Timing Is Everything

A Field Study of Subway Service Reliability

Michael T. Doyle
Associate Director

August 2000

New York City Transit Riders Council
347 Madison Avenue, New York, NY 10017
(212) 878-7087 • www.pcac.org

Acknowledgements

The author would like to thank the members of the NYCTRC for their invaluable efforts in performing field observations.

The author is grateful for special assistance provided in the field by Executive Director Beverly Dolinsky, Transportation Planner Joshua Schank, and Jonathan Sigall of the NYCTRC staff.

The author would also like to thank NYC Transit Executive Vice-President Barbara Spencer, John Tucker, NYCT vice-president - Operations Planning, and Deborah Hall-Moore, NYCT assistant director - Government and Community Relations, for providing prompt and open access to detailed timetable information to use as a basis for comparison in this study.

Finally, the author would like to acknowledge analytic assistance provided by Straphangers Campaign Staff Attorney Gene Russianoff and technical assistance provided by former PCAC Associate Director Alan Foster.

___

Cover photo kind permission David Pirmann/New York City Subway Resources.

Table of Contents

List of Tables

Executive Summary

NYC Transit publishes schedules for all subway lines. However, trains are generally spaced five to ten minutes apart during most hours of the day, eliminating the need for customers to refer to schedules as long as these waiting times between trains, or **headways**, are reliably maintained. From a subway rider's perspective, a convenient, well-performing subway line is one with headways that can be depended upon to not differ widely from day to day. Thus, measuring the **reliability** of headways is a good way to gauge how useful a line is to riders.

Because service is scheduled at such relatively short headways for most hours of the day, trains which keep close to their schedules should by definition operate reliably. However, NYC Transit does not record on-time performance for stations along the length of a line during day and evening hours, thus there is no way to determine the reliability of headways from on-time performance data. Instead, an independent measure of reliability is necessary.

Recognizing the importance of reliable headways and the lack of data about them, in 1992 and 1996 the New York City Transit Riders Council performed field studies to monitor subway service reliability.[1] Both studies showed numerous, unscheduled gaps in service, although reliability did improve overall between 1992 and 1996.

Much has changed since our last reliability study. The system-wide rollout of MetroCard and the implementation of fare discounts and free bus-to-subway transfers have led to spectacular ridership growth and a gradual restoration of service eliminated in 1996. However, although ridership stands at a 30-year high, service levels do not. As more riders compete for limited space, the Transit Riders Council is concerned that unacceptable gaps in scheduled service may be on the rise. Therefore, we decided to perform a new service reliability study for 2000.

Our 2000 study differs somewhat from our previous studies. In order to allow our results to be directly comparable to the reliability data reported by NYC Transit, itself, this year we have used NYC Transit's own measure of reliability, called **service regularity**[2], to analyze our data. This measure determines the proportion of headways which fall within an acceptable range of the length they are scheduled to be. Because we have changed the way we analyzed our data for the 2000 report, our results are not directly comparable with the results of our previous studies.[3] However, they can be compared with NYC Transit's own reliability data from 1996 and 2000.

---

[1] Jeffrey F. Vernick, (September 1992), ***New York City Transit Authority Advisory Council Service Monitoring Project***, and Alan H. Foster, (July 1996), ***New York City Transit Riders Council 1996 Bus & Subway Service Evaluation***. Available upon request.

[2] For a complete definition of NYC Transit's service-regularity measure, please see **Methodology**.

[3] For complete details of the analysis methods used in the Council's 1992 and 1996 reports, please see **Why We Did the Field Study**.

Findings and Recommendations

The best performing trains in our survey for combined morning peak and midday service were the **C** and **Franklin Avenue Shuttle**, with reliability scores of 100%. The worst performers were the **4** and **5** lines, with reliability scores of just 62 and 56 percent, respectively.

Overall, our results indicate that subway service is reliable for most lines, and is improved for most lines from 1996. Our reliability figures for the B-Division, or lettered trains, were markedly good, especially for midday hours. Our results also indicate that NYC Transit does a good job of maintaining reliable service during planned weekday daytime diversions in service, such as on the **J**/**Z** line during midday. Further, the agency appears to provide a quick response to unexpected major delays.

However, reliability on the **1**/**9**, **4**, **5**, **7**, and **M** lines appears to have worsened since 1996. Furthermore, nine lines exhibited reliability figures below 80 percent, including almost all A-Division IRT (numbered) lines. This means that customers who ride these lines–the **1**/**9**, **4**, **5**, **6**, **7**, **B**, **J**/**Z**, **M**, and **R**–can expect more than one out of every five trains to be delayed. For riders who regularly travel to work using these lines, this means that delays can be expected more than twice a week. This frequency of delays is unacceptable and special attention should be paid to these lines to improve reliability.

Of even greater concern, some of these poorly performing trains–the **4**, **5**, **7**, **J**/**Z**, **M**, and **R**–actually were less reliable on a day-to-day basis than lines which experienced planned service diversions and unplanned delays. This represents a serious problem, because these are lines which many people take or to which they transfer. The Council suspects that systemic problems exist on these lines which are detracting from reliability. These problems should be identified and addressed.

Further, a comparison of our results with NYC Transit's own reliability data suggests that the agency's figures may not adequately represent reliability during peak hours because the agency reports only a single, overall figure for reliability. In the future, NYC Transit should also report a breakdown of their reliability data by time of day, as we have done in this report, to better indicate reliability for peak and off-peak hours.

Finally, we disagree with a proposal by NYC Transit to replace its current service-regularity measure with a new measure of reliability called **wait assessment**.[4] Subway headways are scheduled with reference to the specific ridership needs for each line at specific times of day. The existing service-regularity measure is calculated in proportion to these scheduled headways. However, wait assessment ignores the specific service needs of individual lines, instead relying on a universal three-minute reliability standard. Because of this, the proposed wait-assessment measure cannot accurately represent the impacts on riders of all types of delays–especially delays affecting lines with short headways and delays caused by early trains. Due to this lack of precision, we believe wait assessment should not be adopted by NYC Transit as its new measure of service reliability.

[4] For a definition and full discussion of wait assessment, please see **Appendix A**.

Why We Did the Field Study

NYC Transit publishes schedules for all subway lines. However, trains are generally spaced five to ten minutes apart during most hours of the day, eliminating the need for customers to refer to schedules as long as these waiting times between trains, or **headways**, are reliably maintained. From a subway rider's perspective, a convenient, well-performing subway line is one with headways that can be depended upon to not differ widely from day to day. Thus, measuring the **reliability** of headways is a good way to gauge how useful a line is to riders.

Because service is scheduled at such relatively short headways for most hours of the day, trains which keep close to their schedules should by definition operate reliably. However, NYC Transit does not record on-time performance for stations along the length of a line during day and evening hours, thus there is no way to determine the reliability of headways from on-time performance data. Instead, an independent measure of reliability is necessary.

Recognizing the importance of reliable headways and the lack of data about them, in 1992 the New York City Transit Riders Council undertook the first of what has become a regular series of field studies to monitor subway service reliability.[5] The 1992 study utilized a new, independent measure, termed **service regularity**, developed by the Office of the MTA Inspector General. As conceived by the Inspector General, this measure determines the proportion of evenly spaced headways along a line. Our results found service to indeed be uneven and erratic, with frequent lengthy gaps afflicting many lines.

Service regularity is a measure that can be used to determine the reliability of subway service.

We repeated the study in 1996, again using the Inspector General's measure, after major, budget-related service cutbacks prompted renewed public concerns about reliability. Our 1996 results showed a slight improvement in reliability, but still recorded numerous, unscheduled gaps in service. By 1996, NYC Transit had also adopted its own, separate measure of reliability, which it also called service regularity.[6] Unlike the Inspector General's measure, NYC Transit's service-regularity measure determines the proportion of headways which fall within an acceptable range of the length they are scheduled to be. In some instances, our 1996 results indicated worse reliability than did NYC Transit's own figures.

Much has changed since our last reliability study. The system-wide rollout of MetroCard and the implementation of fare discounts and free bus-to-subway transfers have led to

[5] Jeffrey F. Vernick, (September 1992), ***New York City Transit Authority Advisory Council Service Monitoring Project***, and Alan H. Foster, (July 1996), ***New York City Transit Riders Council 1996 Bus & Subway Service Evaluation***. Available upon request.

[6] For a complete definition of NYC Transit's service-regularity measure, please see **Methodology**.

spectacular ridership growth and a gradual restoration of service eliminated in 1996. However, although ridership stands at a 30-year high, service levels do not. As more riders compete for limited space, the Transit Riders Council is concerned that unacceptable gaps in scheduled service may be on the rise. Therefore, we decided to perform a new service reliability study for 2000.

Our 2000 study differs somewhat from our previous studies. In order to allow our results to be directly comparable to the reliability data reported by NYC Transit, itself, this year we have used NYC Transit's service-regularity measure to analyze our data instead of the Inspector General's measure. For this reason, our results are not directly comparable with the results of our previous studies. However, they can be compared with NYC Transit's own reliability data from 1996 and 2000, and we do so in our findings.

We have also included a discussion of the differences between NYC Transit's current service-regularity measure and a new measure of reliability proposed by the agency called **wait assessment**.[7] In the appendix, we compare the two measures and discuss why wait assessment may not be a good indicator of service reliability as experienced by riders.

[7] For a definition of wait assessment, please see **Appendix A**.

Methodology

The Council observed service for all regular subway lines and the **Franklin Avenue Shuttle** (FAS). Observations were made between 8:00 a.m. and 10:00 a.m. and between Noon and 1:00 p.m. on each of two separate, non-Holiday weekdays. These hours were chosen because they represent some of the busiest peak and off-peak hours in the subway system. Observations were made in one direction, only.

Days, directions, and station locations for observations were chosen at random. However, the first and last two stops on all lines except the **Franklin Avenue Shuttle**[8] were excluded from observation in order to concentrate on service regularity as riders experience it along the main body of a given line. (Please see **Table One**, below, for a summary of observation dates and locations.)

Observers were assigned a specific subway line or lines and were instructed to record the arrival and departure times for all trains during the observation periods noted above. In total, approximately 1,300 train movements were recorded. Arrival times were recorded as the moment train doors opened to allow customers to board and disembark. Departure times were recorded as the moment trains began to exit stations. Both arrival and departure times were recorded in order to help better identify delays.

Some field observations were made at stations for which exact NYC Transit timetable data was not available. For these stations, NYC Transit provided the Council with average running times from the nearest station used as a timepoint in the official timetables so that the Council could correctly interpolate departure schedules.

Headways were then calculated for all observed trains and compared with official NYC Transit timetable information. Headways were calculated between departure times only, as is the practice at NYC Transit, in order to maintain comparability of results. From this observed and official headway information, service regularity was then calculated using NYC Transit's own definition of service regularity:

SERVICE REGULARITY: the percentage of intervals between trips departing from all scheduled timepoints, not including terminals, which is within ± 50 percent of the scheduled interval (for all scheduled intervals less than ten minutes), or within ± 5 minutes of the scheduled interval (for scheduled intervals of 10 minutes or more.[9]

However, in order to better gauge service regularity along the length of a line, regularity was calculated for all stations surveyed, not just for stations used as timepoints on NYC Transit's official timetables.

During some observation periods, several lines experienced unusual delays and/or

[8] The **Franklin Avenue Shuttle** has only four stops and thus no observation points exist which cannot be considered the first or last two stops.
[9] NYC Transit Committee Agenda, May 1996, and NYC Transit Committee Agenda, May 2000.

delay problems were announced on trains or in stations. The Council contacted NYC Transit to inquire about the nature of these delays in order to better analyze our data. Finally, it is important to note that these observations do not constitute a statistically representative sample of daily subway service. Nevertheless, our results do suggest conditions that were experienced by riders on business days that were essentially selected at random.

Table One: Field Observations For All Routes By Location, Date[10], Time, and Availability Of Official Timetable Data

Route	Destination	Station	8 am-10 am	Noon-1 pm	Timetable Data Available[11]
1/9	South Ferry	79 St	5/26,30		
1	South Ferry	50 St		5/23-24	
2	Flatbush Av	14 St	5/15,17	5/15,17	
3	New Lots Av	14 St	5/15,17	5/15,17	
4	Woodlawn	42 St	6/2	6/6-7	•
4	Woodlawn	City Hall	6/16		•
5	Dyre Av/238 St	42 St	6/2	6/6-7	•
5	Dyre Av/238 St	City Hall	6/16		•
6	Pelham Bay Park/Parkchester	42 St	6/2	6/6-7	•
6	Pelham Bay Park/Parkchester	14 St/Union Sq	7/11		
7	Times Square	45 Rd	5/5,12		
7	Times Square	Grand Central		5/9-10	
A	Lefferts Blvd/Far Rockaway	59 St	5/3,9		•
A	Lefferts Blvd/Far Rockaway	145 Street		5/23,30	•
B	Bedford Park Blvd/145 St	Pacific St	5/11,23	5/12,16	
C	Euclid Av	59 St	5/3,9		•
C	Euclid Av	145 St		5/23,30	•
D	205 Street	7 Av (Brooklyn)	5/3,5		
D	205 Street	47-50 Sts		5/2,5	•
E	World Trade Center	Continental Av	5/5,10	5/17,18	•
F	Coney Island	Continental Av	5/5,10	5/17,18	•
FAS	Franklin Avenue	Park Pl	5/1, 19	5/3,5	
FAS	Prospect Park	Park Pl	5/1, 19	5/3,5	
G	Smith-9 Sts	Northern Blvd	5/19,26	5/19,26	
J/Z	Broad St/Chambers St	Eastern Pkwy	5/24,31	6/7,9	•
L	8 Av	Myrtle Av	5/3,8	5/22; 6/12	•
M	Metropolitan Av	Pacific St	5/11,23		•
M	Myrtle Av	Wyckoff Av		6/5,16	
N	Ditmars Blvd	Pacific St	5/11,23		•
N	Coney Island	City Hall		5/18,26	
Q	21 St-Queensbridge	7 Av (Brooklyn)	5/3,5		
Q	21 St-Queensbridge	47-50 Sts		5/2,5	•
R	Continental Av	Pacific St	5/11,23		•
R	95 St	City Hall		5/18,26	

[10] All dates in 2000.
[11] Denotes stations for which exact timetable data were available from NYC Transit.

Findings

Our 2000 subway service reliability findings are reported below. Overall reliability figures are presented first. Morning peak and midday off-peak reliability figures are next discussed individually in order to better indicate reliability during these hours.

We have interpreted our findings based upon the numbers of delays per week a regular customer might expect to experience if his or her line were running at a given level of reliability. For example, a 90 percent reliability figure means that a regular rider on this line might experience one delayed train per week out of 10 round trips taken to and from work. Likewise, an 80 percent figure would indicate that a regular rider could expect two delayed trips per week. While it is difficult to determine what different riders would consider an unacceptable frequency of delay, the Council feels that two delays per week is the maximum that riders should expect to experience if service is to be characterized as adequate.

Thus, in the discussion of our findings, we have characterized reliability results of 80 percent or above as acceptable. We consider results of 90 percent or above to be very good. Results below 80 percent are considered inadequate.

Overall Service

Table Two, below, summarizes our findings for overall reliability. These figures represent an average of all observations for each line:

Table Two: Routes Ranked By Overall Service Reliability

Route	Overall Service Reliability
C	100%
FAS	100
G	97
L	94
D	93
N	91
A	90
E	89
Q	86
F	84
2	83
3	81
6	79
B	79
R	79
7	75
M	72
1/9	71
J/Z	68
5	62
4	56

These results are fairly good overall, with 12 lines that exhibited reliability levels at or above 80 percent. Very good were the **A**, **C**, **D**, **G**, **L**, **N**, and the **Franklin Avenue Shuttle**, which performed with 90 percent reliability or better. Of these, the best performers were the **C** and **Franklin Avenue Shuttle** which both performed at 100 percent.

Inadequate reliability levels of below 80 percent were found on the **1/9**, **4**, **5**, **6**, **7**, **B**, **M**, **R**, and **J/Z** lines. Of these, three lines–the **4**, **5**, and **J/Z**–operated below 70 percent reliability, with the **4** line exhibiting the worst overall score for any train–56 percent.

Many lines operate along the same trunk lines for significant portions of their routes, and most of these share similar reliability figures. However, one set of these related lines–the **N** and **R**–showed marked differences in reliability. The **N** was a top performer, at 91 percent reliability. However, the R exhibited an inadequate reliability level of 79 percent, 12 percentage points below its sister line.

Morning Peak Service

Table Three, below, summarizes our findings for morning peak service reliability:

Table Three: Routes Ranked By Morning Peak Reliability

Route	Morning Peak Reliability
C	100%
FAS	100
G	100
D	93
L	93
N	88
A	87
E	86
2	84
Q	81
7	80
F	80
B	79
3	78
R	75
6	73
M	68
1/9	67
J/Z	63
4	57
5	54

Our morning reliability figures are also fairly good overall, with 12 lines that exhibited reliability levels of 80 percent or better. Very good reliability, of 90 percent or better, was recorded on the **C**, **D**, **G**, **L**, and **Franklin Avenue Shuttle**. Of these, the top

performers were the **C**, **G**, and **Franklin Avenue Shuttle**, which each exhibited a perfect morning-peak reliability figure of 100 percent.

Inadequate morning-peak reliability levels of below 80 percent were found on the **1**/**9**, **3**, **4**, **5**, **6**, **B**, **J**/**Z**, **M**, and **R**. Morning reliability for the **4** and **5** were worst, coming in at 57 percent for the **4** line, and 54 percent for the **5** line. We also note that the poor performance of the **J**/**Z** and **M** was almost entirely for hours when no service diversion was in effect, not during hours of Williamsburg Bridge-related trackwork when severe diversions are frequently required.

During the study, systematic major delay problems occurred on the **1**/**9** at 79th Street in Manhattan (on May 30) and the **B**, **M**, **N**, and **R** at Pacific Street on the Fourth Avenue trunk line in Brooklyn (on May 23). The **1**/**9** delays were caused by customers holding doors at 103 St. The **B**, **M**, **N**, and **R** delays were caused by a sick passenger at Grand Street. We recognize these unanticipated incidents detracted from the reliability scores for these lines. However, even on the alternate observation day for these lines, May 12, reliability problems were still experienced on the **M** and **R**.

We noted that frequent announcements were made in stations regarding the Fourth Avenue delays, including the locations and arrival times of the next trains and alternate routes. It also appeared that NYC Transit was able to quickly recover from the delays on Fourth Avenue keeping platform crowding to a minimum at Pacific Street. This was not the case for the **1**/**9**, where the Council witnessed severely overcrowded trains and an overcrowded platform.

In addition, we are surprised to find that the **7** line operated with a morning reliability level of just 80 percent. Because the **7** is one of only two lines that travel on an exclusive set of tracks, and has the capacity for the shortest headways in the system, the line is generally expected to perform better. By contrast, the **L**, the other line that has its own exclusive set of tracks, exhibited a morning-peak reliability level of 93 percent.

We are also surprised to find that the lines with the worst morning reliability figures–the **4**, **5**, **6**, **J**/**Z**, and **M**–were lines that experienced no scheduled diversions and no major delays during either of our observation days, suggesting systemic reliability problems with these lines.

Midday Off-Peak Service

Table Four, below, summarizes our findings for midday off-peak service reliability:

Table Four: Routes Ranked By Midday Off-Peak Reliability

Route	Midday Off-Peak Reliability
5	100%
A	100
C	100
E	100
F	100
FAS	100
J/Z	100
L	100
N	100
Q	100
6	93
D	90
G	90
R	90
3	88
1/9	82
B	80
2	79
M	60
4	55
7	55

Our midday off-peak reliability figures stand in sharp contrast to our morning results, with fully 14 lines that perfomed at very good reliability levels of 90 percent or better. This includes 10 lines–the **5**, **A**, **C**, **E**, **F**, **J**/**Z**, **L**, **N**, **Q**, and **Franklin Avenue Shuttle**–that performed at perfect, 100 percent reliability during midday.

Only four lines showed reliability levels below 80 percent: the **2**, **4**, **7**, and **M**. The **4** and **7** shared the worst midday reliability figure, 55 percent. Again, we are very surprised at the performance of the **7**, historically one of the system's most reliable trains. According to NYC Transit, customers diverting to the **7** because of service interruption on the **N** line caused a delay on the **7** line on May 10. However, the previous day, when no major delays were identified, the line's reliability still remained well below 80 percent. We also note that reliability for all four of these lines showed itself to be worse for midday hours, when scheduled service and ridership are lower, than for the busier morning peak.

Equally unexpected is to find the **5** and **J**/**Z** were among the top performers for midday reliability, both at 100 percent. This is especially notable for the **5** which exhibited a morning-peak reliability figure of just 54 percent. However, we note that not much service is offered on either of these lines during midday hours (especially for the **J**/**Z**

which continues to undergo severe midday service diversions due to work on the Williamsburg Bridge), thus it is easier for NYC Transit to maintain reliable headways.

The Council's 2000 Results Compared with NYC Transit's 1996 and 2000 Reliability Figures

Because we utilized NYC Transit's own methodology for calculating service regularity for this 2000 study, our results cannot be compared with the results of our previous studies in 1992 and 1996. However, our 2000 results are comparable with NYC Transit reliability figures. **Table Five**, below, compares our overall morning-peak and midday results for 2000 with NYC Transit's overall reliability figures for both the 1st quarter of 1996, the year of our previous study, and the 1st quarter of 2000:

Table Five: NYCTRC Overall Reliability Figures for 2000 Versus NYC Transit Figures for 1996 and 2000[12]

Route	1st Quarter 1996 (NYC Transit)	1st Quarter 2000 (NYC Transit)		Average 2000 Morning Peak and Midday Service (Council)
1/9	72%	68%		71%
2	74	72		83
3	78	76		81
4	67	65		56
5	72	74		62
6	61	68		79
7	78	79		75
A	67	71		90
B	69	77		79
C	74	87		100
D	75	87		93
E	73	74		89
F	64	80		84
FAS	N/A	N/A		100
G	76	85		97
J/Z	77	86		68
L	87	83		94
M	79	89		72
N	66	83		91
Q	79	89		86
R	74	81		79

It should be noted that NYC Transit's figures represent reliability for all hours between 6:00 a.m. and 9:00 p.m. Our figures are based on data collected over a limited, two-day span specifically between the hours of 8:00 a.m. and 10:00 a.m. and Noon and 1:00 p.m.. However, our figures do represent a snapshot of service on these days during some of the subway system's busiest hours.

[12] NYC Transit 1st quarter figures indicate reliability for January, February, and March of the relevant year, and represent an average for all service between 6:00 a.m. and 9:00 p.m. during those months. They have been rounded to the nearest whole number.

It appears from the table that reliability has improved for most lines since 1996. Unfortunately, reliability seems to have decreased for the **1**/**9**, **4**, **5**, **7**, and **M** lines, echoing concerns which the Council has recently heard from riders of these lines.

Differences can also be seen between the 2000 numbers collected independently by the Council and NYC Transit. These differences stem from the fact that NYC Transit's figures represent a greater range of hours, allowing more chances for service variations to manifest themselves. For some lines, including the **2**, **3**, **6**, **A**, **G**, **L**, and **N**, our figures are higher. For other lines, including the **4**, **5**, and **J**/**Z**, our figures indicate worse reliability. This suggests to the Council that NYC Transit's method of reporting a single reliability figure for most hours of the day may not accurately reflect reliability during the busiest hours of system use for these lines.

Conclusions and Recommendations

Overall, our results indicate that subway service is reliable for most lines, and is improved for most lines from 1996. Our reliability figures for the B-Division, or lettered trains, were markedly good, especially for midday hours. Our results also indicate that NYC Transit does a good job of maintaining reliable service during planned weekday daytime diversions in service, such as on the **J**/**Z** line during midday. We do note, however, that most planned diversions occur during overnight and weekend hours, time periods not included in this field study. Further, the agency appears to provide a quick response to unexpected major delays, such as those experienced during the morning peak on the **1**/**9** line in Manhattan and on the **B**, **M**, **N**, and **R** lines in Brooklyn.

However, reliability on the **1**/**9**, **4**, **5**, **7**, and **M** lines appears to have worsened since 1996. Furthermore, nine lines exhibited reliability figures below 80 percent, including almost all A-Division IRT (numbered) lines. This means that customers who ride these lines–the **1**/**9**, **4**, **5**, **6**, **7**, **B**, **J**/**Z**, **M**, and **R**–can expect more than one out of every five trains to be delayed. For riders who regularly travel to work using these lines, this means that delays can be expected more than twice a week. The Council finds this frequency of delays unacceptable and believes special attention should be paid to these lines to improve reliability.

Of even greater concern, five of these lines, and some of the worst performers overall–the **4**, **5**, **7**, **J**/**Z**, **M**, and **R**–actually performed more poorly on a day-to-day basis than lines which experienced planned service diversions and unplanned delays. This is most disturbing for the **R** line, which already performed much worse than its sister line in Brooklyn and Manhattan, the **N**. However, the problem is serious for all of these lines, because they are lines which many people take or to which they transfer. Obviously, systemic problems exist on these lines which are detracting from reliability. These problems should be identified and addressed.

Finally, a comparison of our results with NYC Transit's own reliability data suggests that the agency's figures may not adequately represent reliability during peak hours because the agency reports only a single, overall figure for reliability. The Council suggests that in the future NYC Transit also report a breakdown of their reliability data by time of day, as we have done in this report, to better indicate reliability for peak and off-peak hours.

Appendix A: Why Wait Assessment May Not Be a Good Reliability Measure

In May, NYC Transit proposed a new indicator of reliability to replace the agency's current service-regularity measure. According to the agency, the new measure, **wait assessment**, will simplify the assessment of reliability by providing a single standard to be met by all trains, instead of a standard based upon a proportion of individual scheduled headways as is the case with the service-regularity measure.[13] NYC Transit defines wait assessment using a single, three-minute standard:

WAIT ASSESSMENT: The percentage of intervals that are no more than three minutes over the scheduled interval.[14]

All headways which last no more than three minutes longer than they are scheduled to be, including headways shorter than scheduled, are counted as reliable headways using wait assessment as a measure of reliability.

NYC Transit also believes wait assessment will better represent the level of service experienced by customers, and that it is a more reasonable standard because delays shorter than three minutes are ignored. The agency does not feel delays of three minutes or less can be dealt with in a useful manner due to their length.

The Transit Riders Council disagrees with NYC Transit on these points and feels that service regularity may be a better reliability measure for several reasons. First, because wait assessment ignores delays shorter than three minutes, it is far more forgiving than a service-regularity measure when scheduled waits between trains are short. During peak and midday hours, many subway lines have headways of 6 minutes or less. Using the existing service-regularity measure, three-minute long delays on these lines would be counted as unreliable service. However, using the proposed wait-assessment measure, such delayed service would be counted as reliable. This would have the effect of systematically raising reliability figures, suggesting improved service, when in fact no change in service has actually occurred.

To illustrate this effect, we analyzed the reliability of the west side IRT **1**/**9**, a line with headways of five minutes or less during most morning peak and midday hours, using both indicators. During much of our morning observation hours, the **1**/**9** experienced extensive delays, including the diversion of express service to the local track and at least one train which bypassed the station. These delays led to a prolonged crowding condition on the platform at 79th Street and severely crowded trains.

Our results, summarized below in **Table Six**, show a marked difference between the two competing reliability measures. As expected, the service-regularity figure suggests subpar service, indicating that only 67 percent of morning peak **1**/**9** trains arrived with

[13] For the definition of **service regularity**, please see **Methodology**.
[14] NYC Transit Committee Agenda, May 2000.

reliable headways. However, because the originally scheduled headways were short to begin with, the wait-assessment measure ignores almost all of the delay problems and indicates that 93 percent of **1**/**9** service arrived with reliable headways. During midday hours, our results also show an improved reliability rating when the wait-assessment measure is used.

Table Six: Service Regularity Versus Wait Assessment for Morning Peak and Midday Off-Peak Service on the West Side IRT 1/9 Line

Observation Period	Service-Regularity Rating	Wait-Assessment Rating
Morning Peak	67%	93%
Midday Off-Peak	82	91

In order to accurately represent the experience of riders, any measure of reliability must take into account the effect that variations from scheduled service have on riders. When scheduled headways are short to begin with–as they tend to be along highly used lines such as the **1**/**9**–even small variations can lead to very crowded trains and platforms. As our comparison shows, the proposed wait-assessment measure ignores these effects. We do not believe that the delays and resulting crowding we witnessed on the morning-peak **1**/**9** train at 79th Street merit a reliabilty rating of 93%.

However, even during periods of normal service when no major delay condition exists, the wait-assessment measure can seem to inflate reliability figures. Our results for midday reliability on the **1**/**9** differ by nine percentage points depending on which measure is used: service regularity; or wait assessment.

We are also concerned that the wait-assessment measure does not allow early trains to be counted against reliability, no matter how early trains arrive. Any variation from scheduled departure times, whether early or late, can impact passenger loading and platform crowding all along a line. By defining all early trains as reliable, the wait-assessment measure inflates performance figures even further.

Subway headways are scheduled with reference to the specific ridership needs for each line at specific times of day. Because the existing service-regularity measure is calculated in proportion to these scheduled headways, it can more accurately represent the impacts on riders of all types of delay–including delays affecting lines with short headways and early trains. While we understand NYC Transit's concern to develop a less-complicated measure of service reliability, any new measure must retain at least the same level of precision as the current service-regularity measure affords and also must avoid inflating performance results when no real change in service has occurred. We believe wait assessment fails on both of these counts and, therefore, should not be adopted by NYC Transit as its new measure of service reliability.

Appendix B: Service Reliability for Individual Field-Observation Days

The below table summarizes our reliability findings for individual morning and midday observation periods. To achieve the overall results for morning and midday reliability, averages were calculated for the individual observation periods taken as a whole. For the exact dates of each observation period, please see **Table One** in the **Methodology** section.

Table Seven: Morning Peak and Midday Off-Peak Service Reliability for Individual Field-Observation Days

Route	First Morning Observation	Second Morning Observation	*Overall Morning Peak Reliability*	First Midday Observation	Second Midday Observation	*Overall Midday Off-Peak Reliability*
1/9	74%	59%	***67%***	82%	82%	***82%***
2	90	79	***84***	71	86	***79***
3	83	72	***78***	88	88	***88***
4	64	50	***57***	09	100	***55***
5	55	52	***54***	100	100	***100***
6	81	65	***73***	100	86	***93***
7	79	82	***80***	73	36	***55***
A	93	80	***87***	100	100	***100***
B	100	57	***79***	60	100	***80***
C	100	100	***100***	100	100	***100***
D	87	100	***93***	80	100	***90***
E	86	86	***86***	100	100	***100***
F	84	76	***80***	100	100	***100***
FAS	100	100	***100***	100	100	***100***
G	100	100	***100***	80	100	***90***
J/Z	73	53	***63***	100	100	***100***
L	90	95	***93***	100	100	***100***
M	46	91	***68***	60	60	***60***
N	92	83	***88***	100	100	***100***
Q	69	94	***81***	100	100	***100***
R	83	67	***75***	80	100	***90***