

Qualitative and Quantitative Approaches for Studying Transit Stations

By

Gregory Spitz
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
802 295-4999
gspitz@rsginc.com

Elizabeth Greene
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
802 295-4999
egreene@rsginc.com

Tom Adler
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
802 295-4999
tadler@rsginc.com

Rhona Dallison
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
802 295-4999
rdallison@rsginc.com

ABSTRACT

The paper describes the qualitative and quantitative research methods that were used to assess current customer perceptions of, and improvement priorities for, New York City Transit's subway stations. The qualitative research used a mobile focus group facility which was moved to each of the stations being studied. The quantitative research used a survey instrument that included trade-off scenarios structured to support maximum difference scaling conjoint (MaxDiff) to evaluate transit customers' preferences for proposed station improvements.

The focus groups were designed to include both a guided station tour that allowed participants to observe key station features followed immediately by a traditional focus group discussion of those features. The mobile focus group facility provided a very convenient and cost-effective venue for the focus group discussion. Participants were not required to travel beyond the station that was being evaluated, which in most cases was either their home station or along their normal trip route. The quantitative research included a set of MaxDiff scenarios as part of a short intercept survey. The responses to these scenarios supported an analysis of customer preferences among 11 types of station improvements.

INTRODUCTION

Stations can play a significant role in the customer experience of a transit system, as they serve as riders' primary access, egress, and transfer points. Therefore, evaluating and understanding transit customer preferences regarding stations is an important component of transit research. This paper describes qualitative and quantitative approaches that were used to evaluate stations and assess customer perceptions of stations for the Metropolitan Transportation Authority and New York City Transit (MTA/NYCT). The qualitative research used a mobile focus group facility which was moved to each of the stations being studied. The quantitative research used a survey instrument that included trade-off scenarios structured to support maximum difference scaling conjoint (MaxDiff) to evaluate transit customers' preferences for proposed station improvements.

This paper describes the approach used for the MTA/NYCT Station Evaluation Study.

TYPES OF STATIONS EVALUATED

Customer evaluation of transit stations can occur at different stages: pre-construction, pre-renovation, or post-renovation. For the MTA/NYCT Station Evaluation Study conducted in 2005 and 2006 that is discussed in this paper, stations that had been previously renovated were evaluated in order to understand whether those renovations served the needs of transit customers and to determine what other renovations customers might appreciate. The goal of the study was to help plan future renovations to be undertaken for NYCT subway stations and to understand the customers' impressions of the most important improvements that can be made to a station.

Mobile focus groups and MaxDiff analysis as described in this paper can be used at any stage of station evaluation, including for stations scheduled for renovation and stations yet to be constructed.

STUDY BACKGROUND

MTA/New York City Transit is the largest transit service provider in the United States. The New York City Subway carries 4.8 million trips on the average weekday and the infrastructure represents one of the most extensive subway systems in the world. The Subway has 468 stations; more than any other system in the world. The first Manhattan station opened in 1904, followed by extensions to the Bronx in 1905, Brooklyn in 1908, and Queens in 1915. Now more than 100 years old, the New York City Subway is a vast system which requires significant ongoing station rehabilitation efforts.

This MTA/NYCT Station Evaluation Study was undertaken to evaluate ten specific stations in four of the five New York City boroughs. All ten of these stations had been recently renovated (Figure 1).

Figure 1. Ten New York City Subway Study Stations

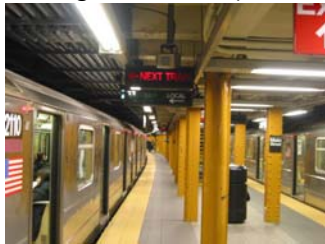
125th Street (Broadway/7th Avenue Line, Manhattan)



DeKalb Avenue (Fourth Ave. – Brighton Lines)



Flushing/Main Street (Queens)



Stillwell Avenue/Coney Island (Brooklyn)



161st Street Yankee Stadium (Bronx)



Delancey Street/Essex Street (Manhattan)



Gates Avenue (Jamaica Line)



Atlantic Avenue/Pacific Street (Brooklyn)



East Tremont Ave./West Farms Square (White Plains Rd. Line)



Roosevelt Avenue/74th Street (Queens Blvd. – Flushing Lines)



The qualitative analysis comprised two focus groups per subway station; the quantitative analysis surveyed at least 300 regular station users per subway station. As with many market research studies, the qualitative element of the study informed the development of the quantitative part of the study. Specifically, the qualitative study informed the development of a list of station improvements (such as better lighting, wider stairwells, improved cleaning and maintenance, etc.) to test in the quantitative part of the study. The focus group assisted in identifying which issues were important to users and the quantitative research was designed to understand the relative importance of each improvement. Additionally, the qualitative analysis provided a better understanding of customers' assessment of their rehabilitated stations through their own words in focus groups, allowing for better interpretation of subsequent quantitative results.

Each station was evaluated separately and an aggregate analysis of all stations was also conducted. One significant challenge presented by this study was the creation of a single list of improvements that would be valuable for any future station rehabilitation. Stations vary dramatically in structure (e.g., belowground versus elevated), configuration (e.g., large transfer station versus a small local station), location (e.g., up-and-coming neighborhoods on the Lower East Side versus established neighborhoods in the Bronx, Brooklyn, and Harlem), and the needs of riders who use each station also vary.

Because of the many differences between stations and their geographical distribution, creative methods were required to conduct the MTA/NYCT Station Evaluation Study.

MOBILE FOCUS GROUPS

A significant challenge was presented by the geographical spread of the stations in the study, which were literally “all over the map” of New York City (Figure 2). A key element of the qualitative part of the study was a guided walk-through tour of the station under study, so in the interest of time and convenience, the focus groups (two for each station) needed be held as close to the station as possible. There are a limited number of focus group facilities in the Bronx, Brooklyn, and Queens; and the larger pool of Midtown Manhattan focus group facilities was not convenient to the stations being studied. Travel between the study station and a traditional focus group facility would have necessitated a focus group commitment of four hours of participants’ time or more, likely creating non-response bias in the group as only subway users with significant free time could have participated. These challenges were overcome by renting a special-events bus to use as a mobile focus group facility (Table 1 and Figure 3).

Figure 2. Map of Study Stations



Table 1. Comparison of Traditional and Mobile Focus Group Facilities

	Traditional Focus Group Facility	Mobile Focus Group Facility
Process	Meet participants at station Conduct station walk-through Travel to focus group facility in Manhattan by group using subway or a van Conduct focus group Participants travel home from Manhattan	Meet participants at station Conduct station walk-through Walk to luxury bus outside station Conduct focus group Participants leave focus group and many simply walk home (if the stop is their home station)
Time	Approximately 4 hours per focus group	Approximately 2 hours per focus group
Cost	Focus group facility fee Higher incentive expense to participants Rental van cost Extra staff or extra days to complete two focus groups at each station	Special-events bus rental fee Lower incentive expense to participants
Constraints	Only those with 4 or more hours of available time can participate Travel during congested times getting to/from focus group facility	Only those with 2 hours of available time can participate Created more work for the research team (audio/video, food, respondent compensation)

In addition to introducing time constraints, traveling to and using a traditional focus group facility in Manhattan would have resulted in higher project costs. The bus rental was equivalent to focus group facility quotes for the same number of groups, and transportation costs to move respondents from the station to a fixed facility were eliminated using this method.

The study required that two focus groups be held for each station evaluation. A representative sample of station users was recruited in person at each station a day or two prior to the focus groups (e.g., people who had used the station for years as well as those who had started using it recently). These participants were met at a predetermined place in each station at the start of the focus group sessions and led on the walk-through tour, so that they could familiarize themselves with the various details of the station rehabilitation, some of which they might not have been aware of (e.g., pigeon control devices or non-slip flooring).

After the walk-through, respondents were brought back to street level to the waiting mobile focus group facility that was parked within half a block of the station. Participants were excited to get on the bus, which was similar to buses that celebrities or musicians travel in (in fact, the bus provider claimed the Rolling Stones and other celebrities as their clients).

Figure 3. Special Events Bus for Use as Mobile Focus Group Facility



Upon boarding the bus, respondents were provided food from a small table set up in the middle of the bus. The moderator then directed participants to their seats in the back section of the bus. The small quarters created a more casual atmosphere than traditional focus group facilities, but it worked well coming out of the activity and hustle of the subway station.

The bus had enough space to provide a very comfortable focus group for eight people. Initial groups had ten participants, which were a bit crowded, leading to a decision to reduce the groups to eight participants. Despite space constraints, there was room enough for six clients to view the groups as they occurred. The groups were conducted in the back half of the bus behind a curtain (no one-way mirror) (Figure 4).

Figure 4. Interior Seating for Participants of Mobile Focus Group



The groups were videotaped and the video feed sent to a monitor in the client area of the bus. The groups were also voice recorded, and the sound feed was brought to a speaker system in the client seating area. Having the audio-visual technologies in plain view and the clients in close proximity to the groups made client presence quite conspicuous in the mobile focus group facility, unlike in a typical focus group facility. However, once the focus groups were underway, respondents seemed to have no reservations about sharing their opinions and were quite relaxed.

Parking was one challenge, especially in Manhattan and Flushing, Queens, locations. The contingency plan was to drive the bus around if parking was tight or unavailable. This only occurred once and was ultimately resolved by a helpful New York City Police Department traffic officer who found a good parking location for the bus.

The mobile focus group facility very successfully met the needs of this station project. It did have some drawbacks, but these were far outweighed by the significant benefits provided.

MAXDIFF ANALYSIS

Many traditional studies have customers prioritize their needs through ratings and rankings (e.g., “On a scale of 0 to 10, with 10 being best, how important is it for this station to have escalators?”). This type of analysis is quite common and for a station evaluation study would be useful for understanding how customers feel about existing station characteristics. However, it does not help understand customer desires for future improvements or the amount by which certain characteristics are valued in comparison to others.

For example, nearly every participant in the MTA/NYCT Station Evaluation Study wanted to add or increase the number of escalators at their station. The more significant question, however, is what they would trade-off for such an improvement? Would they be willing to see a fare increase if the result were more escalators? Additionally, what value do escalators have when compared to other potential station improvements? Would respondents prefer more escalators to better lighting, surveillance cameras, or artwork in the station? Not

only do participants need to communicate the value an escalator has for them, they also must value it compared to other potential improvements that are being considered for the station (Louviere).

The trade-off methods of conjoint analysis and choice modeling were developed to deal with the limitations of traditional rating methods. The word “conjoint” is used because these analysis techniques force respondents to evaluate a number of attributes conjointly, or in the same context. Conjoint analysis and one commonly used variant, discrete choice modeling, force respondents into choices among attribute levels. The resulting choices are modeled, typically using the logit form, so that coefficients are estimated to determine the relative value of various attribute levels. From these coefficients, different future scenarios can be simulated so that decisions on the best bundle of attributes for a particular product or service can be made (Adler). Trade-off methods produce more robust, accurate measures of customer priorities and minimize or eliminate the scale and cultural biases that plague ratings methods (Louviere, Cohen).

The conjoint technique used for the MTA/NYCT Station Evaluation Study was maximum difference scaling conjoint (MaxDiff, and also known as best-worst conjoint) analysis. It allowed for evaluation of various station features and an understanding of how much value customers placed on certain features relative to others. A past study of Chicago’s rapid rail stations used a combination of adaptive conjoint analysis (ACA) and discrete choice conjoint to evaluate potential improvements (Adler), but MaxDiff was used for the following reasons:

- Simple for the respondent to understand and evaluate, even in a busy subway station
- Can be administered using in-person interviews with responses recorded on a paper form on a clip board (no computer necessary)
- Only requires respondents to answer a few comparison questions (versus other conjoint techniques that generally require more questions)
- Can be used to evaluate a relatively large number of items and is simpler than other conjoint methods to set up

Applying other types of conjoint analysis would have been very difficult for this study as respondents were intercepted directly in the subway station and had limited time to answer the questionnaire (as their train typically arrived within three to five minutes after being intercepted). While the study could have been designed with computer interview stations in the subway station, this clearly would have been more time-consuming for respondents and a more expensive approach (Table 2). A paper based MaxDiff implementation was believed to be the most appropriate solution for this particular project.

Table 2. Comparisons between Popular Conjoint Methods

	MaxDiff	Choice Model	Adaptive Conjoint Analysis (ACA)
Can use paper to conduct the experiment?	Yes, easy	Yes, but difficult	No
Requires rating and ranking questions prior to the trade-off experiment?	No	No	Yes
Trade-off experiment is fairly simple?	Yes, very simple	No	Yes, but not as simple as MaxDiff

MaxDiff made the trade-off process much easier and, like the mobile focus group facility, was an appropriate choice for this particular application. Respondents were asked to answer four simple trade-off scenarios in order to test a set of general station statements or attributes. MaxDiff works by compiling a subset of the statements to be tested into a “scenario,” in this case four features, and asking respondents to select a “most important” and “least important” statement for each scenario. For this study, each scenario contained four statements, and each statement therefore ran in multiple scenarios. The exercise results in a utility score for each statement, which can then be used to show the relative importance of each statement versus all other statements. For this study, each respondent was shown four scenarios (each with four statements). An example scenario is below (Table 3).

Table 3. MaxDiff Example, Question 1 of 4

Most Important	Of the following four items, which one is most important to you? Which one is least important?	Least Important
<input type="checkbox"/>	Adding surveillance cameras and emergency call buttons	<input type="checkbox"/>
<input type="checkbox"/>	Wider passages and stairways	<input type="checkbox"/>
<input type="checkbox"/>	Art within the stations such as mosaics or stained glass	<input type="checkbox"/>
<input type="checkbox"/>	Clarity of announcements from public address speakers	<input type="checkbox"/>

Respondents were asked to trade off the following attributes (statements), determined in consultation with MTA/NYCT and the research consultant, as these attributes were deemed to be important measures to meet the objectives of this particular study:

- Availability of subway maps on station platforms and walls
- Clarity of announcements from public address speakers
- Enhanced lighting on station platforms
- Wider passages and stairways
- Improved visibility and open sightlines in the station
- Station art, such as mosaics or stained glass
- Using high quality/attractive materials throughout the station, such as granite
- Adding surveillance cameras and emergency call buttons
- Escalators in the station
- Improved cleaning and maintenance
- 5 minute travel time savings on your subway trip

The way attributes were selected for this study included significant input from the focus groups (e.g., improved cleaning and maintenance was added to the list as it came up in many groups), input from MTA/NYCT as to what was realistic for them to provide in future stations (as not everything customers might state they want in the focus groups is possible, for a variety of reasons), and input from the consultant as to attributes that would help further quantify the analysis (e.g., the addition of a travel time savings attribute allows the analysis to put a value on each attribute in terms of travel time).

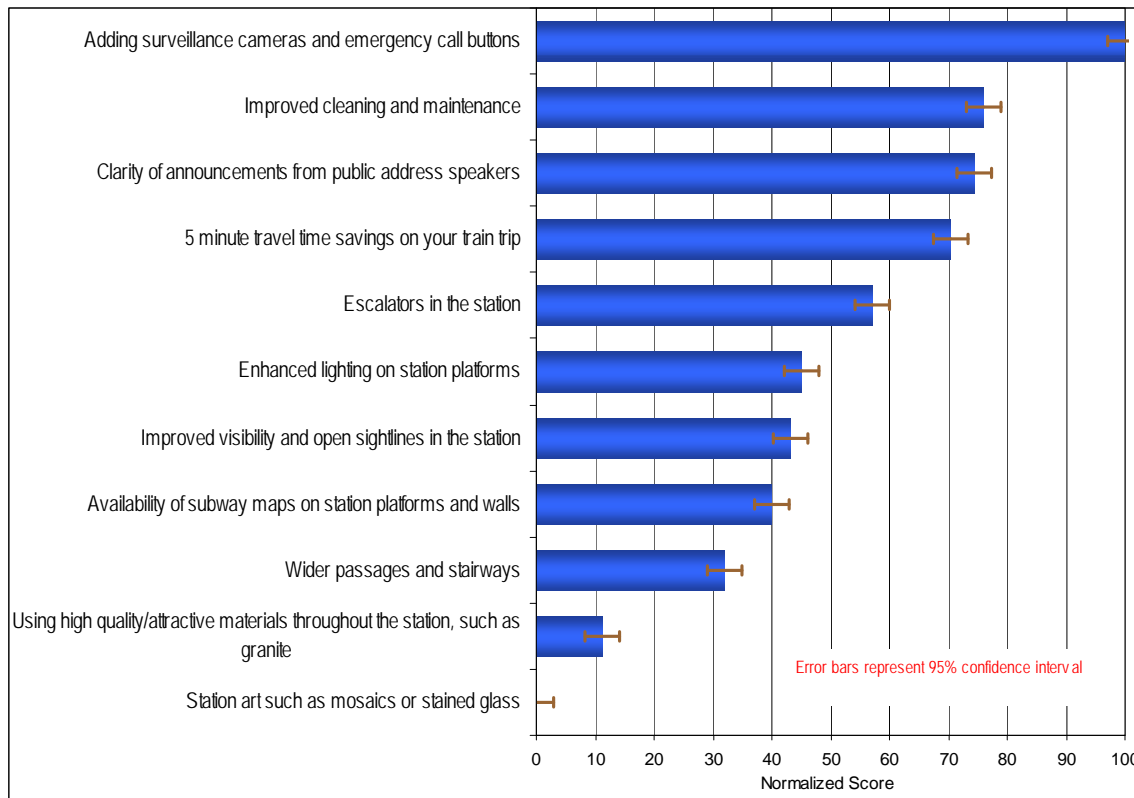
It is important to note that these attributes are stated in relatively general terms. For this study, more general attributes were intentionally selected because this was the level of detail necessary to understand what subway station attributes were of greatest importance to subway customers. For example, it was not desired or even possible to come up with a specific definition of the “enhanced lighting” attribute because a variety of stations were involved in the study and specific lighting improvements differ by station. What was important to researchers was the understanding of how important enhancing lighting is for a subway station, and by how much, relative to other general improvements. This research allows MTA/NYCT to then know what is most important in a renovation. Further and more specified research can be conducted as needed to understand exactly what type of lighting enhancements, for example, are best.

The results of the MaxDiff analysis can be seen in the following chart, which shows the relative importance of each of the attributes listed above (Figure 5). The MaxDiff analysis can be segmented by different users so long as there is sufficient sample to estimate a model (300 respondents were used for each segment in this study). The chart below shows some generic MaxDiff results to show the reader how this method can be analyzed. Specific results are not available for this paper as the study findings have not yet been made public. This is not due to any question of the validity of the results or the research methods used to obtain them. It is simply due to the review policies of MTA/NYCT. This study has only recently concluded and the internal

review process is still ongoing, so results cannot be made public until that process concludes. So, the results for this paper are shown at an aggregate level here.

In order to make the chart easier to interpret, the scale has been normalized to be between 0 and 100. The relative importance is preserved in the rescaling, as is the error around each attribute (represented by the red error bars for 95% confidence).

Figure 5. MaxDiff Results (shown for illustration purposes only, not based on real data)



While these results are generic, their order of preference and magnitudes are real. One interesting result of the study involves safety and security features of subway stations. A standard performance rating analysis/quadrant chart was also conducted for this study, which was conducted in addition to the conjoint analysis, to understand current performance at each station. This standard analysis showed mostly above-average performance scores regarding personal security and safety. However, the MaxDiff results show that the greatest utility would be derived from installation of surveillance cameras and call buttons in the stations.

These findings might appear to be contradictions, but they represent an important example of what the MaxDiff analysis can tell the analyst that standard analysis cannot: relative to all other improvements, surveillance cameras and call buttons still are the most important new improvement for respondents, even though respondents have higher than average satisfaction/performance ratings for personal security and safety attributes. By placing attributes relative to each other (i.e., they are in the context of a real trade-off), MaxDiff and other conjoint techniques identify the relative importance of the features that are tested. These different conclusions are also likely to be more valid than traditional analyses because they are done by forcing trade-offs and are therefore more robust and not prone to issue of scaling bias. An important problem in traditional rating analysis is that most everything appears important and scores tend to cluster (Cohen). Conjoint methods are a way to solve these scaling issues so that the most important attributes stand out.

CONCLUSIONS

Market research to evaluate station attributes presents a number of challenges. Mobile focus group facilities for qualitative research and the MaxDiff method for quantitative research are useful approaches for addressing some of the most difficult challenges presented by station research.

As seen in this paper, both methods created significant benefits for this study and did so within the project's budget constraints, while meeting the project's objectives. As in most projects, one of the most challenging tasks is to find the most appropriate techniques to conduct a project under various limitations such as time and budget. For this study, mobile focus groups and MaxDiff methods were two approaches that worked well in achieving that balance.

REFERENCES

1. Adler, T., F. Falzarano and R. Hazlett, "Quantifying the Value of Transit Station and Access Improvements for Chicago's Rapid Transit System," presented at the Transportation Research Board's 80th Annual Meeting, January 2001
2. Cohen, Steven H. 2003. "Maximum Difference Scaling: Improved Measures of Importance and Preference for Segmentation." Sequin, Washington: Sawtooth Software, Inc.
3. Cohen, S.H. and Neira, L. 2003. "Measuring Preference for Product Benefits Across Countries: Overcoming Scale Usage Bias With Maximum Difference Scaling." ESOMAR 2003 Latin America Conference Proceedings. Amsterdam: The Netherlands.
4. Louviere, Jordan J. and Islam, Towhidul. 2004. "A Comparison of Importance Weights/Measures Derived from Choice-Based Conjoint, Constant Sum Scales, and Best-Worst Scaling." CenSoC Working Paper No. 04-003.