine. Perhaps the MBTA's reputation is worse than its actual service.
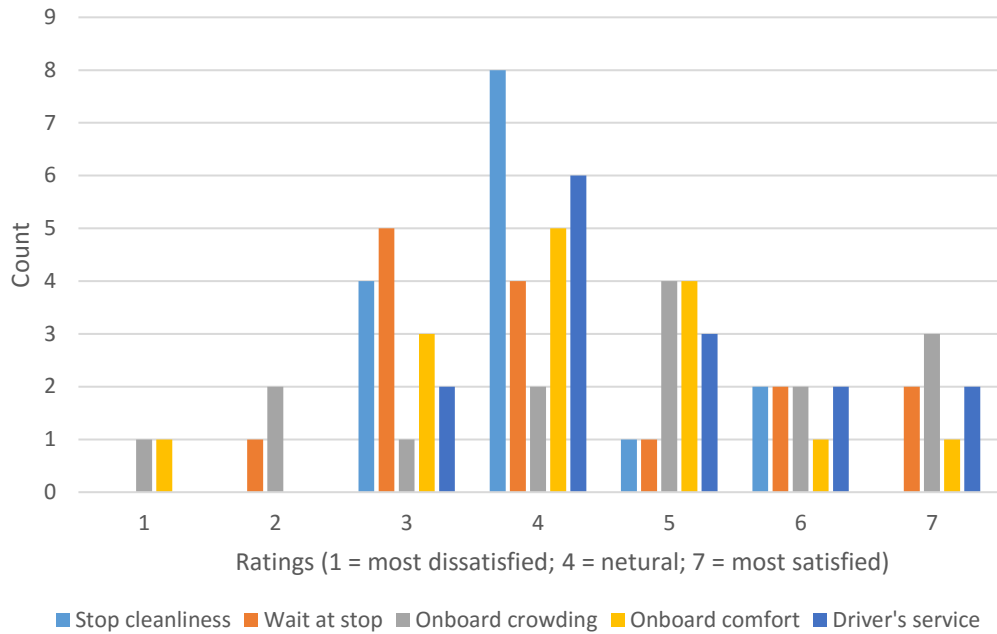
Figure 7.10 - Satisfaction with service, as reported in onboard questionnaire responses
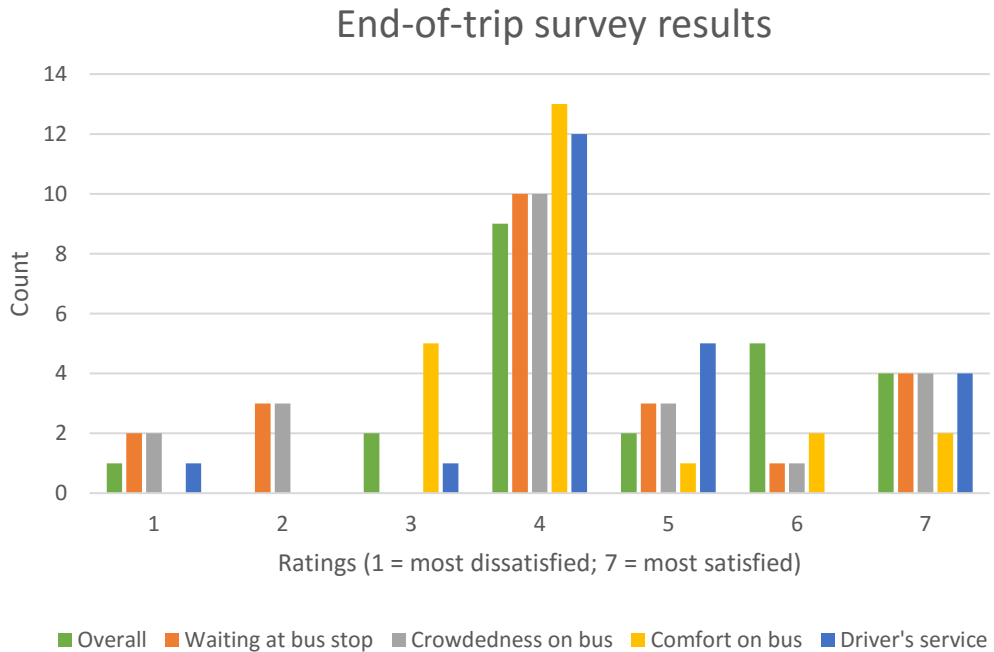
Figure 7.11 - Satisfaction with service, as reported in end-of-trip questionnaire responses

As satisfaction with wait time, crowding, and onboard comfort tends to be affected

by time of the day, I investigate these potential relationships by constructing two categorical time period variables, to capture whether the trip took place during peak (7 am – 9:30 am, 4 pm – 7 pm)[11] or off-peak periods. FMS-TQ records the time at which each questionnaire is pushed to the user, making this categorization very easy. Given the small sample size, I use Fisher's exact test to examine the proportions of satisfaction levels (less satisfied than neutral (<4), neutral (4), or more satisfied than neutral (4)) by time periods. The test results, shown in Table 7.3, do not reveal any statistically significant differences between peak and off-peak for any service attribute. Keep in mind that the dataset is a biased sample, and the reported satisfactions of the bus experience likely also depend on vehicle headways and weather conditions (affecting pleasantness of the wait), rather than time period alone.

I also hypothesize that the ratings on trips on the Waterfront branches (SL1, SL2, and Silver Line Way shuttle) would be higher than those along the Washington Street corridor (SL4, SL5). I based this conjecture on the fact that the Waterfront branches run in a dedicated tunnel for a considerable part of the routes, and, judging from personal experience, the station/stop facilities and vehicles are cleaner than the SL4 & SL5 counterparts. The results from the Fisher's exact test, shown in Table 7.4, however, do not portray any significant differences among the two Silver Line corridors.

---

[11] These morning and evening peak periods end 30 minutes later than the time periods defined in the MBTA Service Delivery Policy (7-9am, 4-6:30pm). As some FMS-TQ questionnaires are generated after the end of the bus trip, I increase the time period by 30 minutes to capture trips that may have started during the official peak periods but finish after.

Table 7.3 - Satisfaction ratings on service, by time period, and Fisher's exact test results

Onboard questionnaires

Stop cleanliness			Wait			Crowding		
Rating	Peak	Off-peak	Rating	Peak	Off-peak	Rating	Peak	Off-peak
<4	2	2	<4	4	2	<4	2	2
4	6	2	4	3	1	4	2	0
>7	2	1	>7	3	2	>7	6	3
p value	0.78		p value	1		p value	0.78	

Comfort			Driver		
Rating	Peak	Off-peak	Rating	Peak	Off-peak
<4	2	2	<4	2	0
4	3	2	4	5	1
>7	5	1	>7	3	4
p value	0.52		p value	0.24	

End-of-trip questionnaires

Overall			Wait			Onboard crowding		
Rating	Peak	Off-peak	Rating	Peak	Off-peak	Rating	Peak	Off-peak
<4	4	0	<4	2	1	<4	4	1
4	6	4	4	6	3	4	7	3
>7	6	3	>7	8	3	>7	5	3
p value	0.42		p value	1		p value	1	

Onboard comfort			Driver's service		
Rating	Peak	Off-peak	Rating	Peak	Off-peak
<4	4	1	<4	1	1
4	9	4	4	8	4
>7	3	2	>7	7	2
p value	1		p value	0.68	

Table 7.4 - Satisfaction ratings on service, by routes, and Fisher's exact test results

Onboard questionnaires

Stop cleanliness				Wait				Crowding		
Rating	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**
<4	1	0		<4	1	3		<4	2	3
4	2	5		4	1	2		4	1	2
>7	1	1		>7	2	1		>7	1	1
p value	0.67			**p value**	0.74			**p value**	1	

Comfort				Driver		
Rating	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**
<4	2	1		<4	0	1
4	1	3		4	2	3
>7	1	2		>7	2	2
p value	0.74			**p value**	1	

End of trip questionnaires

Overall				Wait				Onboard crowding		
Rating	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**
<4	3	0		<4	2	1		<4	3	1
4	5	4		4	2	3		4	5	3
>7	3	3		>7	7	3		>7	3	3
p value	0.54			**p value**	0.68			**p value**	0.84	

Onboard comfort				Driver's service		
Rating	**SL 1/2, SLW**	**SL 4/5**		**Rating**	**SL 1/2, SLW**	**SL 4/5**
<4	1	3		<4	0	1
4	8	3		4	6	4
>7	2	1		>7	5	2
p value	0.33			**p value**	0.60	

To capitalize on the high resolution data and use that to better understand trip-specific characteristics that may influence respondents' trip experience and their reported satisfaction, I also attempt to match the survey responses to specific trips. The FMS system records all questionnaire responses received via the app, along with the user IDs and time and location stamps; it also records the phone's entries into and exits from any FMS-TQ beacon range with the corresponding timestamps of these events. The matching exercise thus entails three steps:

1) Reconstructing the user's Silver Line trips from the beacon interaction data collected by FMS-TQ;

2) Associating survey responses to the reconstructed bus trips; and

3) Locating the corresponding bus runs in the MBTA's database.

Step 1 intends to infer the user's origin and destination stops, wait time, travel time, and bus number for the trip, using the beacon data, and the corresponding user ID and timestamps (Figure 7.12). Step #2 enriches survey data with contextual trip information,

by matching the user responses to specific bus trips, by referencing the timestamp and user ID information. Step #3, performed by SQL with access to MBTA's official trip records, would return the official bus trip ID, route number, direction, and operator ID (Table 7.5). Figure 7.13 provides a schematic illustration of the components and logic of the inferences undertaken.

Beacon data table

row_id	user_id	entry_time	exit_time	major	minor	label
6968	60	4/18/2016 20:34	4/18/2016 20:35	14815	42474	Bus 1030
6969	55	4/19/2016 21:17	4/19/2016 21:18	12025	47255	Logan Terminal C
6970	55	4/19/2016 21:19	4/19/2016 21:20	12025	22272	Logan Terminal E
6971	55	4/19/2016 21:28	4/19/2016 21:30	12025	41318	Silver Line Way
6972	55	4/19/2016 21:31	4/19/2016 21:32	12025	60794	World Trade Center
6973	55	4/19/2016 21:33	4/19/2016 21:34	12025	15857	Courthouse
6974	55	4/19/2016 21:15	4/19/2016 21:36	14815	33420	Bus 1122
6975	55	4/19/2016 21:36	4/19/2016 21:36	12025	62940	South Station SL1
6976	69	4/21/2016 17:20	4/21/2016 17:23	12025	42990	Tide St
6977	69	4/21/2016 17:23	4/21/2016 17:34	14815	23792	Bus 1115

A different trip (different user_id & date/time)

One trip (Origin: Logan Terminal C; Destination: South Station; Bus #: 1122; Travel time: 19 min)

A different trip (different user_id & date/time)

Figure 7.12 - Illustrative example of trip inference (Step #1)

Table 7.5 - Illustrative example of result of data fusing and inference

Survey response collected by FMS-TQ							
user_id	survey_id	overall	waiting	crowding	comfort	service	question_t
55	7	4	3	4	4	4	4/19/2016 21:46

Trip info inferred from FMS-TQ data						
origin_stop	origin_arrive_t	board_bus_t	wait_time	destination stop	arrive_dest_t	bus_num
ogan Terminal	4/19/2016 21:17	4/19/2016 21:15	N/A	South Station	4/19/2016 21:36	1122

Trip info from MBTA's database			
trippiece_id	operator_id	route_num	direction_id
50579240	15510	SL1	Outbound
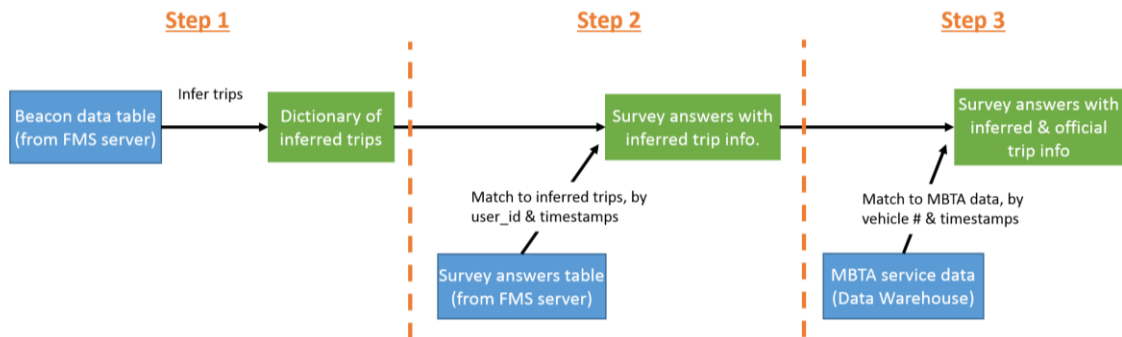
Figure 7.13 - Schematic illustration of data analysis process

The above information signifies a major advance over the Singapore pilot, for which fusing such data from different sources was not possible. Table 7.6 shows the

number and percentage of onboard and end-of-trip questionnaires, respectively, that I have been able to successfully match to trips, as well as infer objective metrics for. The numbers are lower than expected. While testing confirmed that the app easily picks up the beacon signals at boarding and alighting stops, the beacon data collected from the pilot have been less consistent. In instances where FMS-TQ did not register the origin and/or destination stop beacon signals of a trip, I am unable to infer the associated origin, destination, wait time, and/or travel time. The match to the MBTA records has also been less than 100% successful; inspection of the data points to shortcomings in the backend engineering. First, the FMS-TQ's data table did not originally register the ID of the bus beacon associated with each response. This error was detected in early May and corrected on May 13th. Since the correction came after the data cut-off date for this thesis, I still relied on the methods earlier-described to infer the bus beacon ID and vehicle number. This loss of ground truth data has domino effects, as the match to MBTA records relies heavily on having the correct bus vehicle number. While my algorithm has been able to derive the associated vehicle number for most of the responses collected, the accuracy is further compromised by a separate engineering bug. Up until May 4, 2016 (when it was corrected), this bug resulted in the iOS app recording the wrong timestamps in the survey responses table – instead of the time at which the questionnaire was generated and pushed to users, the app recorded the time when users opened the questionnaire on their phone.

Table 7.6 - Number of onboard questionnaire responses with successful inference/match to MBTA data

Onboard									
Total	False positive								
	2								
	Effective		Metrics inferred from FMS-TQ data						Matched to MBTA trip records
17			Bus #	Wait time	Origin stop	Destination stop	Travel time	All metrics	
	15	# of questionnaires	14	11	12	10	14	8	10
		% of effective questionnaires	93%	73%	80%	67%	93%	53%	67%
End-of-trip									
Total	False positive								
	2								
	Effective		Metrics inferred from FMS-TQ data						Matched to MBTA trip records
25			Bus #	Wait time	Origin stop	Destination stop	Travel time	All metrics	
	23	# of questionnaires	21	16	18	13	20	11	18
		% of effective questionnaires	91%	70%	78%	57%	87%	48%	78%

One of the key objectives for the Boston pilot is to answer the question: "how do passengers' reported satisfaction with bus experience correlate with objectively measured factors – such as wait time, travel time, smoothness of the bus ride, and adverse weather conditions?" Given the small number of responses collected, and the even smaller set of responses for which I have been able to infer wait time and travel time, formal statistical modeling is not yet possible. Rather than present model estimations on a very small sample, instead I propose a general model specification, which can hopefully be estimated on the model as the number of observations increases with the number of responses, and as I continue to refine the capability to infer the objective metrics. Figure 7.14 depicts the proposed model, a structure equation model with two latent variables – objective service quality and user-perceived service quality, with the former positively influencing the latter. Objective service quality would be indicated through three variables that we can derive from FMS-TQ (wait time, smoothness of ride, travel time) and two inferable through MBTA's official service records (route and vehicle operator). As service performance is not limited to these six attributes only, the model also recognizes the existence of unobserved factors (D1). Objective service quality influences the quality perceived by the user, which is also impacted by time of day and trip purpose (e.g. whether the user is in a rush to work), weather conditions (e.g. wait in a cold, windy

environment versus on a nice day), and unobserved factors (D2). The perceived quality is hypothetically indicated by the user's reported overall satisfaction with the service, as well as with individual service attributes that we ask about in the questionnaires.
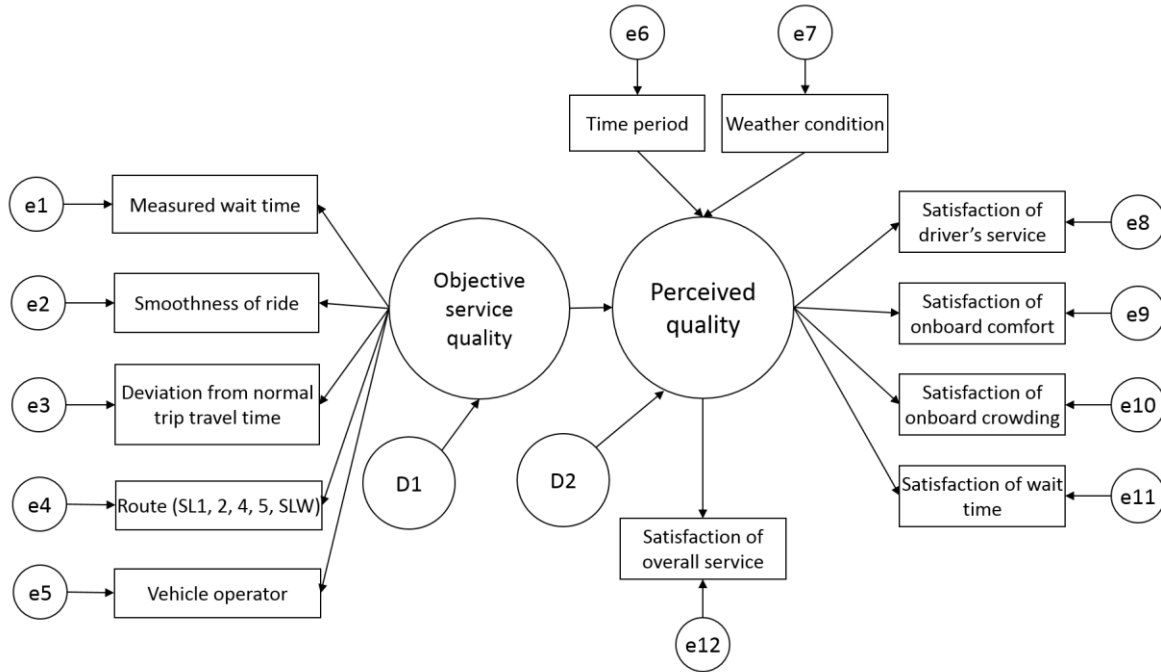
Figure 7.14 - Proposed model for future data analysis

7.4 Discussion

7.4.1 On the Boston pilot

The Boston implementation of FMS-TQ has opened up many possibilities for capturing new insights on users' bus trips. We directly collect passengers' assessment of service quality through the trip-based questionnaires and can also deduce the trip origin, destination, start time, wait time, travel time, vehicle number, route number, and direction for a portion of the trips. Additional, rich detail can be mined from the data. For example, accelerometer data collected by the FMS-TQ can reveal relative smoothness of the bus ride; further analysis of MBTA's data can reveal vehicle headway information, and improve origin-destination inference currently done using only FMS-TQ data; obtaining the weather condition information through a historical weather data API could add that potentially

139

important dimension to users' reported satisfaction. As I am unable to extract and analyze these additional sources in time for this thesis, I look forward to exploring and developing this capability in the weeks to come.

At the same time, the fusion of these various data sources and subsequent inferences is more limited than we had hoped when first conceptualizing the FMS-TQ system. By design, after repeated testing and reconfiguration, the system should be able to consistently detect and record the stop and bus beacon signals; however, the beacon records that we have received so far are rather inconsistent. In some cases, the origin and/or destination stop beacon data are missing, preventing inference of the corresponding wait time and travel times. While I suspect this is in part due to imperfect use – e.g., participants not remembering to turn on their Bluetooth before beginning their trip – our methods and underlying technologies themselves are not perfect. The order in which the beacon data are recorded in the table is not consistent, making it hard to reconstruct the users' bus trips by algorithm alone. For the analysis in this thesis, I performed a follow-up manual check to correct some of the inference errors, but this would not be practical for a much larger set of data.

Despite these shortcomings, the FMS-TQ pilot in Boston represents an important advance. The use of Bluetooth beacons has enabled for the first time the automatic detection of an individual's public transport trip in real-time. Subsequently, this allows for the app to solicit user feedback without manual initiation, and for researchers to derive trip-specific attributes through data collected by the user's smartphone. As a result, the dataset contains both user-reported satisfaction and corresponding objective performance attributes, opening up new possibilities for research inquiries.

7.4.2   A few overall thoughts on Singapore and Boston pilots

The success of a user-feedback platform needs to be judged not solely by its technical capabilities, but equally importantly by its ability to gain traction and generate meaningful data for the service operators. Both the Singapore and Boston pilots have illustrated the challenges of driving participation and adoption. In Singapore, we recruited 26 voluntary users by publicizing through the LTA's email list; in Boston, where we

originally imagined a much bigger participant pool given the public pilot, we have enlisted about the same number of active users during the first three weeks. I think the challenge is three-fold. First, most people may not be compelled to sign up for an initiative and download an app solely to give feedback on their bus rides. Given this lack of enthusiasm or necessity, external incentives will likely be crucial at motivating participation. This brings me to the second challenge: inadequate incentives. Similar smartphone-based studies before ours, such as ones by Carrel et al (2015) and Dunlop et al (2015), have all committed substantial rewards to participants beforehand, in exchange for their commitment to participating. Even with this set-up, participation rates in those studies dropped off drastically once the users satisfied the minimum requirements. As discuss in Chapter 3, fixed cash payments tend to outperform lotteries, prepaid incentives are superior to postpaid rewards, and the ideal remuneration scheme for long-term panel surveys is likely a combination of pre- and post-paid incentives. For both the Singapore and Boston pilots, the rewards are chances to win a gift card or monthly pass through lottery. For limited-duration pilots for a new technology, I believe it's important to be able to collect a large amount of data – even at a higher financial cost – so researchers are better able to test their hypotheses and examine the merits and limits of the technology. Thus, for both the Singapore and Boston pilots, ideally I would have taken the mixed incentive approach: guaranteeing upfront a substantial reward (e.g. $50) for anyone who signs up, asking for their commitment in completing a minimum number of surveys, and offering additional rewards for contributions above the minimum. Lastly, the app can be designed in much leaner ways to reduce buy-in and usage burden. While I am very appreciative of the resources invested by the FMS team for this project, the outcome – from a purely user experience and adoption perspective – might have been better had we delivered the same functionalities as a lighter, stand-alone app rather than an extension of the complicated FMS app. In this alternative design, the app may be branded to better match the pilot, allow people to participate without the compulsory sign-up, or impose additional battery drain on the phone as a result of the location-based tracking features.

For data to be meaningful to public transport agencies, they also need be representative of the general ridership. Although prior to the project I had recognized the likely demographics biases in a smartphone-based pilot – the extent of the actual biases

nonetheless surprises me. The Singapore pilot, being limited to LTA employees, never intended to strive for representativeness. Participants of the Boston pilot, up until the cutoff date for this thesis, are overwhelmingly White, higher-income, iPhone users. After seeing this significant bias, we have recently gone to the Dudley station area for targeted outreach, and plan on conducting more in-person outreach in the near future, especially focusing on driving participation among demographics currently underrepresented in the pilot. Given the inherent barriers for lower-income populations to participate in smartphone-based studies, as discussed in Chapter 2, the representation issue will likely persist throughout this pilot. This is a call for reflection on the practical merits of smartphone app-based surveying, and for more research on ways to reduce representativeness bias in such surveying methods.

Despite the current implementation challenges, I hope that academics and public transport agencies will continue to explore appealing, user-friendly ways to deliver the key principles and functionalities of FMS-TQ. When smartphones even more widespread and integrated into people's daily lives, and when much of public transport trip planning and payment migrate to mobile channels, soliciting riders for occasional feedback on their trip experience may become more feasible and meaningful.

Chapter 8

A Planner's Reflection

In the words of one MIT professor: "Research is like a random walk. There can be much uncertainty and it may take a long time to get somewhere." As a scholar and a prospective urban planner, I will reflect in this chapter on the winding journey of the FMS-TQ project, synthesizing what it takes to implement innovative projects with public agencies on two continents. It distills down to five lessons about the importance of institutionality, consensus-building, iterative improvements, adapting to diverse partners, and grounding research in practice.

Lesson #1: Institutionality eats ingenuity for breakfast

The support from our agency partners has been absolutely critical for the deployment of our two pilots. A central enabling component of the FMS-TQ system is the Bluetooth beacons, installed on buses and/or bus stops. Since public transit systems are predominantly operated and/or regulated by the government, collaboration with relevant authorities is a must.

The siting of the pilots in Singapore and Boston offers a telling, and surprisingly similar, illustration of agency's impetus for innovation. A fellow student once asked me why the LTA and the MBTA both want to work with us on this project. After all, these two agencies seem a world apart on paper: one considered a "world-class" system, the other being the subject of public derision most recently. In my opinion, the answer lies in large part in their shared sense of crisis. As mentioned in Chapter 5, the LTA feels great pressure to improve the local transit service and its public image, given the service shortfalls that have become more frequent in recent years. While Singapore's public transport system is still admittedly one of the best in the world, it has fallen short in the eyes of people who hold it to high expectations. In this context, the LTA sees the FMS-TQ well-aligned with their objectives to pursue more customer-centric initiatives. The MBTA, similarly, feels the same kind of pressure. As described in this thesis's opening and Chapter 5, the MBTA is

eager to repair its public image following the recent public disappointments. With a new leadership team -- Secretary of Transportation, MBTA's General Manager and board members – the MBTA and MassDOT are keen to set a more data-driven, innovative, and customer-oriented tone. The Strategic Initiatives team, our counterpart at the MBTA, has been especially interested in ways to better measure service performance in context of customer experience. Again, we struck the perfect timing. After hearing about the project from Professor Zegras in late August, the MBTA met with the MIT team several weeks later and quickly jumped onboard with this collaboration. Of course, our existing, long-term relationships with both the LTA and the MBTA have helped cement these collaborations with unprecedented speed and support.

Despite the similarities in motivation, there still exist huge institutional differences between the two transit agencies that have significant implications for our pilots. Recall my earlier note that the LTA, as of 2015, did not own any buses or operate bus services in Singapore. As a result, the design of our pilot was restricted to using four LTA-owned bus stops as beacon sites. On the other hand, since the MBTA owned and operated the buses, they were able to give us the authorization to install beacons on buses and at all stops. Leadership from the MBTA Board, Operations team, Vehicle Engineering, as well as the Massachusetts Department of Transportation (MassDOT) have enthusiastically embraced and supported the project; some of these colleagues have been instrumental in helping us cut the conventional bureaucratic tapes. The power of this invaluable institutional support becomes the most obvious in getting the approval for installing beacons at the bus stops. We initially ruled out this possibility, considering that the local bus stops were owned by the City rather than the MBTA. But an MBTA board member, upon hearing about this barrier and weighing it against the benefits of having beacons at the stops, got us the necessary approval from the City within a week. More than once have we been pleasantly surprised by the entrepreneurial spirit of our MBTA colleagues in helping us make things happen -- things that we originally thought to be subject to complex processes and constraints end up being quite simple. This illustrates that bureaucracy, as much as it tends to involve cumbersome procedures, also has the precise power to execute and make things happen on a large scale.

Despite the good will to expedite the process, bureaucracy remains very much a reality. Our agency partners have been very open about the institutional constraints, notably when it involves money changing hands. Despite the low financial requirements of our project (the beacons cost only $20 each), had the LTA or the MBTA been the one to procure them, the project might have lost its characteristically speedy momentum. We have been lucky to be able to fund and procure the necessities through the Singapore-MIT Alliance for Research and Technology (SMART). Both pilots include a sweepstake for participants, which we handle on our own as well to in order to free up more cumbersome processes required of public agencies. Nonetheless, as each implementation represents a collaboration of three large institutions – MIT, SMART, and the respective government authority – political economy always remains front and center. In fact, even by the time we were all ready to launch the Boston pilot, we were informed about a series of significant bureaucratic requirements from MIT, which sent us scrambling and delayed the public launch by another two weeks. Learning to be aware of and manage all the bureaucratic procedures of various institutions has been a major demand for this project, and reflects one of the fundamental skills of a transportation planning career.

For us transportation professionals, bureaucracy is an inevitable part of our entire career. Though transportation projects are traditionally politically-intensive and subject to complex and long processes, our project has demonstrated that things can also be expedited and simplified. One of the keys to successful implementation conforms to the aphorism "being at the right place at the right time" – or, in other words, finding government partners who are keen for innovation and improvements in a time of urgency.

Lesson #2: Building consensus

Related to the topic of institutional context is managing the differences among the collaborating organizations. Whether in Singapore or Boston, it has been a privilege to benefit from the talents and resources of three institutions for this project – MIT, SMART, and the LTA/MBTA. At the same time, these entities and teams have different objectives, modus operandi, and considerations. These differences have manifest themselves in many key decision points of the design and implementation processes; it is hence crucial to know

everyone's priorities, belief systems, and constraints, and to try to find the middle ground that leads to a solution.

The hardest consensus building happened during the survey design phase for both pilots. The questionnaires used in the Singapore pilot went through 14 drafts between the LTA and SMART research teams. Despite inevitable differences in opinions, we largely accepted each other's suggestions, recognizing that we could not find the perfect solution until we deployed the platform for testing. In comparison, the survey design process for Boston was much more intensive. After reviewing several templates suggested by me, the MBTA proposed a design that consisted of three to seven questions per questionnaire, depending on the survey type. Both the lead MIT Principal Investigator and I concurred with this scheme, as we thought it struck a good balance between soliciting granular feedback and minimizing user burden. The research team at SMART also signed off on it. However, a few days into coding the survey in the app, a SMART co-Principal Investigator voiced his strong disagreement with the design. He was concerned that the questionnaires involved too many consequential screens, and demanded a simplification. It became apparent to me that I had to broker an agreement that would preserve the capability to solicit detailed, actionable passenger feedback, while reducing the number of screens involved, in order to satisfy everyone's demands. In the many rounds of suggestions and revisions that followed, I was careful to always first acknowledge my understanding of the intentions of my counterparts' suggestions, then explain the remaining issues with the current iteration with regards to the objectives of the project, propose a solution that would meet the two sides in the middle, and end with stating our openness to further feedback.

Ten days and many, many email exchanges later, the research team was able to come to a finalized design together, internally. But the story did not end there, as I still faced the dreadful task of delivering the news of significant survey design changes to our MBTA partners. Luckily, the lead Principal Investigator came to my aid, and clearly explained the academic reasons behind the changes. The MBTA colleagues accepted the proposal, and, a few minor suggestions later, the survey design was finally locked down.

The consensus-building process comes with no sugarcoat. It is largely frustrating and time-consuming, and I have certainly made a mistake or two that warranted apology

emails afterwards. Do I wish that I did not have to spend so much time and energy on it? Yes, certainly. But I am glad that I did not forgo this process in pure pursuit of "speed." This is a necessary part of doing any multi-stakeholder projects correctly and sustainably. The FMS-TQ, like much academic research, represents the exploration of unknown territories for us, and the process of arriving at a solution collaboratively has enriched our considerations and lowered the risks for future disputes. Take the story about the questionnaire design, for instance – looking back, I am glad that we had someone who pushed us hard to make the app as compact as possible for our users. I also recognize that what we went through for this project must be much simpler compared to the consensus-building required of typical transportation projects, where many people's economic or political interests would be at stake and there may be no way to reach a singular agreement among all. For this reason, I am glad to have had the opportunity to practice this in a simpler context.

Lesson #3: Hardware + software + MBTA database + public users = a very hard project

Looking back, I now realize how ambitious it was to undertake a Master's research project that involves both hardware installation and software development, as well as the general public's participation *in addition to* data from the public transport agencies' sources. Each element has brought significant amount of work and risks to the project, much beyond what I imagined in the beginning.

The most difficult part about the hardware lied in the bus fleet installation (in the Boston pilot). While the Estimote beacons can hardly be easier to install – simply peeling off the film on the back and sticking the self-adhering beacon on a clean surface – installing and subsequently calibrating them for the entire fleet of buses in operation was much more challenging. We went to the Southampton Garage multiple times, because even late at night we are only able to find a fraction of the buses in the garage each time. In the subsequent development and testing stages, we twice had to adjust the beacon settings. The beacons themselves have no connectivity to the Cloud, so they can be only adjusted via the Estimote app on a smartphone, with the phone being in the beacon's signal range. So more trips to

the bus garage and hours of bus scouting on the street ensued. We also had no way of knowing whether the beacons installed at bus stops withstood with forces of nature and theft, so we had to periodically make rounds to all Silver Line stops to inspect them.

Developing the app also involved much more troubleshooting than I imagined. Despite the fact that we had done a pilot already in Singapore, we ended up essentially coding the app from scratch for the Boston implementation due to the drastic differences in design. As mentioned in Chapter 4, the strength of the beacon signals presented some initial challenges to our app's algorithms. In initial testing, we found that the phone may not be able to detect the bus beacon signals consistently in events of extreme crowding (recall that human bodies attenuates the beacon's signal). We thus adjusted the setting of the signal strength of the onboard beacons from -12 dBm (weak) to 0 dBm (strong). While this resolves the earlier issue, it brought on a new problem, making the beacon signals so strong that they could be detected by users outside the bus – even by those on another bus nearby. In the event of denied boarding or bus bunching, the user might be in range of another bus's signal for long enough to be recognized as riding that bus. Given that bus bunching and denied boardings are both characteristic of the Silver Line during the morning rush period (especially at South Station), the engineering team experimented with alternative algorithms to filter out these noisy signals, without achieving a satisfactory outcome. After many trials, we decided to re-adjust the signal strength setting on the bus beacons to -8 dBm (normal), which ultimately proved to be the satisfactory balance in conjunction with the sophisticated app algorithms now in place. For about a month and a half, my life revolved around the typical commuter schedule, so I could ride the Silver Line during the morning and evening peaks to test the app in worst-case scenarios. I was even dreaming of testing the app in my sleep. After each testing, I would upload the debug logs and type up my notes to the frontend engineer in Singapore, who would then check the logs and let me know what happened behind the app on each of my Silver Line trips. We would devise a way to fix the problems, the engineer would implement the fix within the following few days, and the process repeats. At an incomplete count, our engineer issued over 15 test versions of the app over two months.

Product development takes many iterations, testing, time, and patience. It will likely take a long time to get things right, especially if the setup makes testing and adjustments

time-consuming. If I could go back in time, I would tell my slightly-younger, less-wise self to try to simplify the setup so it is more realistic to manage and complete as a Master's student. I would also tell her that it was absolutely the right choice to have believed in, invested in, and persisted through this project.

Lesson #4: Speak the lingo

The diversity of people involved in this project has taught me to think in their mindsets and speak in their language. This not only makes the working process more effective and pleasant, but it also gives me -- the "outsider" -- more credibility in interfacing with the stakeholders. I will share here my experiences working with two particular groups who come from very different backgrounds from mine: software engineers and bus garage mechanics.

Given the exploratory nature of the FMS-TQ project, open questions and design changes were frequent encounters. Wanting to be open and considerate to everyone's opinions, I often shared such questions and alternative plans with the software engineers on our team. But as the engineers grew more comfortable with working with me, they revealed to me their world of thinking. They told me that as engineers, they would experiment creatively with various ways to implement a task, but we as academics and decision-makers needed to first provide them with a definitive blueprint. It was within our expertise and responsibility to decide the objectives and capabilities of a platform, and avoid swaying from it afterwards if possible. I was very appreciative of this feedback, and from then onwards tried to follow that modus operandi as much as possible. In subsequent deliberations, I always pushed the research team to come to a final decision (echoing the earlier lesson on building consensus) before conveying the information to the engineers. In the testing stage, I did my best to aggregate comments from all testers and test trips and translate them into action items before sharing the results with the engineers. Of course, I am still guilty of putting the engineers through way too many iterations and constant requests, but I think the process would have been more cumbersome had I not known the engineer's modus operandi.

Over my many visits to the Southampton Garage for the Boston pilot, I have picked up some local lingo in order to build relationships with the workers there. Ever since the

first time that I stepped into the massive compound, I am aware of being a misfit – physically[12], age-wise, and in terms of education level. I needed to assert my knowledge and credibility, but at the same time did not want my MIT halo to reinforce my differences from the mechanics. While some of these gaps are not likely to be bridged, I think adapting to their vocabulary go a long way. For example, instead of talking about the "Silver Line buses", I start referring to them as "Silvers," which are further distinguished as "DMAs", "CNGs" and "hybrids," or equivalently "1000s", "1100s," and "1200s." After repeated visits, I tried to demonstrate my (very limited) knowledge of the buses and the garage.

Lesson 5: The hands-on experience gives you a greater appreciation of all the nuts and bolts that keep this massive transit system running, everyday

Despite all the bureaucracy, differences, and the physically and mentally tiring work involved, this project has given me the precious opportunity to appreciate the complexity within a city's transportation backbones. I have previously checked transit apps many times to see when the next bus would be arriving, but not until I spent hours tracking the fleet for beacon installation and testing in Boston did I begin to pay attention to how the buses are dispatched and operated. I have taken the Silver Line prior, but the time spent at the Southampton Garage – weaving through the parked buses, walking under buses hoisted in mid-air, opening up panels to install the beacons – reveals the true enormity of the system and the sweat that goes into maintaining the vehicles so they can safely serve the ridership.

What I gained from embedding myself in Boston's public transport system also led to regrets for my involvement in the Singapore pilot. I worked on that project almost entirely remotely, save a two-week stay in Singapore during MIT's spring break. As a result, I did not have nearly as good of an understanding of the local context as I did for Boston. I did not know enough about Singapore's public transport challenges, our partner's institutional priorities and modus operandi, and the minds and hearts of the riders. Working remotely also separated me from the on-the-ground trivialities of designing,

---

[12] I was informed by one mechanic that there were no women – other than a cleaner – who worked at the garage. I have indeed not seen a single woman there.

testing, and deploying the project, weakening my sense of ownership and appreciation of this discovery effort.

I have now lost count of the number of test rides I have logged on all branches of Boston's Silver Line, but I have gained a deep realization of the Silver Line's importance in serving three key corridors in the city. Prior to the project, the only experiences that I had with the Silver Line were my occasional trips to/from Logan Airport. I had never been to the south part of the Seaport District nor Roxbury; more embarrassingly, I was not even aware that Boston had a BRT down Washington Street! Like most people, my familiarity with the local transit system had been centered on the handful of subway and bus lines within the sphere of my routine destinations. As I rode the Silver Line trip after trip to test the app, the diversity and quantity of people who relied on these routes dawned on me: luggage-clad travelers, employees of the offices and factories in the Seaport District, ethnic minorities from the Roxbury and Chinatown communities. One of the most marvelous things about the public transport network is its expansiveness and connectivity -- there are so many routes all throughout the city, serving countless people and their diverse mobility needs. As we go about our daily routines, it is easy to forget about the rest of the network out there and the egalitarian ubiquity that is inherent and unique to public transportation services

The most inspiring thing that I have gained from this journey is a renewed sense of optimism for our public transport systems. Working with colleagues at the LTA and the MBTA has given me a glimpse into the talent and dedication that go into running and improving the system. I look back on the criticism and vilification of the T during the previous winter in Boston – some of which continues today – and I now better understand why this happens. So much of the hard work that has gone into a public transport system, especially one that suffers from chronic underfunding and delayed maintenance, is hidden from public view. Of course, there exist many real issues, which we should not sugarcoat by any means. Rest be assured, however, that the agency has its heart in the right place, and is actively trying to improve. What it needed is a favorable climate and political well to support its many ambitions, and this is something that we, as beneficiaries of transit service, should help lobby and fight for.

Chapter 9

Conclusion

This chapter evaluates the ambitious objectives encapsulated in the FMS-TQ project. I discuss areas for future work and conclude by pondering the place for FMS-TQ and similar platforms in future urban transportation.

9.1   Evaluation

By my own assessment, I have met the initial objectives, as encapsulated in the four research questions listed in Chapter 1, with varying success:

Elucidate current sector practices: objective well-accomplished**.** As presented in Chapter 2, I gained a deep understanding of the landscape of service quality and customer assessment in the public transport sector. More importantly, I was able to synthesize this understanding with the discourse in management and marketing fields, and marry it with the theories of customer feedback to articulate the value of service quality co-monitoring for public transport services.

Explore technical feasibility for innovation: objective met. We succeeded in developing a platform to detect bus trips in real-time and solicit customer feedback. I was also able to, from the data collected, derive additional associated trip attributes for select questionnaire responses. I consider this an important achievement. However, given the accuracy issues discovered during the data analysis, as discussed in Chapter 7, our system fell somewhat short of our original expectations.

Examine correlation between real-time and recalled satisfaction: objective not met. The Singapore pilot did not collect enough paired responses for me to analyze the potential relationships in a statistically rigorous way. Descriptive inspection of the data revealed no obvious correlations.

Investigate relationships between report trip satisfaction and trip attributes:
investigation in progress. While not enough data have been collected in the first three
weeks of the Boston pilot, I have so far demonstrated the capability to derive trip attributes
associated with questionnaire responses, and look forward to applying the proposed model
to a larger dataset by the conclusion of the Boston pilot.

Last but not the least, it is no small feat to have deployed two pilots with our agency
partners in two countries. This is a great reflection of MIT's motto, *Mens et Manus* (Mind
and Hand) – turning the pursuit of knowledge into practical applications.

9.2   Areas for Future Research

The immediate next steps are to continue marketing and outreach for the Boston
pilot, develop a method to analyze accelerometer data from FMS-TQ and extract historic
weather conditions, and refine data fusing and inference described in Chapter 7. I will also
evaluate various modelling techniques and specifications, including the structural equation
model proposed at the end of Chapter 7, to identify the appropriate model for the full
dataset upon the pilot's conclusion.

I also look forward to see future improvements of the FMS-TQ product in several
areas. First, reducing -- or even eliminating -- the dependency on Bluetooth beacons
would likely improve the platform's scalability. While our use of iBeacons has enabled, for
the first time, real-time detection of public transport trips by smartphones, there are
numerous practical downsides as well. Implementation, especially installing beacons on
buses and stops, is extremely dependent on the support and jurisdiction of local public
transport agencies. As discussed in my reflection in Chapter 8, this hardware component
can consume significant time and effort in testing, calibration, and future maintenance. If
we were to implement this platform for a city's entire bus network, it could well be
impractical to install and maintain iBeacons at every bus stop. As FMS improves its
algorithms for travel mode identification, and as the computational capability of our
smartphones and computers continues to become more powerful, it may be possible to
achieve real-time inference solely using phone sensor data and backend GTFS

information.

Looking out to the longer term, I envision the functionality of, and spirit behind, the FMS-TQ platform to take a different shape. Rather than being a standalone app, the capabilities can be integrated into public transport apps. Imagine the likely near-future where most passenger use their smartphones to plan for trips, check real-time information, and pay for rides. Under that scenario, it would not be very difficult to add a feedback feature onto this app. Users would be able to proactively provide information on their trip experience at any time, and they may get prompted occasionally by the app to give feedback on a trip. This approach would likely impose minimal burden on the passengers, while having a better chance of collecting more responses given the large user base.

Further work on the institutional and policy implications are as important as improving the technology itself. How should public transport operators or regulators use the new data to monitor performance and improve service? What are the likely barriers and the corresponding solutions? What would be a fitting framework to measure success? These questions delve deep into the scholarship on public administration and institutional design, and call for interviews with stakeholders from within the system to better understand the organizational and political constraints and opportunities.

9.3   Looking Forward: Enriching Mobility as a Service

As the project came to fruition over the last one and a half years, it has coincided with the rise to prominence of the concept of Mobility-as-a-Service (MaaS). Traditionally, we have thought of our daily travel needs as trips from origin to destination, but with emerging technology, business models, and social norms, the thinking on meeting mobility demands is gradually shifting towards a more service-oriented paradigm. MaaS is about offering travelers seamless mobility solutions based on their exact travel needs (Kamargianni, Matyas, Li, & Schafer, 2015). Integration of trip planning and payment for multiple transport modes is gradually opening up the possibility for users to plan and pay for their travels more efficiently. Technological platforms are poised to combine and

coordinate all public and private transportation services within a city, presenting them to users in a personalized, streamlined, and convenient fashion. Under the right information-sharing agreements with MaaS providers, local government can also benefit from harmonized data from different transport services, which would inform policy decisions at the regional level. As I write, a number of companies – MaaS.Fi, Sidewalk Labs – have already entered the MaaS field.

Why is all this important for the topic of this thesis? The "-as-a-Service" model, at its core, entails providing clients with a solution for their needs. This implies putting customers' needs, experiences, and satisfaction front and center in the service delivery. As discussed in Chapter 2, public transport has traditionally lacked the means to demonstrate empathy to customers – one of the five pillars of service quality in the RATER framework. Hence, MaaS is a call to service providers to develop capabilities to collect better data on customer experience, build customer satisfaction into the organization's decision-making processes. This is precisely the spirit underlying FMS-TQ – to enrich the meaning of and capacity for the word "service" in public transport services.

Bibliography

Abou-Zeid, M., Witter, R., Bierlaire, M., Kaufmann, V., & Ben-Akiva, M. (2012). Happiness and travel mode switching: Findings from a Swiss public transportation experiment. *Transport Policy*, *19*(1), 93–104. http://doi.org/10.1016/j.tranpol.2011.09.009

Agrawal, A., Granger-Bevan, S., Newmark, G., & Nixon, H. (2015). *Comparing Data Quality and Cost from Three Modes of On-Board Transit Passenger Surveys* (No. 6–2015). Mineta Transportation Institute Publications. Retrieved from http://scholarworks.sjsu.edu/mti_publications/190/

Anderson, M. (2015, October 29). The Demographics of Device Ownership [Pew Research Center]. Retrieved from http://www.pewinternet.org/2015/10/29/the-demographics-of-device-ownership/

Andrews, M., Pritchett, L., & Woolcock, M. (2012). Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA). *Center for Global Development*, (Working Paper 299). Retrieved from http://ssrn.com/abstract=2102794

Apple. (2014, June 2). Getting Started with iBeacon, version 1.0. Retrieved from https://developer.apple.com/ibeacon/Getting-Started-with-iBeacon.pdf

APTA. (2015). *2015 Public Transportation Factbook*. American Public Transit Association. Retrieved from http://www.apta.com/resources/statistics/Documents/FactBook/2015-APTA-Fact-Book.pdf

Axhausen, K. W., Löchl, M., Schlich, R., Buhl, T., & Widmer, P. (2007). Fatigue in long-duration travel diaries. *Transportation*, *34*(2), 143–160. http://doi.org/10.1007/s11116-006-9106-4

Axhausen, K., Zimmermann, A., Schonfelder, S., Rindsfuser, G., & Haupt, T. (2002). Observing the rhythms of daily life: A six-week travel diary. *Transportation*, *29*(2), 95–124.

Babakus, E., & Boller, G. W. (1992). An empirical assessment of the SERVQUAL scale. *Journal of Business Research*, *24*(3), 253–268. http://doi.org/10.1016/0148-2963(92)90022-4

Baltes, M. (2002). *Surveying For Public Transit: A Design Manual for Customer On-Board Surveys* (No. NCTR-416-083). National Center For Transit Research (NCTR). Retrieved from http://www.worldtransitresearch.info/research/2539/

Barabino, B., Deiana, E., & Tilocca, P. (2012). Measuring service quality in urban bus transport: a modified SERVQUAL approach. *International Journal of Quality and Service Sciences*, *4*(3), 238–252. http://doi.org/10.1108/17566691211269567

Barkin, R. (2009, February 1). Organized for service. Retrieved March 16, 2016, from http://americancityandcounty.com/technology/hardware_software/crm-311-systems-government-service-response-200902

Batarce, M., Munoz, J. C., de Dios Ortuzar, J., Raveau, S., Mojica, C., & Alberto Rios, R. (2015). Valuing Crowding in Public Transport Systems Using Mixed Stated/Revealed Preferences Data: the Case of Santiago. In *TRB 94th Annual Meeting Compendium of Papers*. Washington, D.C. Retrieved from http://docs.trb.org/prp/15-4703.pdf

Berry, L. L., & Parasuraman, A. (1997). Listening to the Customer -- The Concept of a Service-Quality Information System. *Sloan Management Review*, *38*(3), 65–76.

Berry, L. L., Zeithaml, V. A., & Parasuraman, A. (1990). Five Imperatives for Improving Service Quality. *Sloan Management Review*, *31*(4), 29–38.

Bitner, M. J., & Hubbert, A. R. (1994). Encounter Satisfaction versus Overall Satisfaction versus Quality: The Customer's Voice. In *Service Quality: New Directions in Theory and Practice* (pp. 72–94). 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. Retrieved from http://sk.sagepub.com/books/service-quality/n3.xml

Block-Schachter, D. (2012). *Hysteresis and urban rail : the effects of past urban rail on current residential and travel choices* (Ph.D Dissertation). Massachusetts Institute of Technology, Cambridge, MA. Retrieved from http://hdl.handle.net/1721.1/73697

Boroyan, N. (2014, February 23). Care to Take the MBTA Customer Opinion Survey? *BostInnov*. Retrieved from http://bostinno.streetwise.co/2014/02/23/take-the-mbta-customer-opinion-panel-signup-survey-boston-transit/

Boston Region Metropolitan Planning Organization. (n.d.). Cities and Towns - Boston Region MPO 2010 Census Demographic Profile. Retrieved May 15, 2016, from http://www.ctps.org/datacatalog_share/subjects/cities-and-towns

Bovaird, T., & Loeffler, E. (2012, October). We're all in this together: User and community co-production of public outcomes. University of Birmingham Institute of Local Government Studies. Retrieved from http://www.birmingham.ac.uk/Documents/college-social-sciences/government-society/inlogov/discussion-papers/inlogov-co-production-chapter.pdf

Bregman, S. (2012). *Uses of social media in public transportation*. Washington, DC: Transportation Research Board of the National Academies.

Brown, T. J., Churchill, G. A., & Peter, J. P. (1993). Improving the measurement of service quality. *Journal of Retailing*, *69*(1), 127–139. http://doi.org/10.1016/S0022-4359(05)80006-5

Budiono, O. (n.d.). *Customer Satisfaction in Public Bus Transport -- a study of travelers' perception in Indonesia*. Karlstad University.

Butts, K. (2014). *Design and deploy : iterative methods in adapting mobile technologies for data acquisition : a case study in St. Louis, Missouri* (Master's thesis). Massachusetts Institute of Technology. Retrieved from http://hdl.handle.net/1721.1/97344

California DMV. (2012, January 24). DMV Wants Your Feedback. Retrieved May 11, 2016, from https://www.dmv.ca.gov/portal/dmv/detail/pubs/newsrel/newsrel12/2012_01

Cantwell, M., Caulfield, B., & O'Mahony, M. (2009). Examining the Factors that Impact Public Transport Commuting Satisfaction. *Journal of Public Transportation*, *12*(2).

Carrel, A., Mishalani, R., Sengupta, R., & Walker, J. (2015). In Pursuit of the Happy Transit Rider: Dissecting Satisfaction Using Daily Surveys and Tracking Data. *Submitted for Possible Publication in the Journal of Intelligent Transportation Systems*.

Carrel, A., Sengupta, R., & Walker, J. (2015). The San Francisco Travel Quality Study: Tracking Trials and Tribulations of a Transit Taker. *Submitted for Possible Publication in Transportation*.

Carson, B. (2015, June 3). More than 1 million people have now worked as an Uber driver. *Business Insider*. Retrieved from http://www.businessinsider.com/uber-hits-1-million-drivers-2015-6

Celik, E., Aydin, N., & Gumus, A. T. (2014). A multiattribute customer satisfaction evaluation approach for rail transit network: A real case study for Istanbul, Turkey. *Transport Policy*, *36*, 283–293. http://doi.org/10.1016/j.tranpol.2014.09.005

Ching, A. (2012). *A user-flocksourced bus intelligence system for Dhaka*. Massachusetts Institute of Technology. Retrieved from http://hdl.handle.net/1721.1/77836

Chow, W. (2014, June). *Evaluating Online Surveys for Public Transit Agencies Using a Prompted Recall Approach*. Massachusetts Institute of Technology. Retrieved from http://hdl.handle.net/1721.1/89853

Churchill, G. A., & Surprenant, C. (1982). An Investigation into the Determinants of Customer Satisfaction. *Journal of Marketing Research*, *19*(4), 491. http://doi.org/10.2307/3151722

Cialdini, R. B. (2006). *Influence: the psychology of persuasion* (Rev. ed., [Nachdr.]). New York, NY: Collins.

City of Boston. (n.d.). Focus on Basic City Services | Boston About Results (BAR) | City of Boston. Retrieved March 19, 2016, from http://www.cityofboston.gov/BAR/basicservices.asp

City of Boston - Downtown - Boston, MA. (n.d.). Retrieved March 13, 2016, from https://www.yelp.com/biz/city-of-boston-boston-3

Clark, B., & Rokakis, M. (2014). Do 311 Systems Shape Citizen Satisfaction with Local Government? *SSRN*. Retrieved from http://ssrn.com/abstract=2491034

Collins, C., Hasan, S., & Ukkusuri, S. (2013). A Novel Transit Rider Satisfaction Metric: Rider Sentiments Measured from Online Social Media Data. *Journal of Public Transportation*, *16*(2).

Cook, J. (2015, February 11). Uber's internal charts show how its driver-rating system actually works. *Business Insider*. Retrieved from http://www.businessinsider.com/leaked-charts-show-how-ubers-driver-rating-system-works-2015-2?r=UK&IR=T

Couzin, G., & Grappone, J. (2014). *Five stars: putting online reviews to work for your business*. Indianapolis, Ind: Sybex.

Credo Business Consulting, Siemens AG. (2014). *The Mobility Opportunity*. Retrieved from http://www.credo-group.com/downloads/Mobility%20Opportunity%20Study.pdf

Cronin, J. J., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, *76*(2), 193–218. http://doi.org/10.1016/S0022-4359(00)00028-2

Cronin, J. J., & Taylor, S. A. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing*, *56*(3), 55–68.

de Oña, J., de Oña, R., Eboli, L., & Mazzulla, G. (2013). Perceived service quality in bus transit service: A structural equation approach. *Transport Policy*, *29*, 219–226. http://doi.org/10.1016/j.tranpol.2013.07.001

de Oña, R., Eboli, L., & Mazzulla, G. (2014). Monitoring Changes in Transit Service Quality over Time. *Procedia - Social and Behavioral Sciences*, *111*, 974–983. http://doi.org/10.1016/j.sbspro.2014.01.132

Del Castillo, J. M., & Benitez, F. G. (2012). A Methodology for Modeling and Identifying Users Satisfaction Issues in Public Transport Systems Based on Users Surveys. *Procedia - Social and Behavioral Sciences*, *54*, 1104–1114. http://doi.org/10.1016/j.sbspro.2012.09.825

dell'Olio, L., Ibeas, A., & Cecin, P. (2011). The quality of service desired by public transport users. *Transport Policy*, *18*(1), 217–227. http://doi.org/10.1016/j.tranpol.2010.08.005

Department of Statistics, Singapore. (2015). *Singapore in Figures 2015*. Department of Statistics, Singapore. Retrieved from https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/reference/sif2015.pdf

Dudley, E., Lin, D.-Y., Mancini, M., & Ng, J. (2015, July). Implementing a citizen-centric approach to delivering government services. McKinsey. Retrieved from http://www.mckinsey.com/industries/public-sector/our-insights/implementing-a-citizen-centric-approach-to-delivering-government-services

Dunlop, I., Casello, J., & Doherty, S. (2015). Tracking Transit Rider Experience: Using Smartphones to Measure Comfort and Well-being Throughout the Trip`. In *Public Transportation Marketing and Fare Policy*. Washington, D.C.

Eboli, L., & Mazzulla, G. (2007). Service Quality Attributes Affecting Customer Satisfaction for Bus Transit. *Journal of Public Transportation*, *10*(3), 35–46.

Eboli, L., & Mazzulla, G. (2008). A Stated Preference Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology*, *31*(5), 509–523. http://doi.org/10.1080/03081060802364471

Eboli, L., & Mazzulla, G. (2009). A New Customer Satisfaction Index for Evaluating Transit Service Quality. *Journal of Public Transportation*, *12*(3).

Eboli, L., & Mazzulla, G. (2010). How to Capture the Passengers' Point of View on a Transit Service through Rating and Choice Options. *Transport Reviews*, *30*(4), 435–450. http://doi.org/10.1080/01441640903068441

Eboli, L., & Mazzulla, G. (2011). A methodology for evaluating transit service quality based on subjective and objective measures from th passenger's point of view. *Transport Policy*, *18*, 172–181.

Eboli, L., & Mazzulla, G. (2012). Performance indicators for an objective measure of public transport service quality. *European Transport / Trasporti Europei*, (51). Retrieved from http://hdl.handle.net/10077/6119

Efthymiou, D., Kaziales, M., Antoniou, C., & Tyrinopoulos, Y. (2014). Measuring the Effects of Economic Crisis on Users' Perceptions of Public Transport Quality. *Transportation Research Record: Journal of the Transportation Research Board*, *2415*, 1–12. http://doi.org/10.3141/2415-01

Elmore-Yalch, R. (1998). *A handbook: integrating market research into transit management*. Washington, D.C: National Academy Press.

eMarketer. (2015, August). Number of smartphone users in the United States from 2010 to 2019 (in millions). In *Statistica*. Retrieved from http://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/

Estimote. (2013, July 16). Preorder for Estimote Beacons available, shipping this summer. Retrieved from http://blog.estimote.com/post/57087851702/preorder-for-estimote-beacons-available-shipping

European Committee for Standardization. (2002). *European Standard EN 13816. Transportation - Logistics and services - Public passenger transport - Service quality defintion, targeting and measurement* (No. EN 13816:2002 E). European Committee for Standardization.

Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research*, *15*(2), 195–219. http://doi.org/10.1108/10662240510590360

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, *46*(4–5), 687–724. http://doi.org/10.1016/S0014-2921(01)00208-2

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, *60*(4), 7. http://doi.org/10.2307/1251898

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, *65*(1), 45–55. http://doi.org/10.1037/0022-3514.65.1.45

Friman, M., Edvardsson, B., & Gärling, T. (2001). Frequency of negative critical incidents and satisfaction with public transport services. I. *Journal of Retailing and Consumer Services*, *8*(2), 95–104. http://doi.org/10.1016/S0969-6989(00)00003-5

Friman, M., & Fellesson, M. (2009). Service supply and customer satisfaction in public transportation: The Quality Paradox. *Journal of Public Transportation*, *12*(4), 57–69.

Gigler, B.-S., & Bailur, S. (Eds.). (2014). *Closing the Feedback Loop: Can Technology Bridge the Accountability Gap?* The World Bank. Retrieved from http://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-0191-4

Goodyear, S. (n.d.). The Invention of 3-1-1 and the City Services Revolution. Retrieved May 11, 2016, from http://www.citylab.com/city-makers-connections/311/

Governor's Special Panel. (2015). *Back on Track: An Action Plan to Transform the MBTA*. Commonwealth of Massachusetts. Retrieved from http://www.mass.gov/governor/docs/news/mbta-panel-report-04-08-2015.pdf

Habib, K. M. N., Kattan, L., & Islam, M. T. (2011). Model of personal attitudes towards transit service quality. *Journal of Advanced Transportation*, *45*(4), 271–285. http://doi.org/10.1002/atr.106

Halpern, S. D., Kohn, R., Dornbrand-Lo, A., Metkus, T., Asch, D. A., & Volpp, K. G. (2011). Lottery-Based versus Fixed Incentives to Increase Clinicians' Response to Surveys: Lottery-Based versus Fixed Incentives. *Health Services Research*, *46*(5), 1663–1674. http://doi.org/10.1111/j.1475-6773.2011.01264.x

Harman, C. (2014, April 22). iBeacons and BLE. Retrieved from https://calvium.com/2014/04/ibeacons-and-ble/

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52. http://doi.org/10.1002/dir.10073

Hensher, D. A., Stopher, P., & Bullock, P. (2003). Service quality—developing a service quality index in the provision of commercial bus contracts. *Transportation Research Part A: Policy and Practice*, *37*(6), 499–517. http://doi.org/10.1016/S0965-8564(02)00075-7

Holbrook, A., Krosnick, J., & Pfent, A. (2008). The Causes and Consequences of Response Rates in Surveys. In *Advances in Telephone Survey Methodology*. John Wiley & Sons, Inc.

Holzer, M., Schwester, R., McGuire, A., & Kloby, K. (2006). State-level 311 Systems: Leveraging Service Enhancement and Performance Measurement at the State Level. *The Book of States 2006*, 409–413.

Housing & Development Board. (2016). *Housing & Development Board 2015 Annual Report - Key Statistics* (Annual report). Singapore Government. Retrieved from http://www.hdb.gov.sg/cs/infoweb/about-us/news-and-publications/annual-reports

Howe-Teo, R. (2015, July). *Singapore's Smart Mobility 2030: Big Data and Car-Lite Society*. Retrieved from www.rahs.gov.sg/public/www/download.ashx?id=1361

Iacobucci, D., Amy, O., & Kent, G. (1995). Distinguishing Service Quality and Customer Satisfaction: The Voice of the Consumer. *Journal of Consumer Psychology*, *4*(3), 277–303.

I Paid A Bribe. (n.d.). I Paid A Bribe | Report Now. Retrieved March 20, 2016, from http://www.ipaidabribe.com/#gsc.tab=0

iBeaconinsider. (2014, March 19). What is iBeacon? A Guide to iBeacons. Retrieved from http://www.ibeacon.com/what-is-ibeacon-a-guide-to-beacons/

IBM. (n.d.). IBM - Infographic: The Life of the Connected Commuter. Retrieved March 20, 2016, from http://www.columnfivemedia.com/work-items/infographic-the-life-of-the-connected-commuter

ICMA. (2013). *ICMA State of the Profession 2012 Survey Results*. Retrieved from http://icma.org/en/icma/knowledge_network/documents/kn/Document/305096/ICMA_2012_State_of_the_Profession_Survey_Results

Iseki, H., & Taylor, B. (2010). Style versus Service? An Analysis of User Perceptions of Transit Stops and Stations. *Journal of Public Transportation*, *13*(3), 23–48. http://doi.org/10.5038/2375-0901.13.3.2

Kamargianni, M., Matyas, M., Li, W., & Schafer, L. (2015). *Feasibility study for "Mobility as a Service" concept in London*. UCL Energy Institute. Retrieved from https://www.bartlett.ucl.ac.uk/energy/docs/fs-maas-compress-final

Karlaftis, M. G., Golias, J., & Papadimitriou, E. (2001). Transit quality as an integrated traffic management strategy: Measuring perceived service. *Journal of Public Transportation*, *4*(1), 27–44.

Kaufman, S. (2014). *Co-Monitoring for Transit management -- Using Web-Based Input for Transit Management*. Rudin Center for Transportation, NYU Wagner School.

Kim, Y. K., & Lee, H. R. (2011). Customer satisfaction using low cost carriers. *Tourism Management*, *32*(2), 235–243. http://doi.org/10.1016/j.tourman.2009.12.008

Kittelson & Associates, National Research Council (U.S.), Transit Cooperative Research Program, United States, & Transit Development Corporation (Eds.). (2013). *Transit capacity and quality of service manual* (3rd ed). Washington, D.C: Transportation Research Board.

LACMT. (n.d.). Metro Research. Retrieved March 4, 2016, from https://www.metro.net/news/research/

Land Transport Authority. (2011, November). Passenger Transport Mode Shares in World Cities. *JOURNEYS*. Retrieved from http://www.lta.gov.sg/ltaacademy/doc/J11Nov-p60PassengerTransportModeShares.pdf

Land Transport Authority. (2013a). *Land Transport Master Plan 2013*. Land Transport Authority. Retrieved from https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/ReportNewsletter/LTMP2013Report.pdf

Land Transport Authority. (2013b, January 17). Two New Rail Lines and Three New Extensions to Expand Rail Network by 2030. Retrieved May 16, 2016, from http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=38dc4ca3-5e70-4bf8-97bc-87f78e6303e7

Land Transport Authority. (2015a). Public Transport Ridership. Retrieved from https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/PT%20Ridership.pdf

Land Transport Authority. (2015b). Statistics in Brief, 2015. Retrieved from https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/Statistics%20in%20Brief%202015%20FINAL.pdf

Land Transport Authority. (2016a). *COE Bidding Results 2016* (LTA Publications and Research). Land Transport Authority. Retrieved from https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/COE_Result_2015_2016.pdf

Land Transport Authority. (2016b). Vehicle statistics (yearly). Retrieved from http://www.lta.gov.sg/content/ltaweb/en/publications-and-research.html

Land Transport Authority. (2016c, March 9). Public Transport Customer Satisfaction Survey 2015 - Annex-A. Retrieved April 24, 2016, from http://www.lta.gov.sg/data/apps/news/press/2016/20160309_PTCSSAnnex-A.pdf

Land Transport Authority. (2016d, May). Cost (S$) For Cars Registered in Apr 2016. Retrieved May 16, 2016, from
http://www.onemotoring.com.sg/publish/onemotoring/en/lta_information_guidelines/buy_a_new_vehicle/car_cost.MainPar.30963.File.tmp/Car_Cost_Update.pdf

Land Transport Authority. (n.d.). Bus Service Enhancement Programme (BSEP). Retrieved May 16, 2016, from //www.mytransport.sg/content/mytransport/ptp_drawer_container.html

Land Transport Authority. (n.d.). Bus services. Retrieved May 16, 2016, from //www.mytransport.sg/content/mytransport/ptp_drawer_container.html

Land Transport Authority. (n.d.). Electronic Road Pricing (ERP). Retrieved May 16, 2016, from https://www.lta.gov.sg/content/ltaweb/en/roads-and-motoring/managing-traffic-and-congestion/electronic-road-pricing-erp.html

Land Transport Authority (Singapore). (2015). Public Transport Customer Survey 2014. Retrieved from
http://www.lta.gov.sg/data/apps/news/press/2015/20150311_COSPTCSS14-AnnexA.pdf

Leung, G. M., Johnston, J. M., Saing, H., Tin, K. Y. K., Wong, I. O. L., & Ho, L.-M. (2004). Prepayment was superior to postpayment cash incentives in a randomized postal survey among physicians. *Journal of Clinical Epidemiology*, *57*(8), 777–784. http://doi.org/10.1016/j.jclinepi.2003.12.021

Leung, S. (2013a, August 16). A snarl in Seaport District's success story. *Boston Globe*. Retrieved from https://www.bostonglobe.com/business/2013/08/15/seaport-district-faces-gridlock-development-outpaces-its-transportation-plan/V213FxpY9dmLGtAwQmkQGP/story.html

Leung, S. (2013b, October 4). The Seaport's Silver Line solution. *Boston Globe*. Retrieved from https://www.bostonglobe.com/business/2013/10/03/silver-line-silver-lining-seaport-district-growing-traffic-woes/Bc0TSNoAGr93IL2Ihlwa0N/story.html

Levine, C., & Fisher, G. (1984). Citizenship and Service Delivery: The Promise of Coproduction. *Public Administration Review*, *44*(Special Issue: Citizenship and Public Administration), 178–189.

Li, C., & Zhao, J. (2016). Humanizing Travel: How E-Hail Transforms Driver-Passenger Dynamics. *Submitted to Transport Policy*.

Love, R. (2013, August 6). Why Does GPS Use More Battery Than Any Other Antenna Or Sensor In A Smartphone? Retrieved April 28, 2016, from

http://www.forbes.com/sites/quora/2013/08/06/why-does-gps-use-more-battery-than-any-other-antenna-or-sensor-in-a-smartphone/

LTA. (2013, October 7). Household Interview Travel Survey 2012: Public Transport Mode Share Rises to 63%. Retrieved March 26, 2016, from http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=1b6b1e1e-f727-43bb-8688-f589056ad1c4

LTA. (2014). Statistics in Brief, 2014. Retrieved from https://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/Statistics%20in%20Brief%202014.pdf

LTA. (2015). Rail Length (km) At End-of-Year. Retrieved from http://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/FactsandFigures/Rail%20Length.pdf

MARTA. (n.d.). *Relationship Between Passenger Environment & Customer Satisfaction*. Retrieved from http://www.apta.com/mc/rail/previous/rail/presentations/Presentations/William%20Cooper.pdf

Martin, W. E. (2014). Point of Contact. *Government Technology*. Retrieved from http://www.govtech.com/magazines/pcio/Point-of-Contact.html

Massachusetts Taxpayers Foundation. (2015). *The End of Its Line*. Massachusetts Taxpayers Foundation. Retrieved from http://www.masstaxpayers.org/sites/masstaxpayers.org/files/MTF_The%20T-The%20End%20of%20Its%20Line.pdf

MassDOT. (2015). *MassDOT's Annual Performance Report* - (No. Fiscal Year 2015). MassDOT Office of Performance Management and Innovation. Retrieved from http://www.massdot.state.ma.us/Portals/0/docs/infoCenter/performancemanagement/AnnualPerformanceReport_2015.pdf

MBTA. (2014). *Ridership and Service Statistics* (No. Fourteenth Edition). MBTA. Retrieved from http://www.mbta.com/uploadedfiles/documents/2014%20BLUEBOOK%2014th%20Edition.pdf

MBTA. (2016, April). MBTA Performance Dashboard. Retrieved May 8, 2016, from http://www.mbtabackontrack.com/performance/index.html#/detail/customersatisfaction

McGoogan, C. (2016, January 26). Uber will monitor how safe your driver is through their smartphone. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/technology/2016/01/26/uber-will-monitor-how-safe-your-driver-is-through-their-smartpho/

minimaxir. (2014, August 4). Distribution of Yelp Ratings for Businesses, by Business Category. Retrieved March 13, 2016, from https://www.reddit.com/r/dataisbeautiful/comments/2clcbv/distribution_of_yelp_ratings_for_businesses_by/

Moore, R., & Moore, M. (2004). The Impact of Firm Response Time to Email Communication. *Marketing Management Journal*, *14*(2), 1–12.

Morpace International, Cambridge Systematics, National Research Council (U.S.), Transit Development Corporation, & Transit Cooperative Research Program (Eds.). (1999). *A handbook for measuring customer satisfaction and service quality*. Washington, D.C: National Academy Press.

Munar, A. M., & Jacobsen, J. K. S. (2014). Motivations for sharing tourism experiences through social media. *Tourism Management*, *43*, 46–54. http://doi.org/10.1016/j.tourman.2014.01.012

New York City Transit. (2014). *2014 Customer Satisfaction Survey - Subway*. Retrieved from http://web.mta.info/mta/news/books/docs/2014-Subway-CSS-Board-Presentation.pdf

Newcombe, T. (2014). Is the Cost of 311 Systems Worth the Price of Knowing? *Governing*, (March 2014). Retrieved from http://www.governing.com/topics/urban/gov-311-systems-cost.html

Newman, J. (2013, April 1). Peak Battery: Why Smartphone Battery Life Still Stinks, and Will for Years. *Time*. Retrieved from http://techland.time.com/2013/04/01/peak-battery-why-smartphone-battery-life-still-stinks-and-will-for-years/

NJ Transit. (2015). NJ TRANSIT Customer Satisfaction Survey. Retrieved from http://www.njtransit.com/var/var_servlet.srv?hdnPageAction=ScoreCardSurveyTo

Oliver, R. L. (1981). Measurement and evaluation of satisfaction processes in retail settings. *Journal of Retailing*, *57*(3), 25–48.

Oliver, R. L. (1993). Cognitive, Affective, and Attribute Bases of the Satisfaction Response. *Journal of Consumer Research*, *20*(3), 418–430.

Oracle. (2006). *NYC 311: Automated Services for the City that Never Sleeps* (Siebel case study). Oracle.

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*, *24*(6), 1073–1087. http://doi.org/10.1016/0305-750X(96)00023-X

Ostrom, E., Parks, R. B., Whitaker, G. P., & Percy, S. L. (1978a). The Public Service Production Process: A Framework for Analyzing Police Services. *Policy Studies Journal*, *7*(s1), 381–381. http://doi.org/10.1111/j.1541-0072.1978.tb01782.x

Ostrom, E., Parks, R. B., Whitaker, G. P., & Percy, S. L. (1978b). The Public Service Production Process: A Framework for Analyzing Police Services. *Policy Studies Journal*, *7*(s1), 381–381. http://doi.org/10.1111/j.1541-0072.1978.tb01782.x

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, *49*(4), 41. http://doi.org/10.2307/1251430

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, *64*(1).

Parikh, A., Behnke, C., Vorvoreanu, M., Almanza, B., & Nelson, D. (2014). Motives for reading and articulating user-generated restaurant reviews on Yelp.com. *Journal of Hospitality and Tourism Technology*, *5*(2), 160–176. http://doi.org/10.1108/JHTT-04-2013-0011

Parks, R. B., Baker, P. C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., … Wilson, R. (1981). Consumers as Co-producers of Public Services: Some Economic and Institutional Considerations. *Policy Studies Journal*, *9*(7), 1001–1011. http://doi.org/10.1111/j.1541-0072.1981.tb01208.x

Pedersen, T., Friman, M., & Kristensson, P. (2011). The role of predicted, on-line experienced and remembered satisfaction in current choice to use public transport services. *Journal of Retailing and Consumer Services*, *18*(5), 471–475. http://doi.org/10.1016/j.jretconser.2011.06.013

Percy, S. L. (1978). Conceptualizing and Measuring Citizen Co-Production of Community Safety*. *Policy Studies Journal*, *7*(s1), 486–493. http://doi.org/10.1111/j.1541-0072.1978.tb01797.x

Pew Research Center. (2015). *The Smartphone Difference*. Retrieved from http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/

Poushter, J. (2016, February 22). Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Retrieved from http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/

Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing*, *18*(3), 5–14. http://doi.org/10.1002/dir.20015

Prioni, P., & Hensher, D. (2000). Measuring service quality in scheduled bus services. *Journal of Public Transportation*, *3*(2), 51–74.

ProximityOne. (2016). *Metropolitan Area Situation & Outlook Report* (Boston-Cambridge-Newton, MA-NH Metropolitan Statistical Area (CBSA 14460)). Retrieved from http://proximityone.com/metros/2013/cbsa14460.htm

Raveau, S., Ghorpade, A., Zhao, F., Abou-Zeid, M., Zegras, C., & Ben-Akiva, M. (2016). Smartphone-based survey for real-time and retrospective happiness related to travel and activities. Presented at the Transportation Research Board 95th Annual Meeting, Washington, D.C.

RSG. (2015). *REGIONAL TRANSPORTATION AUTHORITY: 2013-2014. CUSTOMER SATISFACTION STUDY*. Retrieved from https://www.rtachicago.com/files/documents/plansandprograms/2014%20RTA%20Customer%20Satisfaction%20Report%20FINAL%203-6-15.pdf

Rudi Ball, Ajinkya Ghorpade, Kalan Nawarathne, Rui Baltazar, Francisco Camara Pereira, Chris Zegras, & Moshe Ben-Akiva. (2014). Battery Patterns and Forecasting in a Large-scale Smartphone-based Travel Survey. http://doi.org/10.13140/RG.2.1.3230.8641

Rust, R. T., & Zahorik, A. J. (1993). Customer satisfaction, customer retention, and market share. *Journal of Retailing*, *69*(2), 193–215. http://doi.org/10.1016/0022-4359(93)90003-2

SAGE stats. (n.d.). Boston-Cambridge-Newton, MA-NH. Retrieved May 15, 2016, from http://data.sagepub.com/sagestats/document.php?id=place-42485

Sager, R. (2014, May). *A Survey of Transit Agencies on Web-based Feedback Tools and their Role in Addressing Riders* (Masters thesis). Georgia Institute of Technology.

Schaeffer, C. (n.d.). How 311 Systems Improve Constituent Services and Earn Citizen Trust. Retrieved March 15, 2016, from http://www.crmsearch.com/government-311.php

Schaller, B., National Research Council (U.S.), Transportation Research Board, Transit Cooperative Research Program, United States, Federal Transit Administration, … Schaller Consulting. (2005). *On-board and intercept transit survey techniques*. Washington, D.C.: Transportation Research Board.

SeeClickFix. (2010, January 9). SeeClickFix | Distributed 311 for Distributed Governance. Retrieved March 15, 2016, from https://gov.seeclickfix.com/2010/01/09/distributed-311-for-distributed-governance/

SeeClickFix. (2015, March 10). SeeClickFix | New Haven Traffic and Parking has integrated SeeClickFix with VEOCI work order system! Retrieved March 15, 2016, from https://gov.seeclickfix.com/2015/03/10/new-haven-traffic-and-parking-has-integrated-seeclickfix-with-veoci-work-order-system/

SEPTA. (2013). *2012 Customer Satisfaction Survey Final Report*. Retrieved from http://www.septa.org/strategic-plan/reports/customer-satisfaction-survey-2012.pdf

SFMTA. (2014, March). *Muni On-Board Customer Survey*. Retrieved from https://www.sfmta.com/sites/default/files/agendaitems/3-14-14%20Item%2011%20Muni%20On-Board%20Customer%20Survey.pdf

Shaaban, K., & Khalil, R. F. (2013). Investigating the Customer Satisfaction of the Bus Service in Qatar. *Procedia - Social and Behavioral Sciences*, *104*, 865–874. http://doi.org/10.1016/j.sbspro.2013.11.181

Sharma, N., & Patterson, P. G. (1999). The impact of communication effectiveness and service quality on relationship commitment in consumer, professional services. *Journal of Services Marketing*, *13*(2), 151–170. http://doi.org/10.1108/08876049910266059

Singapore Government. (2015, December 30). LTA to acquire buses, take over procurement contracts from SBS for S$187m. Retrieved May 16, 2016, from http://www.gov.sg/news/content/the-business-times-lta-to-acquire-buses-take-over-procurement-contracts-from-sbs-for-187m

Singapore Institute of Planners. (n.d.). Singapore Insitute of Planners. Retrieved May 17, 2016, from http://www.planning.org.sg/Singapore/

Smith, G. (2008). Does gender influence online survey participation? *ERIC Document Reproduction Service*, (ED 501717). Retrieved from http://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1003&context=elementary_ed_pub

Spitz, G., Niles, F. L., & Adler, T. J. (2006). *Web-based survey techniques: a synthesis of transit practice*. Washington, D.C: Transportation Research Board.

StateScoop. (n.d.). New York City 311 system sees record use in 2014. Retrieved March 21, 2016, from http://statescoop.com/new-york-city-311-system-use-hits-record-peak-2014

Stathopoulos, A., & Marcucci, E. (2014). *De Gustibus Disputandum Est* : Non-Linearity in Public Transportation Service Quality Evaluation. *International Journal of Sustainable Transportation*, *8*(1), 47–68. http://doi.org/10.1080/15568318.2012.758523

Stewart, A. (2014). *Visualizing urban accessibility metrics for incremental bus rapid transit projects* (Master's thesis). Massachusetts Institute of Technology. Retrieved from http://hdl.handle.net/1721.1/92057

Straits Times. (2015, December 30). LTA to buy SBS buses, take over contracts. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/transport/lta-to-buy-sbs-buses-take-over-contracts

Strauss, J., & Hill, D. J. (2001). Consumer complaints by e-mail: an exploratory investigation of corporate responses and customer reactions. *Journal of Interactive Marketing*, *15*(1), 63–73. http://doi.org/10.1002/1520-6653(200124)15:1<63::AID-DIR1004>3.0.CO;2-C

Stuart, K., Mednick, M., & Bockman, J. (2000). Structural Equation Model of Customer Satisfaction for the New York City Subway System. *Transportation Research Record: Journal of the Transportation Research Board*, *1735*, 133–137. http://doi.org/10.3141/1735-16

Tan, C. (2015, December 24). MRT's past "teething problems." *The Straits Times*. Retrieved from http://www.straitstimes.com/opinion/mrts-past-teething-problems

Temkin, B. (2014). *State of Voice of Customer Programs, 2014*. Temkin Group.

Too, L., & Earl, G. (2009). Public transport service quality and sustainable development: a community stakeholder perspective. *Sustainable Development*, n/a-n/a. http://doi.org/10.1002/sd.412

Tripathi, V., Kumar, A., & Gunjan, N. (2012). Measuring Service Quality in Indian Public Transport System: A Comprehensive Review. *Asia-Pacific Marketing Review*, *I*(1), 113–119.

Trudel, V. (2014, March 14). Qu'est-ce que les beacons et à quoi servent-ils? Retrieved from http://www.chalifour.net/technologie/beacons/

Tyrinopoulos, Y., & Antoniou, C. (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy*, *15*(4), 260–272. http://doi.org/10.1016/j.tranpol.2008.06.002

U.S. Census. (2014). 2014 American Community Survey 1-Year Estimates. Retrieved from http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

Wallin Andreassen, T. (1995). (Dis)satisfaction with public services: the case of public transportation. *Journal of Services Marketing*, *9*(5), 30–41. http://doi.org/10.1108/08876049510100290

Walter, E., Steyrer, C., & Wiesel, S. (2010). Customer Satisfaction with Loyalty Card Programs in the Austrian Clothing Retail Sector - an Empirical Study of Performance Attributes Using Kano's Theory. In P. Schnedlitz, D. Morschett, T. Rudolph, H. Schramm-Klein, & B. Swoboda (Eds.), *European Retail Research* (pp. 71–95). Wiesbaden: Gabler Verlag. Retrieved from http://link.springer.com/10.1007/978-3-8349-8938-3_4

Warner, S. B. (1978). *Streetcar suburbs: the process of growth in Boston ; 1870 - 1900* (2. ed). Cambridge, Mass. ;London: Harvard Univ. Press.

Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., & McSpurren, K. (1996). Charities, No; Lotteries, No; Cash, Yes: Main Effects and Interactions in a Canadian Incentives Experiment. *Public Opinion Quarterly*, *60*(4), 542. http://doi.org/10.1086/297772

Washington Metropolitan Area Transit Authority. (2014, April 10). Public Outreach and Customer Research. Retrieved from https://www.wmata.com/about_metro/board_of_directors/board_docs/041014_4APublicOutreachandCustomerResearch.pdf

Weinstein, A. (2000). Customer Satisfaction Among Transit Riders: How Customers Rank the Relative Importance of Various Service Attributes. *Transportation Research Record: Journal of the Transportation Research Board*, *1735*, 123–132. http://doi.org/10.3141/1735-15

Weinstock, A., Hook, W., Replogle, M., & Cruz, R. (2011). *Recapturing Global Leadership in Bus Rapid Transit: A Survey of Select U.S. Cities*. Institute of Transportation and Development Policy (ITDP). Retrieved from https://www.itdp.org/wp-content/uploads/2014/07/20110526ITDP_USBRT_Report-HR.pdf

Whelton, A., Dietrich, A., Gallagher, D., & Roberson, A. (2007). Using customer feedback for improved water quality and infrastructure monitoring. *American Water Works Association Journal*, *99*(11).

Whitaker, G. (1980). Coproduction: Citizen Participation in Service Delivery. *Public Administration Review*, *40*(3), 240–246.

Wilson, J., & Kelling, G. (1982, March). Broken Windows. *Atlantic Monthly*.

Wirtz, J., & Tomlin, M. (2000). Institutionalising customer-driven learning through fully integrated customer feedback systems. *Managing Service Quality: An International Journal*, *10*(4), 205–215. http://doi.org/10.1108/09604520010341654

Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology &Tourism*, *10*(4), 283–295.

Zegras, C., Stewart, A., Forray, R., Hidalgo, R., Figueroa, C., Duarte, F., & Wampler, J. (2016). Designing BRT-oriented development. In *Restructuring public transport through bus rapid transit: an international and interdisciplinary perspective*. Bristol ;Chicago, Il: Policy Press.

Zegras, P. C., Butts, K., Cadena, A., & Palencia, D. (2015). Spatiotemporal dynamics in public transport personal security perceptions: digital evidence from Mexico City's periphery. In *ICT for transport: opportunities and threats*. Cheltenham: Elgar.

Zhao, F., Pereira, F. C., Ball, R., Kim, Y., Han, Y., Zegras, C., & Ben-Akiva, M. (2015). Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board*, *2*, 45–56. http://doi.org/10.3141/2494-06

Zimmerman, J., Tomasic, A., Garrod, C., Yoo, D., Hiruncharoenvate, C., Aziz, R., … Steinfeld, A. (2011). Field trial of Tiramisu: crowd-sourcing bus arrival times to spur co-design (p. 1677). ACM Press. http://doi.org/10.1145/1978942.1979187

Appendix A

Singapore pilot questionnaires

Test in black appears on the screen. Text in red denote questionnaire logic and flow control.

Entrance questionnaire

A1. On average, how many days a week do you take…?
- a. Bus

  (If answer > 0)Do you usually check real-time bus arrival info before the trip?
- b. Train (e. g. MRT/LRT)
- c. Taxi

A2. What's the top reason for you to take the bus?
(Choose 1 answer only)
- Faster than alternatives
- Cheaper than alternatives
- More convenient than alternatives
- I have no other choice
- Other (comment)

A3. How satisfied are you with your experience using the bus in your city?
(Rate on scale of 1-5: 1=very dissatisfied, 3 = neutral, 5 = very satisfied)

A4. How well do the following statements describe you?
(Answer with "Strongly disagree", "disagree", "neutral", "agree", or "strongly agree")
- a. I like using public transport.
- b. I think about the environmental impact when I travel
- c. The only good thing about travelling is arriving at the destination.
- d. I use my travel time productively.
- e. Travelling stresses me out.
- f. Traveling is generally tiring for me.

A5. How well does the following statement describe you?
(Answer with "Strongly disagree", "disagree", "neutral", "agree", or "strongly agree")
- a. I am generally satisfied with my life.

A6. What is your gender?
(Answer choices: Male, Female, Prefer not to answer)

A7. What is your employment status?
(Answer choices: Employee, Self-employed, Work from home, Unemployed, Student, Retired, Prefer not to answer)

A8. How many cars are available to your household?
(Answer choices: 0, 1, 2 or more, Prefer not to answer)

At-the-stop questionnaire

B0. Are you waiting at (location) bus stop now?
Did you wait at (location) bus stop at around (time)?
- Yes
- No, I passed by the stop
- No, I didn't go to that stop at all

(If No, go to statement "Sorry for the inconvenience! Click Submit to exit.)

B1. What is the purpose of this trip?
What was the purpose of that trip?
(Choose 1 answer only)
- Work
- School
- Shopping/Eating
- Leisure activities
- Errands
- Going home
- Other

B2. Will/did you make any transfers on this trip?

B3. How do you feel now?
How did you feel while at the stop?
(Rate each of the following dimension on a scale of 1-5, or pick a point on a toggle bar with the following descriptions at each end)
   a. Hurried - relaxed
   b. Anxious – calm
   c. Tired - energetic
   d. Bored – excited
   e. Upset – content
(Ask no more than 2 pairs in a given survey)

B4. How would you rate…
   a. Your trip to the bus stop? (Rate on scale of 1-5: 1 = very inconvenient, 5 = very convenient)
   b. Bus information at the stop? (Rate on scale of 1-5: 1 = very poor, 5 = very informative)
   c. Bus stop condition? (Rate on scale of 1-5: 1 = very inadequate, 5 = very proper)
   d. Crowding at the stop? (Rate on scale of 1-5: 1 = very crowded, 5 = very spacious)

Onboard questionnaire

C0. Are you travelling by bus from (bus stop name) now?
   Did you travel by bus from (bus stop name) at around (time)?
   - Yes
   - No, I passed by that bus stop
   - No, I waited at that stop, but left on foot
   - No, I waited at that stop, but left by taxi/car
   - No, I did not go to that stop at all
   (If No, go to statement "Sorry for the inconvenience! Click Submit to exit.)

C1. What is the purpose of this trip?
   What was the purpose of that trip?
   (Choose 1 answer only)
   - Work
   - School
   - Shopping/Eating
   - Leisure activities
   - Errands
   - Going home
   - Other

C2. How satisfied are you with your wait for that bus?
   How satisfied were you with your wait for that bus?
   (Rate on scale of 1-5: 1 = very dissatisfied, 3 = neutral, 5=very satisfied)

   (If answer is 1, 2 or 3, go to Question C2. Otherwise, skip to Question C3)
C3. Tell us why you are less than satisfied with the wait? (Select all that apply)
   Tell us why you were less than satisfied with the wait? (Select all that apply)
   - Wait time was too long
   - Inaccurate bus arrival info
   - Bus too crowded to board
   - Unsatisfactory bus stop condition
   - Other (comment)
   - Rather not say

C4. How crowded is the bus?
How crowded was the bus?
(Participant to select an answer from the multiple choices, in the form of 5 pictures visualizing varying crowding levels on a bus)

C5. How comfortable do you feel on the bus now?
How comfortable did you feel on the bus?
(Rate on scale of 1-5: 1 = very uncomfortable, 3 = neutral, 5=very comfortable)
(If answer is 1, 2 or 3, go to Question C5. Otherwise, skip to Question C6)

C6. Tell us why you are less than satisfied with the comfort? (Check all that apply)
Tell us why you were less than satisfied with the comfort? (Check all that apply)
- Cannot get a seat on the bus
- Too many people on the bus
- Ride is not smooth
- Temperature is too hot/cold
- Other (comment)
- Rather not say

C7. How satisfied are you with the service provided by the driver?
How satisfied were you with the service provided by the driver?
(Rate on scale of 1-5: 1 = very dissatisfied, 3 = neutral, 5=very satisfied. + "Don't know")

Post-alighting questionnaire

D0. Did you travel by bus around (time) today from (bus stop name)?
- Yes
- No, I passed by that bus stop
- No, I waited at that stop, but left on foot
- No, I waited at that stop, but left by taxi/car
- No, I did not go to that stop at all

(If No, go to statement "Sorry for the inconvenience! Click Submit to exit.)

D1. What is the purpose of this trip?
What was the purpose of that trip?
(Choose 1 answer only)
- Work
- School
- Shopping/Eating
- Leisure activities
- Errands
- Going home
- Other

D2. How satisfied are you with the bus travel time?
(Rate on scale of 1-5: 1 = very dissatisfied, 3 = neutral, 5=very satisfied)

D3. How long do you think this trip took?
_____ minutes

D4. How convenient was it travelling by this bus?
(Rate on scale of 1-5: 1 = very dissatisfied, 3 = neutral, 5=very satisfied)

D5. Which of the following activities did you do while travelling? (check all that applies)
- Idling/thinking
- Reading
- Emailing/texting
- Playing games
- Talking via phone
- Talking in-person
- Listening to music
- Sleeping
- None of the above

D6. Overall, how satisfied are you with this bus experience?
(Rate on scale of 1-5: 1 = very dissatisfied, 3 = neutral, 5=very satisfied)

D7. My experience taking this bus made me happier.
(Rate on scale of 1-5: 1 = strongly disagree, 3 = neutral, 5 = strongly agree)

End-of-day questionnaire

F0. Were you travelling by bus at around (time) today?

F1. Overall, how satisfied are you with this bus trip?
(Rate on scale of 1-5: 1 = very satisfied, 3 = neutral, 5=very dissatisfied)

F2. How did the bus experience compare to your expectations?
(Choose from responses: "much better than expected", "better than expected", "as expected", "worse than expected", "much better than expected")

F3. Would you take this trip again with the same choice of bus routes?
(Can only select 1 answer)
- Yes
- Yes, but only at a different time of day
- Yes, I have no choice.
- Maybe
- No

F4. Travelling helped me accomplish my goals today.
(Rate on scale of 1-5: 1 = strongly disagree, 3 = neutral, 5 = strongly agree)

APPENDIX B

Boston pilot questionnaires

Test in black appears on the screen. Text in red denote questionnaire logic and flow control.

Entrance Survey

A1. Which MBTA services do you use regularly? (Check all that apply)
- ☐ The Silver Line
- ☐ Bus (excluding Silver Line)
- ☐ Subway
- ☐ Commuter Rail
- ☐ Commuter Ferry
- ☐ The RIDE
- ☐ None

(If None is select, de-select any choices that are already selected)

(If A1 = none) Thank you for your interest. This app is currently being piloted for regular MBTA users.

(If A1 SL answer is checked) A2. Approximately how often do you take the Silver Line?
- o Less than once per week
- o One or two times per week
- o Three or more times per week

(If A1 SL answer is not checked) Thank you for your interest. This app is currently being piloted for Silver Line riders. Please enter your email if you would like to be notified if the pilot is extended to other modes. (Comment field for email submission)

A3. Do you have a private automobile available for your use?
- o Yes, always
- o Sometimes (shared with household members)
- o Never

A4. Do you own a bike or a bike-share (Hubway) membership?
- o Yes
- o No

A5. How do you generally plan Silver Line trips? (Check all that apply)
- o Walk up
- o Check schedule

      o   Check real-time information

A6. How satisfied are you with your MBTA bus (including Silver Line) experiences?
(Dropdown: Rate on scale of 1-7: 1=very dissatisfied, 7 = very satisfied)

A7. What is your gender?
- o   Woman
- o   Man
- o   Another (please specify) _____
- o   Prefer not to say

A8. What is your age?
- o   Under 18
- o   18 to 21
- o   22 to 34
- o   35 to 44
- o   45 to 64
- o   65 or over
- o   Prefer not to say

A9. How do you self-identify by race? (Check all that apply)
- ☐   American Indian or Alaska Native
- ☐   Asian
- ☐   Black or African American
- ☐   Native Hawaiian or other Pacific Islander
- ☐   White
- ☐   Other (please specify) _____
- ☐   Prefer not to say

A10. Are you Hispanic or Latino/Latina?
- o   Yes
- o   No
- o   Prefer not to say

A11. What is your current household income?
- o   Less than $14,500
- o $14,500 to $28,999
- o $29,000 to $43,499
- o $43,500 to $75,999
- o $76,000 to $108,499
- o $108,500 to $151,999
- o $152,000 or more
- o   Prefer not to say

Real-time onboard survey - waiting but not on service

(If stop beacon flags a person for some time – 3 minutes? – but no bus beacon subsequently flags them) B0. Were you at a Silver Line stop at [location] earlier today?
- o   Yes, I was waiting for the bus
- o   No, I was just passing by
- o   No, I wasn't

(If No) "Thanks for answering! That's all we need today. Click Submit to exit."

(If Yes) B1. Did you get on the Silver Line?
- o   Yes
- o   No, I left

(If "No, I left" was chosen") B2. Did you make the trip using another mode?
- o   Took a different bus (not the Silver Line)
- o   Took the subway
- o   Took a private or corporate shuttle/bus
- o   Walked
- o   Rode a bicycle
- o   Took a taxi/Uber/Lyft
- o   Took another mode
- o   I did not make the trip

Real-time onboard questionnaire

C0. Are you on the Silver Line now?
Were you travelling on the Silver Line earlier today?
- o   Yes
- o   No

 (If No, go to statement "Thanks for answering! That's all we need today. Click Submit to exit.")

C1. What is the purpose of your trip?
What was the purpose of that trip?
(Choose 1 answer only)
- o   Work or work-related meetings
- o   School or educational activity
- o   Personal errand (bank, daycare, etc)
- o   Medical or other appointment
- o   Social/recreational activity
- o   Civic/volunteer activity
- o   Travel (Airport/Amtrak/intercity bus)
- o   Other (comment)

C2. How satisfied are/were you with the following? (each attribute below rated on a scale of 1 to 7 from extremely unsatisfied to extremely satisfied)
  1. Cleanliness of bus stop
  2. Wait time at bus stop
  3. Crowdedness on bus
  4. Comfort on bus
  5. Driver's service

B3. Did any of the following affect your satisfaction? (Check all that apply)
  ☐ No seat on the bus
  ☐ Bus was crowded
  ☐ Ride not smooth
  ☐ Too hot
  ☐ Too cold
  ☐ Bus not clean
  ☐ Feel unsafe
  ☐ Other
  ☐ None

End-of-trip questionnaire

D1. Please rate your Silver Line trip made at [time]
(Rate on scale of 1-7: 1 = very dissatisfied, 7=very satisfied)

  • Overall
  • Waiting at bus stop
  • Crowdedness on bus
  • Comfort on bus
  • Driver's attitude

(Check box) I didn't make the trip

Additional comments _____