ABSTRACT: Public transport is one of the base services that cities offer their citizens. Public services are specific because their primary goal is providing a good services and not profits. Of course they have to operate profitably but their primary purpose is to provide transport to citizens to whom this might be the only transport alternative. This work analyses quality dimensions of tram public transport and empirically assessed the quality of tram transport in Zagreb through an empirical examination in a period of ten days. Measuring each dimension of quality the overall quality is assessed. The results show that the quality on each dimension is satisfactory for each day except for the day number 2. and that duration of the trip does not diminish perceived quality as suggested by previous literature.

KEY WORDS: service quality, service quality dimensions, public transport, control charts.
1. INTRODUCTION

Services in Croatia follow the trend in developed countries in GDP contribution. The most recent statistics from Croatian Bureau of Statistics shows that 84% of GDP is obtained by services. City public transport is in national classification under NKD H 49.31. However this level of detail is not obtainable for only tram transport rather for the whole group NKD H- transportation which accounts with 5.4% of GDP which shows its significance (DSZ, 2018). For majority of services there are many service providers, but for public transport there is no other public provider. We can say that Zagreb Electrical Tram (ZET) is a monopoly and some citizens have no other options. Therefore, this subject is chosen exactly for this reason. Is a quality of public transport in Zagreb satisfactory given its quasi monopolistic position?

Both Woodcock et al. (2009) and Buehler (2009) state that there are many benefits of using public transport instead of cars such as reducing pollution, reducing congestions in rush hours, lower accidents, less trouble with parking and parking accidents. But, in order to promote a public service it has to be of a good quality. Therefore in this work dimensions from the P-TRANSQUAL model will be investigated. P-TRANSQUAL model has four dimensions, namely, Comfort, Tangible, Personnel, and Reliability.

We also have to ask the fundamental question of why people prefer their own cars instead of public transport. Beirão and Cabral (2007) name reasons such as higher comfort in their cars, are more trusting in their abilities as drivers. On the other hand, Poliak et al. (2017) name problems of public transport in terms that they connect only some parts of the city which diminishes the willingness of passengers to use public transport. As such, than, car drivers have greater comfort in choosing the time when they want to travel and the places they want to travel to. But, they also state that there is a large population of passengers that have no cars and no other means to travel to their destination so they have to use public transport. Those are primarily students and elderly people, and in current time also tourists. However, Xia et al (2015) state that because of traffic congestion in major cities and pollution, environmentally aware people should be drawn more to public transport. These congestions, frequent repair of city roads because of large volume of cars on city streets and fuel price should encourage more people to use public transport. Woodcock et al. (2009) also states that there is a measurable reduction in CO$_2$ if people were willing to switch from their cars to public transport. In order for this switch from private cars to public transport is to happen the quality of transport service should be good (Šojat et al. 2017; Lai et al, 2011).

The main purpose of this paper is to explain dimensions of quality in public transport. The main contribution is the construction of the measurement instrument for evaluating quality of public transport for the purpose of measuring quality objectively as possible. On grounds of this measurement instrument, an empirical investigation is performed by observing quality dimensions. On grounds of this measurement data a control chart is constructed which graphically depicts is quality in acceptable range (Wheeler and Chambers, 1992; Woodall et al., 2004).

Because quality is subjective and may depend on waiting time (Zhao et al. 2014) a regression analysis is performed on influence of length of the trip on quality dimensions, to see if waiting and traveling time diminishes perceived quality.
2. SERVICE QUALITY

Academics and practitioners are in agreement that focus on quality in services is one of the most important factors of service success (Ladhari, 2009). Service quality is an important issue because services of good quality enhance customer satisfaction, customer loyalty, and positive word of mouth (Bakti and Sumaedi, 2015). Rowley (1998) states that client’s perceptions of quality depend on their expectations from the service. If the obtained service meets their expectations than the service is good. If the service exceeds their expectations than the service will be perceived as excellent. The main instrument for measuring service quality is SERVQUAL. It is grounded in five dimensions that include (Khorshidi et al, 2016):

1. Tangibility: physical facility, equipment, and appearance of personnel.
2. Reliability: ability to perform the promised service dependably and accurately.
3. Responsiveness: willingness to help customers and provide prompt service.
4. Assurance: knowledge and courtesy of employees and their ability to inspire trust and confidence.
5. Empathy: caring, individualized attention the firm provides its customers.

Redman et al. (2013) argues that the fifth dimension (empathy) in a decision to choose a private car instead of public transport is not relevant rather that basic accessibility, reliability and mobility provision are more important. Chowdhary and Prakash (2007) warn that generalizations in services are not possible but they do depend on the ratio of human or capital intensity. They state that empathy will be important in labour intensive services but less important in capital intensive industries such as public utilities (transport, electricity and the like).

Cronin and Taylor (1992) and Babakus and Boller (1992) analysing the SERVQUAL model state that items that form a quality dimension might be completely different depending on the service industry. Maksimović et al. (2017) in their research from the perspective of public transport users, show that the most important aspects are:

1. Transport speed (the duration of the trip from origin to destination).
2. Service cost (direct costs that a user has in using the transport).
3. Comfort (conditions under which the transport is carried out, starting with the vehicle, the method of payment for the service, the possibility of comfortable accommodation in the vehicle, the courtesy of staff etc.).
4. Insurance (the likelihood that the journey will take place without any incident, without any adverse consequences).
5. Reliability (guarantee of a certain travel time that can be anticipated in advance).

Mazzulla and Eboli (2006) performed a survey asking passengers what are their main attributes to public transport quality. They show that the main attributes are:

1. frequency of buses,
2. number of bus stops,
3. cleanliness of interior, seats, etc.,
4. comfort on bus,
5. security against crimes on bus,
6. availability of shelter and benches at stops,
7. information on services,
8. availability of seats on bus.

From the above discussion it is obvious that there is a need to use a specifically adapted measurement instrument for each service. Therefore in this work a modified Bakti and Sumaedi (2015) P-TRANSQUAL model is used. It consists of four dimension of quality and is consisted of 20 indicators. For the purpose of the replication of this study we show the measurement instrument here as a whole.

1. **Comfort:**
   - IK1 Passengers capacity in public transport
   - IK2 Safety while using public transport services
   - IK3 Comfortable temperature on public transport
   - IK4 Security from crime while using public transport services
   - IK5 Safety related to the behaviour of other passengers
   - IK6 Smoothness of the ride (without abrupt breaks)

2. **Tangible:**
   - IK7 Facilities and equipment of public transport
   - IK8 Condition of public transport seats
   - IK9 Cleanliness of the interior, seats, and windows of public transport
   - IK10 Condition of public transport machine
   - IK11 Public transport cleanliness of the graffiti

3. **Personnel:**
   - IK12 Neatness of personnel
   - IK13 Helpfulness of personnel
   - IK14 Responsiveness of personnel
   - IK15 Courtesy of personnel
   - IK16 Understanding of passengers’ need

4. **Reliability:**
   - IK17 Adequacy of public transport
   - IK18 Waiting time of public transport
   - IK19 Travel time of public transport
   - IK20 Delivery to the destination

The comfort dimension has six items. Capacity means the number of seated and standing places in comparison to the number of people using the transport. This item is especially important in rush hours when people go to or from work and students to and from school. Safety is represented by two factors and deemed important because if a passenger does not feel safe he will also not feel comfortable in the public transport. Comfortable
temperature means does the passenger feel too cold or too hot or does the air conditioning or heating works and used in the right time. Smoothness of the ride means that the driver obeys all the traffic laws, does not drive too fast and then use abrupt breaks. Safety related to the behaviour of other passengers means that other passengers are also behaving politely, that there is no imposition from for example drunk passenger.

The tangible dimension refers exclusively to the appearance, condition and cleanliness of the public transport vehicle. The general state of the vehicle is considered (whether there are any malfunctions or irregularities), cleanliness inside and outside, in particular the cleanliness and condition of seats and windows. The personnel dimension implies the behaviour of the service provider, that is, in this case the driver of public transport. Are the drivers friendly, whether they are willing to help if necessary, whether they answer passenger’s questions regarding transport are being considered and understanding of passenger’s needs. Understanding the passenger’s needs refers, for example, to a situation when a passenger runs to the station to arrive before the transport leaves, and the driver waits for these few seconds in order for the passenger to embark. The reliability dimension is nothing other than fulfilling the promise that the service will be executed accurately and at the time as promised. Thus, in the dimension of reliability, there are indicators of the time of travel and waiting.

According to Larson et al. (1991) in psychology of waiting show that known waiting times seem shorter than unknown waiting times. Therefore it is suggested that customers in general know how long they will wait. With advancement of technology and its affordability, most of train stations in Zagreb have displays with waiting times for trams.

3. METHODOLOGY

Data collection was performed by observation method. It means that the researcher used the public tram and filled the check list for each drive. Additionally, the researcher measured the time of each drive. Data were collected for all three type of public services (tram, bus and rail) but for this article only the results from tram ride are presented. The measurements were performed during 10 days, at various hours from the 1st to the 10th of June while students are still not on their break and the vacations have not yet started for working people. The researcher measured for each item in the modified P-TRANSQUAL template was the item satisfactory or not, meaning that he counted actually the errors in the public transport. After that a control chart is made for this period of 10 days to evaluate and graphically display is the quality acceptable. Control charts are used because they display when something is not in order. If a measurement falls below the lower control limit or above the upper control limit, the errors are no more random but there is some cause why the measurement fell out of the limits, and it is a signal to investigate what caused the anomaly (Benneyan, 1998; Montgomery, 2009; Horvat et al. 2006). Therefore the first goal is to assess is quality of city tram public transport in Zagreb of adequate quality using Statistical Process Charts (SPC). So the first hypothesis will be put in affirmative way, that the quality of Zagreb tram transportation is satisfactory, that is that none of the measurement will fall out of control limits.

H1  Public tram transport in Zagreb is satisfactory, that is, no measurement falls out of the control limits of the SPC chart.
The second hypothesis checks the results of Zhao et al. (2014) that duration of transport and waiting will negatively affect perceived quality. Therefore the second hypothesis is:

H2  Duration of transport negatively affects perceived quality.

Hypothesis H2 will be checked using regression analysis. The hypothesis actually researches how duration of the ride affects perceived quality.

4. RESULTS

Data collection was performed by observation method. It means that the researcher used the public tram and filled the check list with all 20 items of the P-TRANSQUAL model for each drive. After a complete analysis of all measurements a summation table by each quality dimension is made and shown in Table 1. Table 1. depicts number of errors for each dimension for the duration of the experiment period.

<table>
<thead>
<tr>
<th>Quality dimension</th>
<th>Number of observed defects measurement</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comfort</td>
<td>4 10 5 7 3 2 7 7 3 50</td>
<td></td>
</tr>
<tr>
<td>Tangible</td>
<td>1 1 1 1 0 0 0 1 0 8</td>
<td></td>
</tr>
<tr>
<td>Personnel</td>
<td>0 0 2 1 1 0 0 0 0 4</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>1 7 2 1 3 4 1 1 3 23</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>6 18 10 10 8 7 3 9 11 3 85</td>
<td></td>
</tr>
</tbody>
</table>

There are several control charts in use such as XR chart, p chart and chart. XR chart is excluded because attribute measures were researched. That leaves p and c chart. P chart cannot be used because it can be used only if percentages are in question. That leaves the c chart which is most suitable and additionally allows that for a single measurement there is no upper limit on number of defects. Using formulas for c chart (Epprecht et al. 2003):

\[
\bar{c} = \frac{\text{number of observed defects}}{\text{number of samples}} \quad \text{Eq. 1}
\]

Formula for standard deviation for c chart is:

\[
\sigma_c = \sqrt{\bar{c}} \quad \text{Eq. 2}
\]

Upper and lower control limits are measured by

\[
\text{UCL} = \bar{c} + z\sigma_c \quad \text{Eq. 3}
\]

\[
\text{LCL} = \bar{c} + z\sigma_c \quad \text{Eq. 4}
\]
Parameter $z$ is set to 3 and means that 3 standard deviations will be included in the calculus of upper and lower limit. Three standard deviations ensure that there is no false alarm, that is, that random variation caused a measurement to fall outside control limits (Epprecht et al. 2003):

Table 1 is then graphically represented in Figure 1.

![Graphical representation of measurement](image)

Figure 1: Graphical representation of measurement

It can be seen from Figure 1. that most errors are in dimension of Comfort (10 errors) which was measured by (Passengers capacity in public transport, safety while using public transport services, Comfortable temperature on public transport, Security from crime while using public transport services, Safety related to the behaviour of other passengers, Smoothness of the ride (without abrupt breaks)). The second highest number of defects is in dimension of reliability (7 defects) measured by (adequacy of public transport, waiting time of public transport, travel time of public transport and delivery to the destination).

To construct the control chart all errors are summed (last row in Table 1.) and displayed in Table 2. sum of all defects across all four dimensions for the researched period.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Sum</td>
<td>85</td>
</tr>
</tbody>
</table>
Using Eq1-4, average number of defects are evaluated, standard deviation and control limits.

\[
\bar{c} = \frac{\text{number of observed defects}}{\text{number of samples}} = \frac{85}{10} = 8.5
\]

\[
\sigma_c = \sqrt{\bar{c}} = \sqrt{8.5} = 2.9
\]

\[
\text{UCL} = \bar{c} + z\sigma_c = 8.5 + 3\times 2.95 = 17.2
\]

\[
\text{LCL} = \bar{c} - z\sigma_c = 8.5 - 3\times 2.95 = -0.2 = 0
\]

The obtained control chart is presented in Figure 2.

![Figure 2: c control chart for public tram in Zagreb for the period of measurement](image)

According to Figure 2. It can be seen that on the second day there are more defects than other days and the quality of that day measured in number of defects exceeds upper control limit. Especially pronounced errors can be seen from Figure 1 and the figure shows that most defects of that day were in dimension of comfort and reliability. It means that the user of the chart should go back to measurements of the day two and find attributable cause that showed that quality was not acceptable on day two. Since the researcher did not record outside weather conditions such as rain or nearby event, the researcher cannot objectively attribute the cause of a larger number of defects caused that day. This also presents the limitation of the study, that is, outside factors should be taken into account too.

It can be concluded, however, that in ten days period only one day shows inadequate quality (day 2) which does not mean that quality is not good rather that there existed a cause that increased the number of defects that day. Therefore, in nine out of ten cases the quality is in acceptable range and hypothesis H1 was confirmed. But since the hypothesis was
posed in affirmative way that no measurement will fall out of control limits, hypothesis H1 has to be rejected, since the measurement of day 2 fell out of control limits.

The second hypothesis checks the validity of preposition posed by Zhao et al. (2014) that the length and waiting for the service will diminish its perceived quality. Table 3 displays the average length and waiting for the trip for that day and the overall number of defects for that day.

<table>
<thead>
<tr>
<th>Measurement day</th>
<th>Duration in minutes</th>
<th>Number of defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30,26</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>30,08</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>27,14</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>28,15</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>29,44</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>28,13</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>28,38</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>29,53</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>33,29</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>26,54</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure 3. Average duration and overall number of defects per day

From Figure 3, one can see that the average duration was not abnormal on the second day with the highest number of defects rather on the day nine. The hypothesis will the du-
ration diminish perceived quality is checked with regression analysis. The dependent variable is number of defects (negative quality), and the independent variable is the average duration in minutes. The regression was performed using SPSS statistical package and gives following results: The regression coefficient for the whole model is $R=0.428$ and coefficient of determination $R^2$ is 0.183. That means that only in 18.3% of cases does the duration affect number of defects. The whole model is not significant giving $p=0.217$ which signals that the number of defect is not statistically significant predictor of length of the trip. Therefore, H2 can be only partially accepted that indeed as the duration becomes longer the number of defects also rise, but the hypothesis has to be rejected because the whole model is not significant.

5. DISCUSSION OF RESULTS

Even though both hypotheses are rejected that is that none of the measurement will fall out of control limits and H2 that the length of the trip increases the number of defects, the results are still educative. For the purpose of testing hypothesis H1 modified Bakti and Sumaedi (2015) P-TRANSQUAL model is used for the purpose of objectively verifying quality of public tram service in Zagreb city. The results are still positive because in 9 of 10 cases (90%) the quality of public tram transport is satisfactory. The remaining 10%, that is, day 2, additional research into causes of larger number of defects should be performed. It could be that that day it was raining and the comfort was lower than usual. On rainy days the roads are slippery and maybe the reliability part of the quality construct was higher.

Measuring variables that were mentioned in this paper, and measuring additional environmental variables (that present the limitation of this work) one could continuously enhance quality by reducing the causes of defects. With each small improvement quality of service can be improved Montgomery (2009). But this paper also shows that it is not easy to measure quality on 20 items and that the observations necessities a skilled observer to monitor and adequately record observations. Therefore, from the economic perspective it is questionable, is worth the effort of educating a special quality squad, to perform those measurements and in this way objectively assess quality of the public transport.

Hypothesis H2 is also rejected because the model was not significant, which means that hypothesis proved by Zhao et al. (2014) that the duration of the trip will negatively affect perceived quality was not supported by our measurement and analysis. But, because the standardised beta coefficient was positive, that still allows to conclude that really as the length of the trip increases, so will increase number of defects, and diminish the perceived quality of the service. This may explain why for longer trips, passengers prefer alternative means of transport than the city tram.

Limitations of this work

The limitation of this paper is in the fact that environmental variables were not measured which could explain why the measurement of day two fell outside of control limits. In future examinations and replications of this work, one or more environmental variables have to be recorded too.
Managerial implications

According to Meredith and Shafer (2003) companies investing in quality can form longstanding relationships with their clients and have a source of positive advertisement. Additionally, cutting defections in half doubles company growth and improves customer retention by 5%. Since it would be in ZET’s interest to have more passengers it would be in their interest to invest in quality and form a special quality unit. A case of Toronto public transport (reported by Seijts et al., 2012) showed that incorporating a quality team unit can enhance quality and thus attract more passengers.

6. CONCLUSION

Public transport is one of the most important services the city provides for its citizens. It is important that this service is on its highest level so that the satisfaction of the citizens is raised, and more citizens attracted to public transport. In this paper many reasons why quality of public transport is important. It is also important to promote quality of public transport to attract more passengers to use public transport, instead of using their cars. In this way congestions, pollution and accidents could be reduced. In city of Zagreb the network of public transport is well developed and connects all parts of the city. It is also important that dimensions of Comfort, Tangible, Personnel and Reliability are with as little as possible defects so that the overall quality of the service is good. To each individual passenger not each dimension of quality is equally important. For example to one the dimension of comfort may be most important and that he or she feels safe during the ride, that the trams are not overcrowded, and that the temperature is adequate, while to some other dimension of reliability might be more important, such as waiting, duration of the trip, arriving at the right destination and the like.

Tram transport in Zagreb on grounds of this examination shows high level of quality (in 90% of cases the quality was satisfactory) but there is still room for improvement.

Hypothesis H2 that the quality of the transport will fall with the duration of the trip was only partially confirmed. This could mean that for short distances people prefer tram ride instead of the car or other means of transport while for longer rides people maybe think of other means of transport.

The importance of public transport is not only in reducing congestions, rather also in other aspects such as reducing pollution, reducing number of accidents, reducing problems of parking, street noise, and the like. Therefore, results of this study could be used in promoting public transport in order to reduce negative effects on environment using own cars.

REFERENCES


