3.1 Introduction

The aim of this chapter is to investigate quality attributes and dimensions in the area of railway passenger services and to identify the knowledge gap shown in the conceptual literature. This aim is achieved in three steps. The first step uses the conceptual framework identified in chapter Two and evaluates its applicability to model passenger satisfaction in railway passenger services. The second step attempts to unfold contextual issues necessary for possible adjustments to be made in the conceptual framework supported by literature through focus group discussions. The aim of the last step attempts to identify quality attributes and domains in global railway passenger services obtained from past research studies for the purpose of refining them to suit real-life environments, upon which provisional models may be developed to evaluate passenger satisfaction in a dynamic railway passenger setting in Indian Railways through expert opinions. This chapter therefore reviews the contextual research literature pertaining to service quality and passenger satisfaction from the service quality perspective in railway passenger services. The review provides a theoretical framework for the development of RAILQUAL Instrument – incorporating existing research findings – is ultimately offered to conceptualise the formation of the related to service quality attributes in Indian Railway passenger services.

3.2 THE INDIAN RAILWAYS

The Indian Railways (IR), more than 150 years old, is among the one of the largest and oldest railway systems in the world. It has an extensive network with 63,974 route kilometres of route length and played an integrating role in the social and economic development of the country. IR is the principal mode of transportation for the long haul freight movement in bulk, long distance passenger traffic and mass rapid transit in
suburban areas. It occupies a unique position in the social economic map of the country and is considered as a vehicle and a barometer of growth. It is also biggest state owned enterprise in India and contributes about 1% of India’s GNP. Indian Railways is running approximately 13,000 trains a day, including 9000 passenger trains which carries 20 million passengers and 2.44 million of freight between 7000 railway stations each day. To fathom the scale, consider the fact that each day Indian trains travel four times the distance to moon and back (Raghuram G 2008, Sudhir Kumar, 2009).

The journey of railways in Indian sub-continent started modestly in 1853 with 34 kilometres (kms). Iron wheels rolled on rails on 16th April, 1853, where the first-ever train, with a capital of 3.8 million, carrying 400 people in 14 carriages, covered the 21 mile distance in about 75 minutes from Bombay to Thane (Sailaja, 1988, Alivelu, 2008). By 1950, India had a net-work of about 34,000 miles. In 1948, immediately after Independence, there were as many as 42 different railway systems consisting of 13 class I Railways, 10 class II railways and 19 class III Railways. The class of Railway is fixed depending on gross earnings. The major task of the Indian Railways (IR) is to integrate the above mentioned divided railway systems of the sub-continent such as princely state railways, state owned railways and to bring them under one management. The Railway Board in 1950 decided for the regrouping of the Indian Railways into six zonal systems, namely the Northern, the North Eastern, the Southern, the Central, the Eastern and the Western Railways. The unequal distributions of workload on some of the railways have led to further bifurcation of zones. Eastern Railway was split into two zones, namely, Eastern Railway and South Eastern Railway. Similarly, North Eastern was split into North Eastern Railway and North East Frontier Railway. Thus, by the year 1958, there were eight zones on Indian Railways.

The functioning of the new zones continued to be watched closely and based on regular analysis of their working, minor adjustments are carried out whenever necessary, with a
view to improve their utility and efficiency. Particularly close watch is kept on the rapidly increasing workloads of some of the new zones, in order to provide relief and streamline their operation. The formation of South Central Railway in 1966 as the ninth zone, in order to improve the services of southern parts of India, is made with some marginal adjustments from Southern and Central Railways. Carving out of South Central Railway has resulted in stability in the zonal formation at least for little more than three decades in the history of IR. In order to bring about greater efficiency in administration, speedy implementation of the on-going projects, better customer care, reduction of work load on the administrators of each zone, Indian Railways have decided to create seven new zones by territorial re-adjustment of existing zones.

Thus IR has been reorganized into sixteen railway zones by adding seven new zones to the already existing nine zones namely East Central, East Coast, North Central, North Western, South Western, West Central, South East Central from 2003 onwards. Recently Metro Railway, Kolkata has been declared as New 17th Zonal Railway w.e.f 29.12.2010. Five Year Plans from 1950 onwards threw up enormous challenges to the railways for playing key role in the industrial and all round development of the Indian Economy. Thus, what started as a system to the interests of the foreign masters has in the last hundred and fifty eight years, developed into a significant means of transportation for socio-economic development of a welfare society like India.

IR is state owned and operated under the Ministry of Railways (MOR), Government of India (GOI). The MOR functions under the guidelines of Minister of Railways assisted by Minister of State for Railways. The policy formation and management of Indian Railway board comprises of Chairman and six functional Members (Figure 3.1). Wide powers are vested in the board to effectively supervise the running of 17 Zonal railways each headed by a General Manager. Zonal Railways are further divided into smaller operating units called Divisions. There are 68 Operating Divisions in Indian Railways at present, each under a Divisional Railway Manager. In addition, there are a number of production units,
training establishments, construction organization and other railway establishments. These are generally headed by General Managers. 12 subsidiary organizations under the MOR also doing specialised jobs contributing to IR’s growth and progress. (Raghuram G., Rachna Gangawar, 2007)

Figure 3.1 Organisational Structure of Indian Railways
3.2.1 Indian Railway Passenger Business

Indian Railways is a commonly used mode of public transportation in the country. During 2009-10, it carried 7,246 million passengers. Passenger kilometres, which is calculated by multiplying the number of journeys by mean kilometric distance was 903 billion and passenger earnings is Rs 1547.96 crore (Sudhir Kumar, Shagun Mehrotra, 2008).

The trend of passenger traffic since 1950-51 is shown below in Table 3.1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Suburban (all classes)</th>
<th>Non suburban</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Class</td>
<td>Mail Express</td>
<td>Ordinary</td>
</tr>
<tr>
<td>1950-51</td>
<td>412</td>
<td>25</td>
<td>52</td>
</tr>
<tr>
<td>1960-61</td>
<td>680</td>
<td>15</td>
<td>96</td>
</tr>
<tr>
<td>1970-71</td>
<td>1219</td>
<td>16</td>
<td>155</td>
</tr>
<tr>
<td>1980-81</td>
<td>2000</td>
<td>11</td>
<td>260</td>
</tr>
<tr>
<td>1990-91</td>
<td>2259</td>
<td>19</td>
<td>357</td>
</tr>
<tr>
<td>2000-01</td>
<td>2861</td>
<td>40</td>
<td>472</td>
</tr>
<tr>
<td>2006-07</td>
<td>3514</td>
<td>58</td>
<td>713</td>
</tr>
<tr>
<td>2007-08</td>
<td>3689</td>
<td>66</td>
<td>776</td>
</tr>
<tr>
<td>2008-09</td>
<td>3802</td>
<td>76</td>
<td>895</td>
</tr>
<tr>
<td>2009-10</td>
<td>3876</td>
<td>86</td>
<td>983</td>
</tr>
</tbody>
</table>

Source: Indian Railways’ year book 2009-10

Passenger revenue in different classes with corresponding number of passengers and passenger kms in 2009-10 is shown below in Table 3.2.
Figure 3.2 Indian Railway network
Indian Railways plays an important role in the transportation of passengers. Passenger services can be segmented into suburban and non-suburban. It is estimated that the Indian Railways have a share of about 20% in the passenger transport market, in terms of passenger kilometres. Approximately 90% IR passenger revenues come from non-suburban segment. Passenger services account for nearly 60% of IR’s total transport effort, in terms of train kilometres, but yield less than 30% of total revenues. The growth in terms of transportation output in passenger segments has been increasing in recent past particularly in “premium” segment, but overwhelming share of passenger traffic remains in the lower classes. Suburban passenger business is viewed as non-profitable venture for IR due to the low fare structure and low priced monthly season tickets. The subsidy provided to this segment in the year 2009-10 amounted to approximately Rs 16,000 crores (MOR, 2010, yearbook 2009-10). Suburban accounted for approximately 54% of the...
originating passengers in 2009-10, while contributing to only 7% of passenger revenue. Nearly 3.3 billion passengers make use of IR’s non-suburban passenger services annually. The upper class travellers, though comprise just 1.19% of originating passengers, account for nearly 24.21% of the passenger revenues.

3.2.2 Cross - Subsidization

Freight earnings today account for over 66% of the total traffic earnings of IR. Freight tariff on Indian Railways is among the highest in the world as would be seen by a comparison with the freight rates per tonne kilometre of the other world railways. Particularly, as compared to the major freight railways like US Railroads, Chinese and Russian Railways, the freight rates of Indian Railways are extremely high. In fact, the rate of US Railroads is one fourths that of IR’s as seen below in Figure 3.2.

![Figure 3.2 - Freight Rates Comparison](image)

**Figure 3.3a Average Freight Revenue per Tonne Kilometre**

Source: World Bank report as quoted in Indian Railways vision 2020 report

On account of the large volume of passenger traffic in India and the inevitable financial impact on them of any fare increase, increasing passenger fares has always been sensitive
issue, particularly for second class passengers. As a result, passenger fares in India are low compared to most foreign railways. The high freight rates and low passenger fares make the fare-freight ratio of India Railways one of the lowest in the world. Over the past 5 years, it reduced from a level of 0.32 to 0.36 (Table below). The fare-freight ratio in case of China is 1.2 and Korea is 1.4 as shown in Table 3.3 and Figure 3.3.

<table>
<thead>
<tr>
<th>Fare Freight ratio</th>
<th>2004-05</th>
<th>2005-06</th>
<th>2006-07</th>
<th>2007-08</th>
<th>2008-09</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Earnings per NTKM (Paise)</td>
<td>74.84</td>
<td>80.83</td>
<td>85.39</td>
<td>90.98</td>
<td>98.73</td>
</tr>
<tr>
<td>b. Earnings per PKM (Paise)</td>
<td>24.52</td>
<td>24.53</td>
<td>24.75</td>
<td>25.74</td>
<td>25.93</td>
</tr>
<tr>
<td>Fare-Freight Ratio (b/a)</td>
<td>0.32</td>
<td>0.30</td>
<td>0.39</td>
<td>0.28</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Table 3.3 Fare Freight ratio

Source: White paper on Indian Railways 2009-10

Figure 3.3b Ratio of average passenger fare to average freight rates
Source: World Bank report quoted in white paper on Indian Railways 2009-10
3.2.3 Market Share of Indian railways

Even though traffic volumes on Indian Railways have gone up over the years, rail shares has gone down steadily over the past few decades. There is unfortunately no fully authentic data on the market share on rail transport. However, a study conducted by Rail India Transport and Economic Services (RITES) for the planning Commission in 2009 called the ‘Total transport system study on traffic flows and modal costs’ gave a historical overview of the shares of different transport modes in India as show below Table 3.4.

Table 3.4 Trends in Road and Rail Passenger Traffic shares, 1950/51 – 2000/01

<table>
<thead>
<tr>
<th>FY</th>
<th>Railway Pkm billion</th>
<th>% of total</th>
<th>Road Pkm billion</th>
<th>% of total</th>
<th>Total Pkm billion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1951</td>
<td>67</td>
<td>68</td>
<td>31</td>
<td>32</td>
<td>98</td>
</tr>
<tr>
<td>1960</td>
<td>74</td>
<td>51</td>
<td>71</td>
<td>49</td>
<td>145</td>
</tr>
<tr>
<td>1970</td>
<td>113</td>
<td>37</td>
<td>189</td>
<td>63</td>
<td>302</td>
</tr>
<tr>
<td>1980</td>
<td>199</td>
<td>32</td>
<td>421</td>
<td>68</td>
<td>620</td>
</tr>
<tr>
<td>1990</td>
<td>281</td>
<td>21</td>
<td>1,029</td>
<td>79</td>
<td>1,310</td>
</tr>
<tr>
<td>1991</td>
<td>296</td>
<td>22</td>
<td>1,079</td>
<td>78</td>
<td>1,375</td>
</tr>
<tr>
<td>1996</td>
<td>342</td>
<td>20</td>
<td>1,368</td>
<td>80</td>
<td>1,710</td>
</tr>
<tr>
<td>2001</td>
<td>457</td>
<td>13</td>
<td>3,000</td>
<td>87</td>
<td>3,457</td>
</tr>
</tbody>
</table>

Source: Ministry of Railways, Planning commission.

3.2.4 Earnings of the Indian Railways

The years spanning 2003-04 to 2007-08 saw unprecedented growth, led by strong economic fundamentals. Passenger earnings exceeded the growth in passenger kilometres. The passenger earnings grew at a CAGR of 10.52% in the period under review as compared to 9.24% during the period 1999-2004 while passenger kms grew by 9.13% as shown in Table 3.5
### Table 3.5 Growth in Passenger Earnings and Realisation per pkm

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Passenger Earnings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(cr)</td>
<td>14113</td>
<td>15126</td>
<td>17225</td>
<td>19844</td>
<td>21931</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YOY Growth</strong></td>
<td>6.12%</td>
<td>7.18%</td>
<td>13.87%</td>
<td>15.21%</td>
<td>10.52%</td>
<td>10.52%</td>
<td>9.24%</td>
</tr>
<tr>
<td><strong>PKMs</strong></td>
<td>575608</td>
<td>616632</td>
<td>695821</td>
<td>771070</td>
<td>839159</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YOY Growth</strong></td>
<td>6.19%</td>
<td>7.13%</td>
<td>12.84%</td>
<td>10.81%</td>
<td>8.83%</td>
<td>9.13%</td>
<td>6.02%</td>
</tr>
<tr>
<td><strong>Earnings per PKM</strong></td>
<td>24.52</td>
<td>24.53</td>
<td>24.75</td>
<td>25.74</td>
<td>26.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YOY Growth</strong></td>
<td>-0.04%</td>
<td>0.04%</td>
<td>0.90%</td>
<td>4.00%</td>
<td>1.52%</td>
<td>1.27%</td>
<td>3.03%</td>
</tr>
</tbody>
</table>

Source: White paper on Indian Railways 2009-10

**PKM** - Passenger Kilo Metres  
**YOY** - Year on Year Growth Rate  
**CAGR** – Compounded Annual Growth Rate

The lower CAGR of 1.27% in earnings per PKM during the period under review compared to 3.07% in the previous five year period is primarily because of no direct increase in fares and in fact reduction in some classes.
3.3 Refining and modifying SERVQUAL model

Parasuraman et al., (1991a) asserts that the SERVQUAL instrument is a concise multi item scale to be applicable across a broad spectrum of service industries. It can be used to understand better service expectations and customer perceptions. The SERVQUAL developers also stress the practical and diagnostic value of capturing both expectations and perceptions for the sake of identifying service shortfalls rather than explaining variance in an overall measure of perceived service quality by simply using performance based measurement. Pitt, Oosthuizen, and Morris (1992) also tested the SERVQUAL scale in a high-tech industrial setting and found that it is a reliable and valid instrument amid the concerns that the five factor solution evidenced and factor structure is not generic. But the studies of SERVQUAL replication studies show that the scale possesses moderately high reliability in terms of coefficient alpha which serves as bench mark of internal consistency.

As a concluding remark, Lewis and Mitchell (1990) consider that “in the meantime, SERVQUAL remains the most reliable tool available for the measurement of service quality in the 1990s”.

Despite all these strengths, the SERVQUAL instrument has come under much criticism in the areas of conceptual foundation and operational difficulties. Carman (1990) questioned the replication of service quality dimensions under the SERVQUAL methodology across a number of different service industries without modifying wording and adjusting number of items which are unique to certain industry settings. Figure 3.4 shows the SERVQUAL model in an explanatory way and Table 3.6 summarises some of the major criticisms and suggestions for improvement of the SERVQUAL instrument.
Chapter 3: Contextual Research Review

A “GAPS” Model of Service Quality

Fig 3.4 SERVQUAL (Parasuraman, et al., 1991a)
### Table 3.6 Weakness of SERVQUAL Instrument

<table>
<thead>
<tr>
<th>Weaknesses of SERVQUAL</th>
<th>Suggestions for improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SERVQUAL treats all 22 items in the scale equally important</td>
<td>1. Introduction of weighting (Carman, 1990; Lewes and Mitchell, 1990).</td>
</tr>
<tr>
<td></td>
<td>2. Point allocation across 5 dimensions (Parasuraman, Zeithaml and berry, 1991a)</td>
</tr>
<tr>
<td>2. Negatively worded items create difficulties to respondents (e.g. more time to read, more comprehension mistakes and more likely to attach negative emotional connotations)</td>
<td>1. Use of warning to the respondents on the existence of negative/positive wording in the instruction section (Babukus and Boller, 1992)</td>
</tr>
<tr>
<td></td>
<td>2. All negatively worded items to be changed to positively worded items (Parasuraman, Zeithaml and berry, 1991a)</td>
</tr>
<tr>
<td>3. Problems of operationalizing service quality as differential score (e.g. psychological constraint and unstable factor structure)</td>
<td>1. Use of performance based SERVPERF (Cronin and Taylor, 1992)</td>
</tr>
<tr>
<td></td>
<td>2. Use of non-differential score by rephrasing each item it the form of “How” (Brown, Churchill and Peter, 1993)</td>
</tr>
<tr>
<td>4. The word “should” in the expectation statements is unrealistic and causes confusion</td>
<td>1. “Should” expectations changed to “would” expectation (Parasuraman, Zeithaml and berry, 1991a)</td>
</tr>
<tr>
<td>5. Separate lists of statements may be less accurate than comparisons made almost the same point in time</td>
<td>1. Comparison of expectation and perception measures at the same time using bipolar semantic differential graphical scale (Lewis and Mitchell, 1990)</td>
</tr>
<tr>
<td></td>
<td>2. Use of direct measurement. For example “The visual appeal of XYZ’s physical facilities are (much better, better, about the same, worse, much worse) than I expected” (Carman, 1990)</td>
</tr>
<tr>
<td>6. Restricting customers’ responses to a 7 point mask subtle variation in their expectations and perceptions</td>
<td>1. Use of graphical scale (Lewis and Mitchell, 1990)</td>
</tr>
<tr>
<td>7. Adjectives used in the scale cause ambiguity (e.g. How “up-to-date” is it?)</td>
<td>1. Use of bipolar semantic differential scale (Lewis and Mitchell, 1990)</td>
</tr>
<tr>
<td></td>
<td>2. Certain items to be modified. For example “up to date” was changed to “modern looking”</td>
</tr>
</tbody>
</table>
Chapter 3: Contextual Research Review

<table>
<thead>
<tr>
<th></th>
<th>Parasuraman, Zeithaml and Berry, 1991a</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.</td>
<td>Increased length of the questionnaire</td>
</tr>
<tr>
<td></td>
<td>1. Expectations and perception measures to be combined by using bipolar semantic differential graphical scale or using direct measurement approach.</td>
</tr>
<tr>
<td>9.</td>
<td>Dimensionality of 22 item instrument not generic</td>
</tr>
<tr>
<td></td>
<td>1. Number of items and wording for particular services setting to be modified (Carman, 1990)</td>
</tr>
<tr>
<td>10.</td>
<td>Measure should account for multiple service functions</td>
</tr>
<tr>
<td></td>
<td>1. Measures for each service function.</td>
</tr>
<tr>
<td>11.</td>
<td>Only focuses on the process quality attributes</td>
</tr>
<tr>
<td></td>
<td>1. Outcome quality attributes to be included (Rechard and Allaway, 1993)</td>
</tr>
<tr>
<td>12.</td>
<td>Questionable reliability, convergent validity and discriminant validity</td>
</tr>
<tr>
<td></td>
<td>1. Convergent and discriminant validity to be strengthened with additional more stringent evaluation criteria (Babukus and Boller, 1992)</td>
</tr>
<tr>
<td></td>
<td>2. Use of non-difference score to improve reliability and discriminant validity (Brown, Churchill and Peter, 1993)</td>
</tr>
<tr>
<td>13.</td>
<td>Problem in dealing with finite ideal point attributes (i.e. one on which customers’ ideal point is at a finite level and therefore, performance beyond which displease customers.)</td>
</tr>
<tr>
<td></td>
<td>1. Need to evaluate the models in situation which there is a high incidence of finite ideal point attributes (Teas, 1993b)</td>
</tr>
</tbody>
</table>

3.4. Critiques of the SERVQUAL model

The authors in their study published in 1991 (Parasuraman et al., 1991a) refined the model. The amendments were mostly connected to the statements applied by the model. On one hand they modified the text of the scale statements and on the other hand they have exchanged the original statements by new ones and negative statements by positive ones. They held that the high average values of the original model were caused by the normative wording of the statements. The amended definition focuses on what the customers expect of a supplier supplying excellent services (such as excellent companies will insist on error-free records), rather than on how an organization has to conduct an activity (such as companies should keep their records accurately). Besides the statements of the expectation scale, the statements of the perception scale were
Chapter 3: Contextual Research Review

rephrased as well. Within the dimension of tangibility, they have introduced questions aimed at communication materials; in the assurance dimension aimed at the knowledge and expertise of the employees. In addition to the already existing expectation, perception scales, they have introduced a third scale into the model. They asked the respondents to divide 100 points between the dimensions, according to how important they find it relevant to a certain service. They thought that by correcting the service quality values of the individual dimensions by these relative importance values, a more accurate and reliable total service quality value may be derived.

Based upon the results of the tests, carried out after the amendments, the authors refined their original findings on the dimensionality of the model. They argued that the five dimensions of SERVQUAL may be clearly separated, but the factor-analysis often can not reflect the difference of the characteristics. They concluded that the “refinement still reflects the basic five-dimensional structure of the original scale” (Parasuraman et al., 1991a, p. 431.), but the individual factors are not clearly separated, and the correlation between them took a higher value as compared to the original researches (this is particularly true in case of responsiveness and the assurance dimensions). The tangibles factor, proved to be one-dimensional in the original scale, was divided into two-sub dimensions (one is connected to physical appearance the other to the appearance of the employees and communication material).

In 1993 the model was subjected to a further revision mostly based on the criticism directed at the expectation concept. Contrary to the former interpretations, expectations were defined as zone of tolerance, with the two end-points of the desired and adequate service levels (Parasuraman et al., 1993). In the new examination this expectation theory was included in the SERVQUAL model by using these three different models. On the questionnaire of the three-scale model, the responders were asked to form a judgment regarding the desired, the adequate and the perceived service. In the two-scale version the responders were asked to compare the perceived service with the desired and the adequate level, while the one-scale form
Chapter 3: Contextual Research Review

only asked that the perceived performance is evaluated relevant to the desired service level. This time they have used a nine point Likert-scale adding the “I do not know / I do not have opinion” evaluation option. They ceased correcting by the relative value, because weighing did not improve the explanatory potential of the model.

The test results showed that all three forms have high reliability as well as their projective, differentiating and similarity validity is adequate. They found that although the one-scale method has the best projection capabilities, the most information is carried by the three-scale form. Thereby they showed that by applying SERVQUAL’s disconfirmation approach, managers will acquire more adequate and usable information for their decisions on service quality, in comparison with the other forms. Their conclusion, that the original five dimensions of the SERVQUAL cannot be identified clearly, is of great importance. Responsiveness, assurance and empathy overlaps, thus finally three dimensions may be identified: tangibles, reliability and the common dimension of responsiveness and assurance and empathy (Parasuraman et al., 1994b, p. 221.).

Parasuraman et al.’s confident declarations concerning the SERVQUAL model’s general character, reliability and validity resulted in further analyses and critical remarks taking these as starting points among experts of both theory and practice, (Coulthard, 2004, p. 481.).

Based on the first criticisms (Carman, 1990) concerning, among others, necessity of negatively worded statements and reading of the concept of expectation, Parasuraman et al. (1991a, 1994a, 1994b) conducted modifications on the model and re-defined certain statements for the sake of easier intelligibility.

However, most of the repeated researches did not support the authors’ statements. It must be added, however, that these repeated researches most often applied research methods somewhat modified as compared to the original model. The original statements were changed; the 22 statements were re-defined and several ones were detracted from and
Chapter 3 : Contextual Research Review
added to them, depending on the service sector where the research was conducted. Various technologies were introduced also in methodology. For instance, other scales were applied instead of the Likert scale of 7 points and polling by questionnaires as well as the related administration were conducted in various ways.

Parasuraman et al. (1991a) drew attention to the fact that researchers could get a proper result on the quality of services only if they apply the original model in its entirety, without any modification. In their opinion, minor modifications on the definition of the statements do not hurt the integrity of the entire model, however, omission of certain statements or insertion of new statements would question the entirety and reliability of the model. The following are the criticisms on SERVQUAL till date across the world by different researchers.

3.4.1. Theoretical Bases
Criticisms concerning the theoretical bases of SERVQUAL essentially question the validity, reliability and applicability of the model. By comparison of customers’ expectations to the perceived performance of service providers, SERVQUAL interprets the perceived quality as an instrument which is related to customers’ satisfaction but does not coincide with it. However, according to the arguments of Oliver (1980), Cronin, Taylor (1992, 1994) and Iacobucci (1994), the perceived quality is much rather an attitude. Cronin and Taylor stated that the disconfirmation-based SERVQUAL model did not measure the service quality or the customers’ satisfaction, it was a model based on a “flawed paradigm” (Cronin and Taylor, 1992, p. 64.).

Andersson (1992) questions the economic, statistical and psychological bases of the SERVQUAL model’s formation. First, the model disregards the costs of service quality improvement. Next, the method of data collection applied to the chosen statistical method (factor analysis) is false as it applies an ordinal scale (Likert scale) instead of the interval scale, which is more applicable to factor analysis. Deriving from the
application of ordinal scale, it can hardly handle connections and interactions between the individual dimensions. And finally, psychological factors were poorly considered during formation of the model.

3.4.2. Process Orientation

The nordic model (Grönroos, 1984) examined among service quality models and the further methods developed from that (Lehtinen and Lehtinen, 1991; Rust and Oliver, 1994) interpret service quality in a broader sense. They separate the dimensions concerning the service’s result (technical quality, result quality) and process (functional quality, process quality), completing them by the dimension of image (Grönroos, 1984), company quality (Lehtinen and Lehtinen, 1991) or organisational environment (Rust and Oliver, 1994).

In essence, the SERVQUAL model only focuses on the process; in fact, it examines only one segment of service quality, the quality of the servicing process. Although several statements can lead to the reading of further quality-related aspects, the model is still process-directed. Further components displaying the general reading of service quality are missing, such as factors concerning the core service, the service as “product”, the judgement of the organisation as participant of market and society or, even the business policy of the organisation (Sureschandar et al., 2001). Recent researches (Brady et al., 2002; Chui, 2002) miss further factors: they urge on that the model should cover the rate of the service as well as feelings and emotions related thereto.

3.4.3 Dimensions

A major part of criticisms in relation to the SERVQUAL model concerns the dimensionality of the model. The repeated researches (like Babakus and Mangold, 1989; Carman, 1990; Finn and Lamb, 1991, Saleh and Ryan, 1992; Babakus and Boller, 1992; Bouman and van der Wiele, 1992; Gagliano and Hathcote, 1994; S. Llosa et al., 1998; Dabholkar et al., 2000; Cunningham and Young, 2002) most often could not
Chapter 3: Contextual Research Review

reproduce the five original service quality dimensions. Based on the settings of the applied factor analysis methods, several researches could identify from six to nine dimensions (Carman, 1990), while other researches only one (Cronin and Taylor, 1992; Brown et al., 1993). The latter researchers expressly argued in their essay for the one-dimensional service quality. Llosa et al.’s (1998) research did not support this latter assumption but neither did it prove Parasuraman et al.’s (1988) results. Nearly 74 per cent of the persons polled by researchers classified the 22 original statements into 3-6 groups.

The number of dimensions varied on a broad scale depending on the field of service examined. According to Babakus and Boller, the domain of service quality may be factorially complex in some industries and very simple and unidimensional in others” (Babakus and Boller, 1992, p. 265.). Carman’s (1990) research on hospitals identified nine dimensions: (reception of patients, accommodation, food, secrecy, nursing, introduction to treatment, courtesy, directing visitors, planning of dismissal, invoicing). Gagliano and Hathcote (1994) examined the clothing retail industry and defined four dimensions (personal attention, reliability, tangibles, comforts). Bourman and van der Wiele (1992) described three factors in car servicing, namely gentleness to the customer, tangibles and fairness.

Parasuraman et al. (1994) also faced the problem of dimensionality when examining the SERVQUAL model repeatedly. In the repeated research, they could identify only three dimensions instead of the five original ones. Above tangibles and reliability, the three other dimensions (responsiveness, assurance and empathy) fused to one common factor.

3.4.4 Problem of the Factor Structure

It is a further problem that the statements forming the original factors do not clearly fit in the factor to which we would expect them. In Carman’s (1990) essay, for instance, two of the statements originally belonging to the SERVQUAL’s empathy factor came into
the dimension of tangibles during analysis of a dental clinic’s service quality and similar anomalies were experienced during examination of other service sectors as well. This observation was supported by further essays (Buttle, 1996).

The five factors of SERVQUAL are “composed” by 4-5-4-5-4 statements sequentially. The few composing items (statements) attached to each dimension result in the changeability of the factor structure. By application of more statements, the stability of the single dimensions can be increased, which was also accepted by Parasuraman et al. (1991). Carman (1990) applied 40 statements upon analysis of the hospital service quality, Bouman and van der Wiele (1992) did 48 in the field of car services, Dabholkar et al. (1996) did 28 in the case of retail services, while Sureshchandar et al. (2001) did 41.

The definition of the single statements i.e. the positive or negative coding can also affect the factor structure. Thirteen of the SERVQUAL’s 2 statements were worded in positive, while nine in negative form. Each of the “denying” statements belonged to one factor (responsiveness and empathy). Although Parasuraman et al. aimed to decrease the possibility of systematic yes-no answering, later this practice was still rejected (Parasuraman et al., 1991a). This happened in part because this method increased the time to fill in the questionnaires: twenty-two statements had to be assessed twice and in addition, even the negative statements had to be interpreted by the answering person. The other reason was that Babakus and Boller (1992) had proved during application of the factor analysis that the positive-negative definition results in ”method factors” and not dimensions to be derived from the statements themselves (Buttle, 1996, p. 22.).

Analysing the research, we can state that the foregoing research have not found a generally applicable dimension structure describing the service quality universally and comprehensively. Dimensions may vary depending on both the examined field of service and the applied research method.
3.4.5 Role of Expectations

Several researchers have questioned the wording of the statements of the expectation scale. It is not clear what the expectations must refer to: the level expected under ideal, excellent or the given environmental conditions. The wording of the expectation scale upon formation of the model effected that most of the answering persons gave the mark of six or seven to the statements on the Likert-scale of seven degrees (where the two ends meant “do not agree at all” and “totally agree”). Authors were recognizing that the “should” terminology might be contributing to unrealistically high expectation scores, which questioned the model’s applicability, therefore the wording of each statement were modified (Parasuraman et al., 1991a, p.422.). For instance, they indicated the statement “excellent companies will insist on error-free records” instead of the statement “companies should keep their records accurately”. However, Brown et al. (1993) noted that this modification had little effect.

Based on his researches, Teas (1993, 1994) reached the conclusion that answering persons read expectations in different ways. In his opinion, differences between the single expectation-scores related to each statement do not derive from the different judgement of each answering person related to the given statement but much rather from the fact that everybody reads the concept of expectation in another way. He thought that answering persons applied any of the following six readings in relation to expectations (Teas, 1993).

Service attribute importance – how important is the given statement for the answering person

Forecasted performance – possibility of the future realisation of the performance expected by the answering person (can be)

Ideal performance – optimal level, which may be the performance of the service provider
Chapter 3 : Contextual Research Review

Deserved performance – which *should be* the performance of the service provider *considering the investments* for the use of services,

Equitable performance – which *ought to be* the performance of the service provider at given costs,

Minimum tolerable performance – which minimally *must be* the performance of the service provider.

Based on the conclusions reached from analysis of reading of the concept of expectation, Teas (1993, 1994) deemed the expectation an ideal base of comparison in his models (Evaluated Performance and Normed Quality).

In relation to the role of expectations in the model, Teas (1993, 1994) raises further questions. The –1 value of the perceived service quality (P-E) measured according to the disconfirmation (concerning a given statement) can arise based on six different combinations of P (performance) and E (expectation) scores (P=1, E=2; P=2, E=3; P=3, E=4; P=4, E=5; P=5, E=6; P=6, E=7). Do the values determined by the different pairs of scores mean the same perceived service quality? Are the expectation rates universally valid for all service providers in a given sector or do different expectations belong to the different service providers? Does one standard expectation of general validity belong to each (SERVQUAL) statement and dimension or does the customer have other expectations in case of the different services depending on their location, for instance?

Accordingly, Iacobucci et al. (1994) would rather apply some general standard instead of the subjective and ambiguous concept of expectation in the model. Similarly, according to the definition of Voss, Roth, Rosenzweig, Blackmon and Chase, service quality is „based on the meeting or exceeding of certain established service standards” undertaken by the given service provider (Voss, Roth, Rosenzweig, Blackmon and Chase, 2004, p. 213.). Thus, according to their wording, expectations are not based only on the customer’s subject but on the performances undertaken by service providers, which may be influenced by the ability of service provider as well.
Chapter 3 : Contextual Research Review

3.4. 6 Two Administration, Order Keeping

It was a further criticism that the execution of the method, administration of the double scale is difficult, as an answering person must assess the same 22 statement twice: first on the basis of the expectations and then based on the perceived performance. This is not only time-consuming and boring but it often leads to the so-called exhaustion effect, which questions the suitability and truth of the collected data (Bouman and van der Wiele, 1992).

It is neither the same in what order are the answering persons polled on the two scales: first the expectation scale and then the perception scale, eventually in reverse order or perhaps both at the same time. Caruana (2000) proved by analysis of the SERVQUAL’s developed, three-scaled model (Zone of tolerance) that the answers given first to the expectation scales (desired, adequate) influenced the perception rates significantly. These results correspond to the psychological research showing that answering persons are influenced by the previous answers and experiences related to the same question (Strack and Martin, 1987).

3.4.7 Dynamism

Customers insert their previous experiences into their expectations and they modify them flexibly, eventually influenced by technical development. It is not clear how the model captures this continuing, dynamic change of the expectations. According to several longitudinal researches, expectations may be higher and higher by progress of time (Parasuraman et al.’s works, for instance) but they also may decrease (in the medical sector, for instance). Researchers agree that research on service quality should focus on study of dynamic models in the future (Buttle, 1996).
3.4.8. Psychometric problems due to Difference Rates

Some researchers (like Brown, 1993; Spreng and Singh, 1993; Van Dyke et al., 1997) drew attention to psychometric problems concerning analysis of SERVQUAL’s difference rates. According to their arguments, further analysis of a new variable deriving from the difference between two different index numbers (in the specific case the index concerning the perceived performance and that representing the expectations) leads to psychometric problems related to reliability and validity of differences. Thus, the question arises as to what the scale measures infect. In addition, researchers also questioned usability of the generally applied Cronbach-alpha as regards to difference rates (Buttle, 1996).

3.4.9. Problems Deriving From the Applied Likert-Scale

Although most researchers modified the number of statements of the original SERVQUAL model, their wording and the applied methods in their essays, almost all of them insisted on the application of the Likert-scale. However, two main problems have to be stressed in relation to this evaluative scale: the issues of centring and the number of categories (Smith, 1995; Buttle, 1996).

3.4.9A Centring

Answering persons, who, in the lack of knowledge and experiences, can not assess some questions, indicate the centre (meaning rate 4 in the case of SERVQUAL) as the “do not know” choice is missing. Thus, final results are distorted significantly. However, Babakus and Mangold (1992) have shown that, in the lack of the “do not know” choice, a significant number of answering persons leave one or more questions unanswered and cause problems in processing of the questionnaire this way. On the other hand, the lack of the “do not know” choice may move the answering person still to indicate something (despite the fact that he does not know the statements...
Chapter 3 : Contextual Research Review

in question and does not have related experiences). This, however, may lead again to distorted final results and false service quality rates.

Besides that centring can be regarded as a neutral value judgement or a “do not know” answer, centring may raise a further problem. Namely, it means some satisfactory solution from the answering person’s point of view: by filling in the questionnaire quickly, he did what he was asked to do but, as a consequence of the hurry, he did not thought over the statements. Instead, rather choosing the middle course, he did not give too high or too low rates, either. Of course, neither these rates reflect the answering person’s actual attitude, thus they may lead to false conclusions in the long run (Krosnick et al., 2002).

3.4.9B Number of Categories

When Likert-scale is applied, the answering person must indicate how much he agrees or disagrees with a number of statements concerning the examined unit. (Malhotra, 2005, p. 336.). The number of the scale’s categories and the wording of the specific category rates, among others, however, significantly influence answers.

The SERVQUAL scale contains seven categories, which corresponds to the traditional suggestions concerning the number of categories. Only the ends of the single categories are verbalized by definitions of “totally agrees” and “totally disagrees”. Some researchers (Smith, 1992) state that definition of only the ends may move the answering persons rather to choose the extreme rates. This shift to the positive direction was verified by several repeated researches related to the SERVQUAL model and Parasuraman et al. (1994a) acknowledged the phenomenon as well. However, defining each scale category one by one does not definitely improve the accuracy and reliability of the data (Andrews, 1984).

In the repeated researches related to the SERVQUAL model, most researchers (like Finn and Lamb, 1991; Babakus and Mangold, 1992, Dabholkar et al., 1996) applied a scale of
five points as an alternative of the Likert scale of seven points. Other researchers (like Robledo, 2001; Brady et al., 2002) also applied a disconfirmation scale of five points, which however, was formed by re-definition of the ends of the scale („much better than expected”; „much worse …”). Robledo (2001) tested the SERVPEX model consisting of 26 statements and of three dimensions determined by the statements (tangibles, reliability, customers’ relations), by means of analysis of the service quality of airlines, applying the disconfirmation scale of five points. In the essay, the SERVPEX model’s better validity was justified as compared to the SERVPERF (a model based on perceived performance only) and SERVQUAL models.

Notwithstanding the above, the issues of the applied scale, the number of the scale’s categories and the wording of the single categories are still open.

### 3.5 Review of Service quality in Public Transport Industry

Allen and Di Cesare (1976) considered that quality of service for public transport industry contained two categories: user and non-user categories. Under the user category it consists of speed, reliability, comfort, convenience, safety, special services and innovations. For the non-user category, it is composed of system efficiency, pollution and demand. Silcock (1981) conceptualized service quality for public transport industry as the measures of accessibility, reliability, comfort, convenience and safety. Pullen (1993, p261) defined quality of service for local public transport industry as a concept that involves “those attributes of the service which affect its fitness for purpose” and “the attributes, and indeed fitness for purpose, required detailed definition in relation to local objectives and circumstances”.

Traditionally, the performance indicators for public transport industry are divided into two categories: efficiency and effectiveness. Under the efficiency category, the measures are concerned with the process that produce the services while the measures in the effectiveness category are used to determine how well the services provided are with
respect to the objectives that are set for them (Pullen, 1993). Quality of service is one of the performance indicators under the effectiveness category. It is composed of accessibility, reliability, comfort, convenience and safety. Allen and DiCesare (1976) classified the performance measures into three categories: quantity of service, quality of service and cost/revenue of which are further divided user and non-user measures. Passengers waiting time, lost mileage and characteristics of each journey mode (time of arrival, time spend, time of arrival at destination) are commonly used measures to measure quality of service (Pullen, 1993). More recently, output quality measures that have been used for the rail system in Britain include train performance (delays per passenger train), train overcrowding, asset condition (broken rail per train mile) and safety or accident risk (signal passed at danger per train mile (Pollit and Smith, 2002).

Of late, psychometric measures caught the attention of many transport undertakings. The level of transport service (LOTS) was developed to measure the quality of service based on travel speed and comfort. The rating scale is from A (excellent) to F (not suitable). Each attribute was assigned a weight that reflects the importance of each attribute. The weighted score reflects the performance of that attribute. The overall performance (i.e. overall LOTS value) can be obtained to reflect the quality of service. Other measurement systems are discussed in Du Plessis (1984), McKnight et al (1986), Forsyth and Smyth (1986), Miller (1995).

Hanna and Drea (1998), and Drea and Hanna (2000) have studied the quality of service in part of the Amtrak passenger rail system in USA. Their research focus was on the attributes of service quality that influenced the transport choice of the survey respondents, e.g. rail vs. automobile. The attributes used in the first research were comfort, cost, timing (ability travel when I want), location (ability to travel where I want) and in-transit productivity (ability to work while travelling). In the later part of research cost, convenience getting to station, parking availability, Amtrak comfort, seat comfort, ride,
seating area cleanliness, and courtesy of on board staff were the service quality attributes tested.

Tripp and Drea (2002) also used a survey of Amtrak passengers to assess the “direct and indirect” relationship between pre-core/ peripheral and core service performance components and their impact on the likelihood of repeat purchase (p.433). They found that the core experiences on-board that determined the customer’s attitudes to the service provider and subsequently their intention to use the train again. These attributes included announcements, seat comfort, ride, cleanliness of seating area, courtesy of on board staff, rest rooms and café car conditions.

The gap model of service quality and the concept of service quality showed consistency that service quality should be measured on a multidimensional basis. Some transport service quality literature pointed out those different methods could be used for measuring service quality. It depends on the type of users, purposes for using the measures and the environments in which the services were provided. From this point of view, SERVQUAL is an instrument that could be used to fulfil the purpose of measuring perceived service quality from the customer’s perspective in this industry.

On the other hand, inconsistency is also found when comparing these two. They have different dimensions in conceptualization of the service quality. SERVQUAL is much more service oriented. Those commonly used in public transport industry are more industry-based. SERVQUAL is much more humanistic, or customer related, while most of the measures used in public transport industry are much more mechanistic, or have a technical focus, or use more objective measures. This led to the criticism that SERVQUAL could not tell the whole story. In Genestre and Herbig’s (1996) study, it was shown that by adding product quality to SERVQUAL, several strong product related factors are identified. In the airline industry, Young et al., (1994) added industry – based measures to SERVQUAL measures and the predictive power to satisfaction was significantly increased.
In summary, in order to measure the quality of service thoroughly, the attributes used in SERQUAL, the public transport industry, and the railway service sector should be grouped together to form a pool of items for measurement which is named as RAILQUAL instrument.

The value of adopting RAILQUAL instrument is that the diagnostic value is very high. The broad areas that are not doing well and their importance to the evaluation can be unfolded. While SERVQUAL has the ability to capture some dimensions, elements that not captured can be incorporated in the additional measures. The Gap model has already been gone through a complete building process since 1985 and was fully tested afterwards. From the literature review, it is clear that a set of attributes are needed that could be classified into different dimensions to measure service quality. Also these attributes and dimensions will need to be more context-specific than the basic SERVQUAL dimensions. The traditional measures in public transportation industry lack information about the underlying perception of customers, while SERVQUAL model is too service-oriented and lacks information about the service offering. Therefore the combination of the dimensions from these two different aspects of measuring service quality could increase the understanding of the quality construct for the railway passenger service sector.

3.6 Qualitative Research on Railway passenger services

Positivist research methodology derives from the natural sciences where objectivity, measurement, reliability and validity are emphasised (De vaus, 2002; Lee, 1992; Neuman, 2003). Positivism has evolved to encompass different approaches including logical empiricism, post-positivism and behaviourism (Neuman, 2003). Positivist research is generally based on quantitative data and derives from an objective perspective and endeavours to explain and predict occurrences in society by identifying regularities and casual relationships between events (Lee, 1992; Neuman, 2003). Quantitative research facilitates the development of clearly defined statistical relationships between dependant
Chapter 3: Contextual Research Review

Critics of positivism argue that the focus on confirming abstract relationships fails to reflect vagaries of reality (Neuman, 2003) and that the interdependence between variables can provide a unjustified clarity to the results (Lee, 1992). In contrast positivist researchers would argue that the research approach combines “deductive logic with precise empirical observations ... to discover and confirm a set of probabilistic casual laws that can be used to predict general patterns of human activity” (Neuman 2003 pg71).

Interpretive research reflects the theory of symbolic interactionism and encompasses a range of theoretical approaches including hermeneutics, phenomenology, subjectivism and ethno-methodology (Neuman, 2003). The research approach in interpretive research is normally qualitative and involves participant observation and field based research (Neuman, 2003). Although derived from the major work by Dilthey (1883) interpretivism is a relative newcomer to business research methodology and suffers from being “overshadowed by the inordinate recognition given to quantitative research ... if it’s not experimental, empirical, or statistical, it is not research” (Leedy, 1993, p140).

The third approach, critical social science, encompasses the approaches of dialectical materialism, class analysis and structuralism and is critical of both positivist and interpretive approaches (Neuman, 2003). Advocates of the critical social science approach criticise positivism for being impersonal, ignoring the social context and supporting the status quo. They equally criticise interpretive research as being too subjective and focussing on subjective reality. Critical social science endeavours to identify the underlying social structures and myths with the explanation acting as a catalyst for the change of social structures. (Neuman, 2003)

The majority of business research is encompassed with the positivist, or quantitative, and the interpretive, or qualitative, approaches and following discussion will focus exclusively on these approaches. The distinction between the research methods derives from a fundamental and diametrically opposed vies as to the best approach to understand social
From an ontological perspective the research methods contrast objectivity with subjectivity. Quantitative research adopts an objective stance and seeks to identify the precise nature of relationships among social phenomena whilst qualitative research aims to identify the motives and reasons which lead people to act in the ways they do (Lee, 1992; Neuman, 2003). From an epistemological position, the research methods contrast positivism with phenomenology, that is, the quantitative researcher seeks to decompose social world into its separate parts thus enabling the determination of casual relationships between the separate parts, whilst the qualitative researcher seeks the unfolding of social processes and the meaning of social life (Lee, 1992; Neuman, 2003). Although advocates of both approaches may take strong position on the merits of their flavoured methods the importance and validity of both are well recognised (Lee, 1992; Neuman, 2003).

### 3.7 RAILQUAL Research Context

In broad overview, this research investigates the antecedents of service quality within Indian Railways passenger services. The research investigates the direct relationship between reservation and ticketing, platform services, in train service, safety and security, punctuality, employee service with railway passenger service quality, and also investigates the influence service quality on passenger satisfaction, either directly or indirectly and of passenger satisfaction on behaviour intention.

Previous research across a number of domains has identified a sequential relationship between service quality, customer satisfaction and behaviour intentions. This research extends the previous work by investigating the relationship in the railway passenger services sector. The research included the identification of the aspects of the railway passenger service quality that are important to passengers and passenger satisfaction in relation to these aspects. The research also investigates a number of aspects that mediate between the relationship between railway passenger service quality, passenger satisfaction and behaviour intentions including perceived value.
Chapter 3: Contextual Research Review

The role of perceived value is important to evaluation of satisfaction by consumers (Bojanic, 1996; Day, 2002; Rosen and Surprenant 1998) and acts as mediating variable between service quality and customer satisfaction (Bojanic, 1996; Day, 2002; Rosen and Surprenant 1998)

3.8 Chapter Review

This chapter examined the concept of transportation sector service quality in general and railway passenger service quality in particular. The existing measurement approaches discussed. The strengths and weaknesses of SERVQUAL is observed minutely and refining and modifying the model is suggested. After that the total flaws in the SERVQUAL model elaborately presented. Finally the need of modifying the SERVQUAL to be suitable for evaluation of railway passenger service quality is presented and the procedure to develop and test a new measurement instrument RAILQUAL is described in brief. The following chapter will review about identifying and pruning the attributes which are significant predictors of railway passenger service quality with the methodology followed and its relevancy to the research context.

3.9 Chapter References


Chapter 3: Contextual Research Review


Chapter 3 : Contextual Research Review


Indian Railways Vision 2020 report presented by Indian Railways in Parliament 2009-10


Chapter 3 : Contextual Research Review

Leedy, P.D (1993), Practical research : Planning and design, New York, Macmillan


Lewis, Barbara R. and Vincent-Wayne Mitchell (1990), Defining and Measuring the Quality of Customer Service, Marketing Intelligence and Planning, 8 (6), pp11-17.


Neuman, W.L (2003), Social research methods, Qualitative and quantitative approaches, Boston, Pearson Education.


Chapter 3 : Contextual Research Review


Chapter 3 : Contextual Research Review


Sudhir Kumar,Shagun Mehrotra,(2008), Bankruptcy to Billions – How Indian Railways Transformed, Oxford Press


White Paper on Indian Railways presented by Indian Railways in Parliament 2009-10