CrossMark

Formative and reflective measurement models for analysing transit service quality

Laura Eboli[1] · **Carmen Forciniti**[1] · **Gabriella Mazzulla**[1]

Abstract Transit service quality is a complex concept depending on different service aspects, such as service frequency and punctuality, comfort, cleanliness, information and so on. Transit service quality is generally measured through the satisfaction of the users with the service. There are relationships between the overall service quality and the different transit service aspects, and between each aspect and the characteristics describing it. Structural equation models represent a useful tool for exploring this kind of relationship and determining the influence of the different service characteristics on service quality. An investigated issue concerning structural equation models is the contrast between the formative and the reflective approach. The structural models proposed for measuring transit service quality have followed a reflective approach, according to which the latent variable (or the service aspect) is the cause of the observed measures (or the service factors describing the service aspect); but in this paper we investigate on the fact that formative variables could be considered to model the relationship among the service quality characteristics, supposing that the observed measures, which represent the service characteristics, form the latent construct. The findings from the comparison between the results obtained by applying the two different approaches suggest that the reflective model is surely more suitable for describing the phenomenon of passenger satisfaction with transit service quality. However, we retain that if some service aspects can be more conveniently investigated through a reflective approach, other service aspects could follow a formative approach in a better way.

Keywords Transit service quality · Structural equation modelling · Formative and reflective models

✉ Laura Eboli
laura.eboli@unical.it

[1] University of Calabria, Rende, Italy

🍀 Springer

1 Introduction

Structural equation models (SEMs) have represented a very adequate and useful tool for analysing transit service quality, with service quality depending on a series of factors often difficult to define and measure. SEMs are included in that category of service quality measures based on customer's perspective. Customers have the right elements for appropriately judging the used service; moreover, passengers are the direct users of the service, and for this reason an analysis based on their perceptions allow the establishment of the elements retained as critical by the same users, and the possibility of improving more effectively the service (de Oña et al. 2016a, b).

Structural equation modelling (SEM) is a relatively new method whose use is rapidly expanding thanks to user-friendly software like AMOS (Arbuckle and Wothke 1995). SEM is a specific type of regression analysis that explains relationships between independent (exogenous) and dependent (endogenous) variables. It is composed of up to three sets of simultaneous equations, estimated at the same time: (1) a measurement model for the endogenous variables, (2) a measurement model for the exogenous variables, and (3) a structural model. Latent variables are constructs which cannot be directly observed, but they must be defined in terms of underlying observed variables, called indicators. A measurement model defines each latent variable, whereas the structural model represents the relationships between exogenous and endogenous variables. SEM was adopted in several fields of research and generalized by Joreskog (1973) and Wiley (1973). Some applications were proposed, for example, in the field of psychology and social science (MacCallum and Austin 2000; Muthén et al. 2006), in the field of natural science (Mitchell 1992; Grace and Pugesek 1997), and especially in the field of economy and statistics (MacLean and Gray 1998; Eskildsen and Dahlgaard 2000; Boari 2000; Manaresi et al. 2000). In the field of transportation research some applications of SEM were proposed to analyse land-use and transport interactions (e.g. Tschopp and Axhausen 2007; Van Acker et al. 2007; de Abreu e Silva and Goulias 2009; de Abreu e Silva et al. 2012; Eboli et al. 2012). Also in public transport some authors proposed SEM applications, such as Bamberg, and Schmidt (1998), Fillone et al. (2005) and Tam et al. (2005).

More specifically, SEM was adopted for investigating customer satisfaction of public transport services, but there are not many studies in this field. Examples are Andreassen (1995), Stuart et al. (2000), Karlaftis et al. (2001), Eboli and Mazzulla (2007, 2012, 2015), Ngatia et al. (2010), Irfan et al. (2011) and de Oña et al. (2013). Traditionally, the overall quality of a service is affected by many service quality factors. All the factors are generally grouped in macro-factors according to the nature of the service characteristics. As an example, service factors such as comfort of seats and air conditioning on board are included in the same macro-factor usually named "comfort". So, in a traditional SEM describing the relationship between service characteristics and the overall service there are a series of latent constructs representing the macro-factors explained each by a series of observable indicators representing the service quality factors. The dependent latent construct "overall service quality" is generally explained by observable indicators such as

"satisfaction". Definitively, the structural model describes the relationship between a latent endogenous variable representing the overall service quality and some latent exogenous variables representing the service macro-factors. On the other hand, the measurement model relating to the exogenous variables describes the relationship between the macro-factors and the service quality factors, representing the observed variables, adopted for measuring the latent constructs.

An investigated issue concerning the SEM approach is the contrast between formative and reflective models (Bollen 1989; Bollen and Lennox 1991; Diamantopoulos and Siguaw 2006). We can talk about a reflective model when the latent variable is the cause of the observed measures. Though the construct is not directly measurable, it exists independently of its effect indicators. For example, intelligence determines the responses of a subject to a questionnaire designed to assess this aspect, not vice versa. There are some aspects, however, that we cannot consider as latent constructs existing a priori: they are determined by the observed measures, which become the causes of the latent variables (Edwards and Bagozzi 2000). In these cases, the correct model is the formative one: the phenomenon is defined by, or is a function of, the observed variables (Simonetto 2012).

Literature is rich in SEMs describing the relationship mentioned above, where the variables are always reflective. In this paper, the question arises as formative variables can be considered to model the relationship between service quality and the service factors. A priori, we could retain that both approaches can represent the case of transit service quality. More specifically, according to a reflective approach, each service macro-factor (e.g. comfort, information) exists independently of the service factors which can easily describe it. On the contrary, according to a formative approach, the service quality factors could be considered as the causes of the macro-factor. In order to deepen this matter, we propose two models for highlighting the differences between the reflective and the formative approach for analysing transit service quality. Eboli and Mazzulla (2012, 2015) have recently proposed two works aimed at investigating the influence of a series of service quality attributes on the overall service quality of a railway service. In the model structure there are latent exogenous variables representing service quality macro-factors (such as safety, comfort, information and so on), each explained by a series of observed service quality attributes, and a latent endogenous variable representing the overall service quality, explained by observed global indicators. In the models proposed in the present paper we investigated about the same railway service of the above mentioned works. We retain that the proposed work could give an innovative contribution to the literature because there are no works about transit service quality investigating the differences between the two mentioned approaches; instead, we think that there are many differences among the various service quality aspects, which could be better analyses through one of the approaches rather than the other one.

In the following, we propose a review focusing on the studies adopting the reflective approach rather than the formative one. After the review, we propose a section describing the model structures adopted for analysing service quality, based on the use of reflective and formative approaches: a brief theoretical framework is introduced before the description of the specific framework. Then, there is a section

about the application of the proposed models, where we briefly describe the data, and finally we present the results of the models. The paper ends with the conclusions about the work.

2 Formative versus reflective models

The distinction between formative and reflective measures is important because proper specification of a measurement model is necessary to assign meaningful relationships in the structural model (Anderson and Gerbing 1988). The reflective approach dominates the psychological and management sciences, while the formative approach is mostly adopted in economics and sociology. According to the reflective measurement models, causality flows from the latent construct to the indicator. This means that the change in an indicator X reflects the change in the latent construct Y. Scholars argue that if variation in X is associated with variation in Y, exogenous interventions that change Y can be detected in the indicator X; they assume that this relationship between construct and indicator is reflective (Borsboom et al. 2003, 2004). Practically all scales in business and related methodological texts on scale development use a reflective approach to measurement (Bearden and Netmeyer 1999; Bruner et al. 2001; Netmeyer et al. 2003; Spector 1992).

On the other hand, several researchers think that not all latent constructs are measurable with a series of correlated items (Bollen and Lennox 1991; Edwards and Bagozzi 2000; Fornell 1982). Where causality flows in the opposite direction, from the indicator to the construct, a formative index results (Blalock 1964; Diamantopoulos and Winklhofer 2001; Edwards and Bagozzi 2000). In other words, more indicators form a construct without any assumptions concerning the patterns of inter-correlation between these items. In contrast to the reflective model, few examples of formative models are seen in the business literature.

Coltman et al. (2008) present an organizing framework for construct measurement that begins with theoretical justification to define the nature of the focal constructs. There are three theoretical considerations in deciding whether the measurement model is formative or reflective: (1) the nature of the construct, (2) the direction of causality between the indicators and the latent construct, and (3) the characteristics of the indicators used to measure the construct.

Concerning the nature of the construct, the scientific literature suggests that in a reflective model the latent construct exists independently from the measures (Borsboom et al. 2004; Rossiter 2002). On the contrary, in a formative model the latent construct is dependent on the interpretation by the scholar (Borsboom et al. 2003). Examples of formative constructs are the indexes composed by a series of indicators; any change in one or more of these components causes a change in the index value.

According to the direction of causality, which is the second theoretical consideration, reflective models assume that causality flows from the construct to the indicators, while in formative models causality flows from the indicators to the construct. This means that in reflective models, a change in the construct causes a

change in the indicators, while in formative models, a change in the indicators results in a change in the construct (Fig. 1).

Finally, the theoretical consideration regarding the characteristics of indicators suggests that in a reflective model, change in the latent variable must precede variation in the indicator(s). This means that the indicators all share a common theme and are interchangeable; inclusion or exclusion of one or more indicators from the domain does not alter the validity of the construct. On the contrary, in formative models, the domain of the construct is sensitive to the number and types of indicators representing the construct; therefore, adding or removing an indicator can change the conceptual domain of the construct.

3 Methodology

3.1 The model structure

In order to understand better the SEM structure, we briefly report the basic equations introduced by Bollen (1989). Specifically, for the latent variable model the basic equation is the following:

$$\eta = B\eta + \Gamma\xi + \zeta \tag{1}$$

in which η (eta) is an ($m \times 1$) vector of the latent endogenous variables, ξ (xi) is an ($n \times 1$) vector of the latent exogenous variables, and ζ (zeta) is an ($m \times 1$) vector of random variables. The elements of the B (beta) and Γ (gamma) matrices are the structural coefficients of the model; the B matrix is an ($m \times m$) coefficient matrix for the latent endogenous variables; the Γ matrix is an ($m \times n$) coefficient matrix for the latent exogenous variables.

The basic equation of the measurement model for the exogenous variables is Eqs. (2), and (3) is for the endogenous variables:

$$x = \Lambda_x\xi + \delta \tag{2}$$

$$y = \Lambda_y\eta + \varepsilon \tag{3}$$

in which x and δ (delta) are column q-vectors related to the observed exogenous
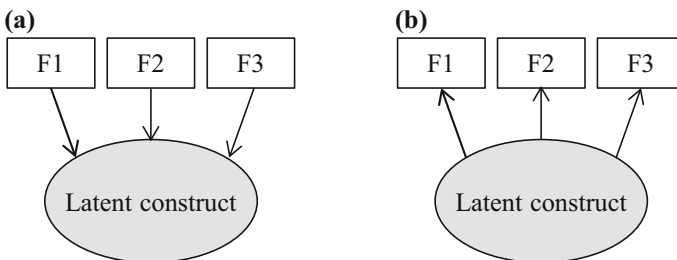

Fig. 1 Formative (**a**) versus reflective (**b**) measurement models

variables and errors, respectively; Λ_x (lambda) is a $(q \times n)$ structural coefficient matrix for the effects of the latent exogenous variables on the observed variables; y and ε (epsilon) are column p-vectors related to the observed endogenous variables and errors, respectively; Λ_y is a $(p \times m)$ structural coefficient matrix for the effects of the latent endogenous variables on the observed ones.

In the proposed model, we supposed the presence of seven latent exogenous constructs representing seven main characteristics of a railway service: safety, cleanliness, comfort, service, additional services, information and personnel. These seven constructs represent the seven macro-factors investigated through the survey addressed to a sample of users of the railway service analysed in the paper. These latent factors are supposed to be linked to a latent construct representing the overall service quality, simulated by a latent endogenous variable, named as service quality. Overall transit service quality generally contains a large number of service factors. Starting from the various studies reported in the literature of the field, among which there are the European Standard EN13816 (CEN 2002) or the manuals of the Transportation Research Board such as Transportation Research Board (2003), we can state that the aspects mainly characterizing bus services are service availability, service reliability, comfort, cleanliness, safety and security, fare, information, customer care and environmental impacts. The seven factors chosen for the proposed study well reflect the selection of the service factors emerging from the literature review. Each latent construct is explained by really observed service quality factors; so, the latent endogenous variables are linked to 33 observed variables, while the latent endogenous variable is linked to two observed indicators of global service quality. The first one is simply the satisfaction rate expressed by each user on the overall service (satisfaction). The second one is represented by the number of factors for which the user has experienced problems in the last 30 days before the interview (critical event).

3.2 The reflective approach

By adopting a reflective approach and by considering the theoretical assumptions described in the previous section, we can apply the following reasoning.

We can assume that the latent constructs, which represent service macro-factors (safety, cleanliness, comfort, service, additional services, information, and personnel) exist independently of the measures, which are the various service factors included in each macro-factor. So, as an example, safety exists independently of the attributes that measure it, such as travel safety, personal security on board, and personal security at station (Fig. 2). As well as, comfort is an existing concept representing a service aspect independently of the service characteristics adopted for measuring it, linked to crowding on board, air-conditioning on board, and so on (Fig. 3).

In terms of direction of causality, we remind that the reflective approach assumes that causality flows from the construct to the indicators, and therefore from the service macro-factor to the characteristics describing it. Hence, a change in the service macro-factor causes a change in the service factor. As an example, the level
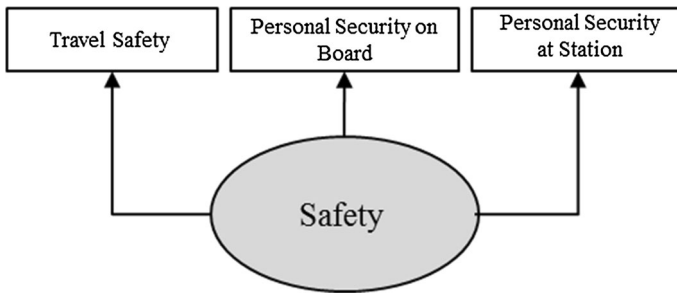
Fig. 2 An example of a reflective measurement model for "safety" latent construct
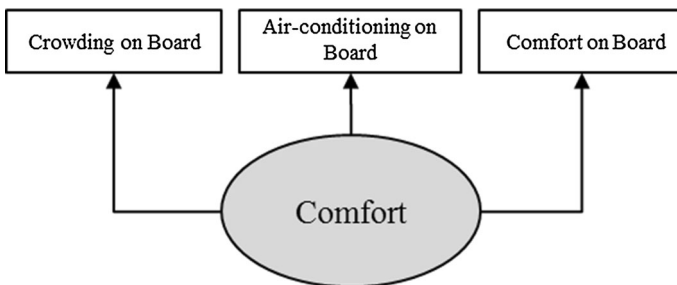

Fig. 3 An example of reflective measurement model for the "comfort" latent construct

of satisfaction about the comfort of the passengers travelling by the transit system influences the level of satisfaction with the comfort on board, comfort of seats and the other characteristics linked to comfort.

Another aspect regards the characteristics of the indicators, which share a common theme and are interchangeable. As an example, the indicators explaining information, such as information at station or information timeliness on board, share the same theme, and the inclusion or exclusion of one of them does not materially alter the content validity of the construct.

3.3 The formative approach

From a different perspective, by adopting a formative approach, we assume that the service macro-factors are dependent on the factors that form them. So, by considering the same example of safety, the formative approach states that safety does not exist as an independent entity but it depends on travel safety, personal security on board and personal security at station. In the same way, we could consider that the concept of comfort is strictly dependent on the various characteristics that forms it. In other words, the meaning of comfort changes based on the service characteristics considered for describing it. By making this reasoning, the service macro-factor could be better included in a formative model. Another interesting example is represented by the service macro-factor "service" (Fig. 4). It seems that this macro-factor better adapts to a formative approach, because its
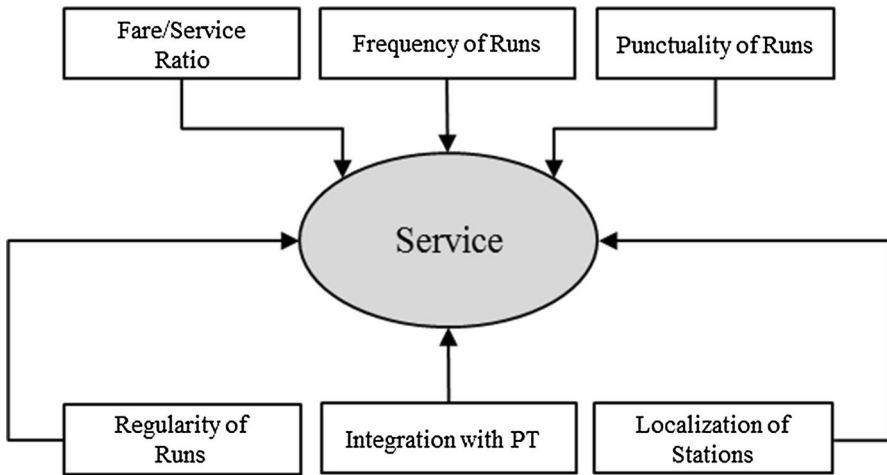
Fig. 4 An example of formative measurement model for the "service" latent construct

concept is not clearly defined like "safety" or "comfort", but it can represent more characteristics which are heterogeneous among them. In our model, this aspect includes service factors such as frequency, punctuality and regularity of runs, as well as aspects such as fare, localization of station and integration with public transport. As we can observe, the meaning of this aspect is formed by the analyst, and it cannot be considered as existing a priori.

Consequently, causality flows from the indicators to the construct; therefore, a change in the service attributes results in a change in the service macro-factor under study. As an example, a change in the satisfaction about cleanliness of seats or of a vehicle, as well as cleanliness of toilets or stations, change the level of satisfaction about the "cleanliness" macro-factor.

In this case, since the indicators define the construct, the domain of the service macro-factor is sensitive to the number and types of service factors representing it. Hence, adding or removing a service attribute can change the conceptual domain of the service macro-factor. As an example, the domain of the macro-factor "personnel" changes if we remove one or more of the service factors describing
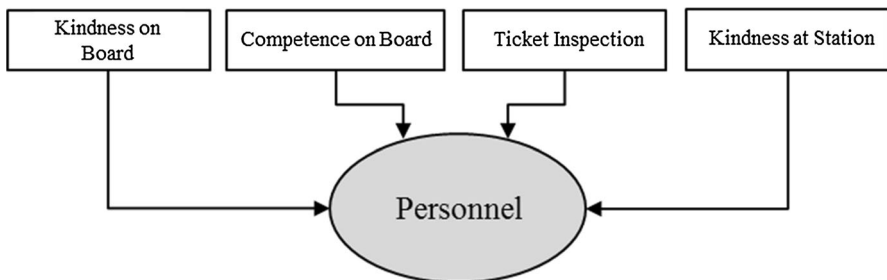

Fig. 5 An example of a formative measurement model for "personnel" latent construct

it, such as kindness on board or in station, rather than competence on board or ticket inspection (Fig. 5).

4 Transit service quality modelling: formative versus reflective approach

4.1 Data

The data adopted for calibrating the proposed models were collected through a survey addressed to passengers of a railway service operating in the north of Italy. The service offers 32 regional lines and 9 suburban lines connecting the towns of the hinterland of the city of Milan, and 2 express lines connecting Milan with the Malpensa airport. On board face-to-face interviews were realized from June to July 2011, during the whole week, between 6.00 a.m. and 10.00 p.m. The final sample contains 16,718 passengers. The number of passengers travelling on the analysed lines is about 550,000. So, the sample rate is of about 3%.

The questionnaire aimed to collect data concerning general information about the train (e.g. time period of the interview, train, line, station), socio-economic characteristics and travel habits of the passengers (e.g. gender, age, professional condition, income, trip scope, frequency), and the passenger perceptions about the used services. Specifically, 33 service quality factors concerning safety, cleanliness, main and additional services, information, and personnel were evaluated by the users, who expressed importance and satisfaction rates on a cardinal scale from 1 to 10. Users also indicated the occurrence of a critical event concerning each service quality factor in the last 30 days before the interview. Finally, they expressed a satisfaction rate about the overall service.

More than half of the interviewed passengers travel by the regional lines, about 40% suburban lines, and the remaining 7% through the Malpensa express services. The major part of passengers (80.5%) were interviewed in a weekday, 14.2% in a before a holiday day, and 5.3% in a holiday. 30.6% of sample travels to work, 13.6% to study, and the remaining 55.8% for other purposes, such as personal activities (about 30%), bureaucratic activities (about 10%), and tourism (about 15%). The sample is almost equally spread between females and males. The major part of the interviewed passengers (more than 70%) are aged lower than 40 years. Most of the sampled people are employed (60%), and a considerable part is composed of students (about 30%). About 35% of sample does not give any kind of information about the income, but most people belong to a class of income lower than 2000 Euros per month (56%). Almost 65% of the sample are habitual users, while 35% of the passengers occasionally travel by train. Definitively, the prevalent passenger is a young employed, with a monthly net income lower than 2000 Euros, who habitually travels by train to reach the place of work (Table 1).

Regarding the importance and satisfaction rates expressed by the interviewed users, we can highlight that all the attributes registered average importance rates in a small range varying from 8.0 to 9.2; this is not a surprising result because, as found in many other scientific studies, users tend to assign high importance to all the

Table 1 Sample characteristics

Characteristics	Statistics
Gender	Male (45.5%), female (54.5%)
Age	16–25 (36.7%), 26–40 (33.8%), 41–65 (25.6%), >65 years-olds (3.9%)
Professional condition	Employee (40.5%), manager (4.2%), entrepreneur (2.1%), freelancer (8.2%), self-employed worker (3.7%), unemployed (3.4%), student (30.5%), housewife (2.4%), pensioner (4.5%), other (0.6%)
Income level	≤1000 (25.0%), 1001–1500 (19.3%), 1501–2000 (12.0%), 2001–3000 (4.9%), 3001–4000 (1.6%), >3000 Euros (2.2%)
Car ownership	Have car (58.1%), have not car (41.9%)
Scope of journey	Work (30.6%), studying (13.6%), other purposes (55.8%)
Frequency of journey	Daily (46.8%), weekly (18.0%), occasionally (35.2%)

factors when they have to state a rate of importance (de Oña et al. 2013). The attributes considered as the most important concern travel safety and personal security. On the contrary, the average satisfaction rates suggest that people judge as not very good most of the service characteristics; in fact, only nine attributes have an average rate higher than six (>6). The service characteristics considered as the most satisfying concern safety, personnel, integration with the other modes of public transport, and localization of the stations. On the contrary, the characteristics judged as the least satisfying concern cleanliness, comfort, and information. Users expressed an average rate of satisfaction with the overall service equal to 5.8.

Finally, some observations can be made about the critical events registered by users. Only the attribute linked to personal security on board registered a very high number of critical events in the last 30 days before the interview (14,715 occurrences out of 16,623 interviewed users), although the satisfaction rate was higher than six; evidently, they feel quite safe on board, and most of the experienced problems are not substantial for them. For aspects such as cleanliness of vehicles, seats and toilet facilities, crowding and air-conditioning on board, and punctuality of runs, the number of registered critical events is notable (over than 6000 occurrences), but for the other service aspects the number of occurrences is relatively less important (about 2000 or 3000 occurrences).

4.2 Models

In this section, we show the results of the models with the final aim to explain the differences between the two adopted approaches.

Figures 6 and 7 show the complete theoretical construct of the reflective model and the formative one, respectively.

Table 1 shows results of the model structured according to the reflective approach. In the first and second column the model variables are reported, with the indication of the direction concerning the relationship between the two categories of variables (endogenous and exogenous); the third column shows the values of the regression weights (RW) of the coefficients; the fourth and fifth column contain the
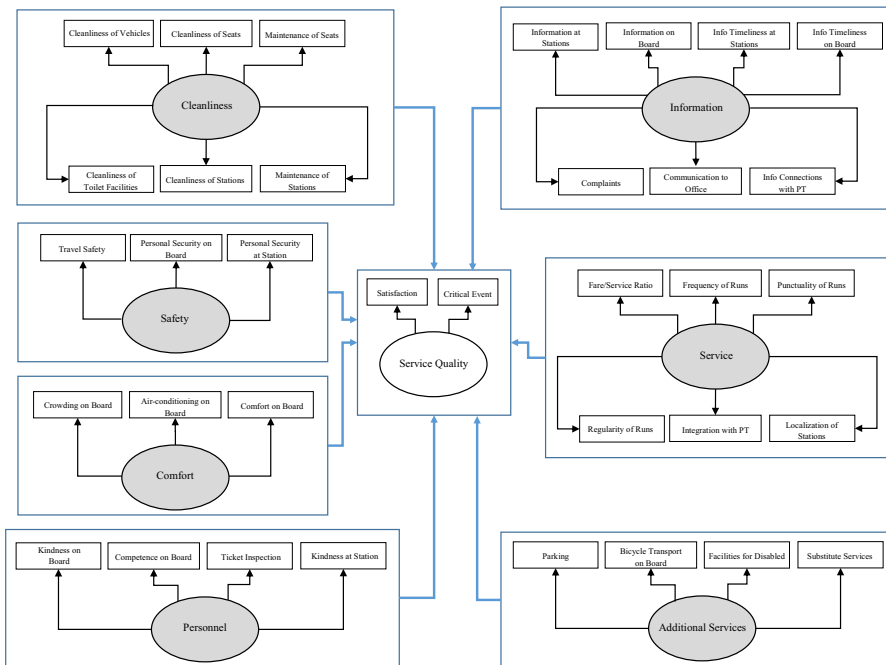
Fig. 6 Theoretical construct of the reflective model

values of the standard error (SE) of each coefficient and the probability levels (P) that the estimated coefficient is significantly different from zero; finally, in the last column the values of the standardized regression weights (st. RW) are shown.

Before describing the contents of Table 1, we report the values of the statistics on the goodness of fit. From our proposed model, we obtained a minimum value of discrepancy function statistically significant according to the Chi squared test. Because the Chi squared test of absolute model fit is sensitive to sample size and non-normality in the underlying distribution of the input variables, various descriptive fit statistics may be used to assess the overall fit of a model to the data. The criteria that researchers often use for verifying if a model is regarded as good are the following, even if Bollen (1989) suggests that these criteria are merely guidelines. The Normed Fit Index (NFI) exceeds 0.90 (Byrne 1994) or 0.95 (Hu and Bentler 1999; Schumacker and Lomax 2004); the Goodness of Fit Index exceeds 0.90 (Byrne 1994; Carreira et al. 2014); the Adjusted Goodness of Fit Index exceeds 0.90 (Carreira et al. 2014); the Comparative Fit Index exceeds 0.93 (Byrne 1994), but some authors report that a CFI ≥0.8 is good enough for structural validity of the model (Browne and Cudeck 1993; Hair et al. 2009); RMS is less than 0.08 (Browne and Cudeck 1993; Hu and Bentler 1999) and ideally less than 0.05 (Steiger 1990); RMSEA is lower than 0.08 for a very good fit (Carreira et al. 2014) and it provides a mediocre fit between 0.08 and 0.10 (MacCallum et al. 1996). The relative Chi square should be less than 2 or 3 (Kline 1998; Ullman 2001).
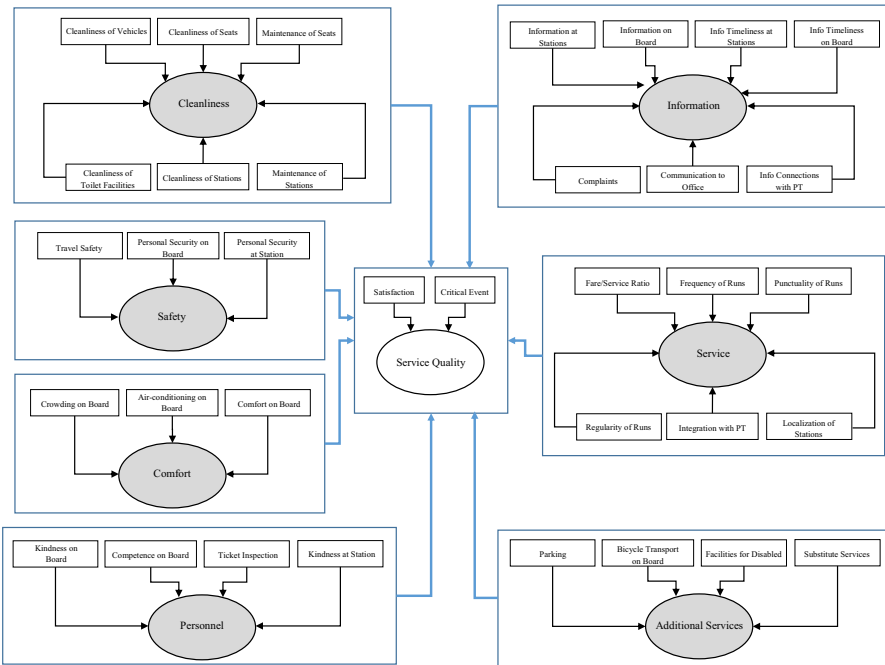
Fig. 7 Theoretical construct of the formative model

In order to obtain better values of goodness-of-fit, as suggested by some authors in the literature, we have inserted in the model structure some correlations among the errors of the service attributes belonging to the same macro-factor which had shown a certain correlation of the satisfaction rates. For our reflective model we obtained a GFI equal to 0.84, an AGFI of 0.80, and a CFI of 0.90. The RMSEA has a value of 0.07 (Table 2). By observing the values concerning our model, and by considering the suggestions reported in the literature, we can state that our model shows good values of goodness of fit. More specifically, we retain that the CFI obtained for our model can be considered as very good, as well as the obtained value of RMSEA.

Table 2 Goodness of fit indexes

Indexes	Values for reflective model	Values for formative model
Goodness of Fit Index (GFI)	0.84	0.61
Comparative Fit Index (CFI)	0.90	0.72
Adjusted Goodness of Fit Index (AGFI)	0.80	0.48
Root mean square error of approximation (RMSEA)	0.07	0.10

All the parameters have a correct sign and assume a statistically different from zero value, at a good level of significance, with the exception of the coefficient of the latent variable representing "additional service". The latent exogenous variables with the highest positive effect on service quality are service (0.380) and cleanliness (0.365). Directly after, there are two service aspects presenting similar coefficients: information (0.312) and comfort (0.297). Personnel shows a weight equal to 0.207. Finally, safety and additional services have the lowest effect on service quality, showing weights of 0.081 and 0.011, respectively. The relationship between the latent endogenous variable and observed variables shows that satisfaction explains the major part of service quality, having a weight of 0.572, but also the critical event indicator has a relevant influence (−0.394) (Table 3).

Table 2 shows results of the model structured according to the formative approach. Before analysing the coefficients, we have to specify that in this case the tests of goodness of fit show not ideal values. In fact, the GFI is 0.61, the AGFI is 0.48, the CFI is 0.72, the RMSEA is 0.10 (Table 2). However, we can surely consider the values of CFI and RMSEA as acceptable. We also observed that not all the parameters are statistically significant, and some of them have not the expected sign. There are three service factors which show a negative sign: cleanliness of stations (−0.039), localization of stations (-0.195), and bicycle transport on board (−0.542), meaning that if satisfaction with these factors increase, the overall satisfaction decreases. However, cleanliness of stations and bicycle transport on board are not significant at a good level of significance.

In the following, we want to analyse the coefficients of some service aspects with the aim to have a comparison between the two kinds of approaches and to confirm that the reflective approach could be not very suitable for describing the relationship concerning some latent constructs. More specifically, by observing the weights of the latent exogenous variables, we can say that also in this case the service aspects having the highest positive effect on service quality are service (0.514) and cleanliness (0.392), even if they show more difference of weight. Directly after, there are three service aspects presenting similar coefficients: comfort (0.330), information (0.278) and personnel (0.246). Also according the formative approach, safety and additional services have the lowest effect on service quality, showing weights of 0.110 and 0.103, respectively. Analogously to the reflective approach, the relationship between the latent endogenous variable and observed variables shows that satisfaction explains the major part of service quality, showing a weight of 0.437, but also the critical event indicator has a relevant influence (−0.284) (Table 4).

5 Discussion of the results

The most interesting aspect of the proposed models is represented by the relationship between the latent exogenous variables and their observed indicators, because the difference between the two investigated approaches concerns just the nature of the relationship between the latent constructs and the observed indicators.

Table 3 Model results

			RW	SE	P	St. RW
Latent endogenous variable		Latent exogenous variable				
Service quality (η_1)	$\Leftarrow$	Safety (ξ_1)	0.047**	0.009	0.000	0.081
Service quality (η_1)	$\Leftarrow$	Cleanliness (ξ_2)	0.166**	0.011	0.000	0.365
Service quality (η_1)	$\Leftarrow$	Comfort (ξ_3)	0.186**	0.017	0.000	0.297
Service quality (η_1)	$\Leftarrow$	Service (ξ_4)	0.221**	0.012	0.000	0.380
Service quality (η_1)	$\Leftarrow$	Additional services (ξ_5)	0.007	0.012	0.533	0.011
Service quality (η_1)	$\Leftarrow$	Information (ξ_6)	0.210**	0.012	0.000	0.312
Service quality (η_1)	$\Leftarrow$	Personnel (ξ_7)	0.119**	0.010	0.000	0.207
Observed endogenous variable		Latent exogenous variable				
F1 (x_1) travel safety	$\Leftarrow$	Safety (ξ_1)	1	–	–	0.760
F2 (x_2) personal security on board	$\Leftarrow$	Safety (ξ_1)	1.126**	0.017	0.000	0.828
F3 (x_3) personal security at station	$\Leftarrow$	Safety (ξ_1)	1.127**	0.020	0.000	0.816
F4 (x_4) cleanliness of vehicles	$\Leftarrow$	Cleanliness (ξ_2)	1	–	–	0.904
F5 (x_5) cleanliness of seats	$\Leftarrow$	Cleanliness (ξ_2)	1.007**	0.004	0.000	0.905
F6 (x_6) maintenance of seats	$\Leftarrow$	Cleanliness (ξ_2)	1.011**	0.006	0.000	0.916
F7 (x_7) cleanliness of toilet facilities	$\Leftarrow$	Cleanliness (ξ_2)	0.859**	0.006	0.000	0.808
F8 (x_8) cleanliness of stations	$\Leftarrow$	Cleanliness (ξ_2)	0.759**	0.007	0.000	0.716
F9 (x_9) maintenance of stations	$\Leftarrow$	Cleanliness (ξ_2)	0.713**	0.007	0.000	0.676
F10 (x_{10}) crowding on board	$\Leftarrow$	Comfort (ξ_3)	1	–	–	0.659
F11 (x_{11}) air-conditioning on board	$\Leftarrow$	Comfort (ξ_3)	1.168**	0.016	0.000	0.752
F12 (x_{12}) comfort on board	$\Leftarrow$	Comfort (ξ_3)	1.170**	0.016	0.000	0.829
F13 (x_{13}) fare/service ratio	$\Leftarrow$	Service (ξ_4)	1	–	–	0.698
F14 (x_{14}) frequency of runs	$\Leftarrow$	Service (ξ_4)	1.062**	0.012	0.000	0.775
F15 (x_{15}) punctuality of runs	$\Leftarrow$	Service (ξ_4)	1.128**	0.014	0.000	0.771
F16 (x_{16}) regularity of runs	$\Leftarrow$	Service (ξ_4)	1.030**	0.013	0.000	0.759
F17 (x_{17}) integration with pt	$\Leftarrow$	Service (ξ_4)	0.890**	0.011	0.000	0.706
F18 (x_{18}) localisation of stations	$\Leftarrow$	Service (ξ_4)	0.833**	0.011	0.000	0.675
F19 (x_{19}) parking	$\Leftarrow$	Additional services (ξ_5)	0.803**	0.015	0.000	0.536
F20 (x_{20}) bicycle transport on board	$\Leftarrow$	Additional services (ξ_5)	0.752**	0.012	0.000	0.635
F21 (x_{21}) facilities for disabled	$\Leftarrow$	Additional services (ξ_5)	1.023**	0.013	0.000	0.727
F22 (x_{22}) substitute services	$\Leftarrow$	Additional services (ξ_5)	1	–	–	0.792
F23 (x_{23}) information at stations	$\Leftarrow$	Information (ξ_6)	1.203**	0.012	0.000	0.806
F24 (x_{24}) information on board	$\Leftarrow$	Information (ξ_6)	1.345**	0.013	0.000	0.876
F25 (x_{25}) info timeliness at stations	$\Leftarrow$	Information (ξ_6)	1.343**	0.013	0.000	0.886
F26 (x_{26}) info timeliness on board	$\Leftarrow$	Information (ξ_6)	1.325**	0.013	0.000	0.867

Table 3 continued

			RW	SE	P	St. RW
F27 (x_{27}) complaints	$\Leftarrow$	Information (ξ_6)	1.041**	0.012	0.000	0.700
F28 (x_{28}) communication to office	$\Leftarrow$	Information (ξ_6)	0.995**	0.009	0.000	0.707
F29 (x_{29}) info connections with PT	$\Leftarrow$	Information (ξ_6)	1	–	–	0.705
F30 (x_{30}) kindness on board	$\Leftarrow$	Personnel (ξ_7)	1.022**	0.012	0.000	0.847
F31 (x_{31}) competence on board	$\Leftarrow$	Personnel (ξ_7)	1.010**	0.009	0.000	0.860
F32 (x_{32}) ticket inspection	$\Leftarrow$	Personnel (ξ_7)	0.947**	0.011	0.000	0.701
F33 (x_{33}) kindness at station	$\Leftarrow$	Personnel (ξ_7)	1	–	–	0.774
Observed endogenous variable		Latent endogenous variable				
Satisfaction (y_1)	$\Leftarrow$	Service quality (η_1)	1	–	–	0.572
Critical event (y_2)	$\Leftarrow$	Service quality (η_1)	−3.089**	0.073	0.000	−0.394

Reflective approach

** Statistically significant at a level of 5%

Starting from the service aspects resulting as the most important for the users (service and cleanliness), we can observe that the coefficients of the reflective model are higher than the coefficients of the formative model. According to the reflective approach, service characteristics are well explained by the service factors linked to the reliability of the service: frequency of runs (0.775), punctuality (0.771) and regularity (0.759); also integration between the services analysed and other local transit services presents a relevant weight (0.706); finally, the factors linked to the fare and localization of station have the lowest weights (0.698 and 0.675, respectively). According to the formative approach, punctuality is the factor mostly influencing service quality, showing a coefficient significantly higher than the others (0.671). Except the factor linked to the fare which obtained a coefficient equal to 0.495, the other factors presents low weights.

According to the reflective model, cleanliness is prevalently understood as maintenance of the seats (0.916), cleanliness of seats (0.905) and of vehicles (0.904); the other indicators have lower weights, up to a minimum value of 0.676, registered for the attribute linked to the maintenance of the stations. So, for the users cleanliness represents the cleanliness on board, it means cleanliness of seats and vehicles. A similar result is obtained from the formative approach, even if cleanliness of vehicles (0.442) and of the seats (0.423) are prevalent as regards maintenance of the seats (0.355).

The successive service aspects affecting the satisfaction about the service are information and comfort, for both the approaches. According to the reflective approach, comfort is prevalently interpreted by the passengers as degree of comfort on board (0.829) and air-conditioning (0.752), even if the level of crowding on board presents a relevant value of the coefficient (0.659). We can conclude that for bus services the level of crowding on board is the indicator mostly used by the researchers who analyse service quality in public transport, but this factor is less

Table 4 Model results

			RW	SE	P	St. RW
Latent endogenous variable		Latent exogenous variable				
Service quality (η_1)	$\Leftarrow$	Safety (ξ_1)	0.021**	0.005	0.000	0.110
Service quality (η_1)	$\Leftarrow$	Cleanliness (ξ_2)	0.047**	0.009	0.000	0.392
Service quality (η_1)	$\Leftarrow$	Comfort (ξ_3)	0.013**	0.004	0.003	0.330
Service quality (η_1)	$\Leftarrow$	Service (ξ_4)	0.070**	0.005	0.000	0.514
Service quality (η_1)	$\Leftarrow$	Additional services (ξ_5)	0.009*	0.005	0.070	0.103
Service quality (η_1)	$\Leftarrow$	Information (ξ_6)	0.020**	0.007	0.004	0.278
Service quality (η_1)	$\Leftarrow$	Personnel (ξ_7)	0.018**	0.009	0.048	0.246
Observed endogenous variable		Latent exogenous variable				
Safety (ξ_1)	$\Leftarrow$	F1 (x_1) travel safety	1	–	–	0.642
Safety (ξ_1)	$\Leftarrow$	F2 (x_2) personal security on board	0.669	0.488	0.170	0.385
Safety (ξ_1)	$\Leftarrow$	F3 (x_3) personal security at station	0.448	0.279	0.109	0.302
Cleanliness (ξ_2)	$\Leftarrow$	F4 (x_4) cleanliness of vehicles	1	–	–	0.442
Cleanliness (ξ_2)	$\Leftarrow$	F5 (x_5) cleanliness of seats	0.991**	0.389	0.011	0.423
Cleanliness (ξ_2)	$\Leftarrow$	F6 (x_6) maintenance of seats	0.828**	0.183	0.000	0.355
Cleanliness (ξ_2)	$\Leftarrow$	F7 (x_7) cleanliness of toilet facilities	0.190	0.138	0.168	0.078
Cleanliness (ξ_2)	$\Leftarrow$	F8(x_8) cleanliness of stations	−0.096	0.088	0.280	−0.039
Cleanliness (ξ_2)	$\Leftarrow$	F9 (X_9) maintenance of stations	0.838**	0.244	0.000	0.345
Comfort (ξ_3)	$\Leftarrow$	F10 (x_{10}) crowding on board	1	–	–	0.142
Comfort (ξ_3)	$\Leftarrow$	F11 (x_{11}) air-conditioning on board	4.300**	1.669	0.010	0.570
Comfort (ξ_3)	$\Leftarrow$	F12 (x_{12}) comfort on board	3.654**	1.180	0.002	0.496
Service (ξ_4)	$\Leftarrow$	F13 (x_{13}) fare/service ratio	1	–	–	0.495
Service (ξ_4)	$\Leftarrow$	F14 (x_{14}) frequency of runs	1.167**	0.080	0.036	0.077
Service (ξ_4)	$\Leftarrow$	F15 (x_{15}) punctuality of runs	1.375**	0.116	0.000	0.671
Service (ξ_4)	$\Leftarrow$	F16 (x_{16}) regularity of runs	0.426**	0.094	0.000	0.194
Service (ξ_4)	$\Leftarrow$	F17 (x_{17}) integration with PT	0.273**	0.091	0.003	0.115
Service (ξ_4)	$\Leftarrow$	F18 (x_{18}) localisation of stations	−0.458**	0.092	0.000	−0.195
Additional services (ξ_5)	$\Leftarrow$	F19 (x_{19}) parking	1	–	–	0.283
Additional services (ξ_5)	$\Leftarrow$	F20 (x_{20}) bicycle transport on board	−2.459**	1.211	0.042	−0.542
Additional services (ξ_5)	$\Leftarrow$	F21 (x_{21}) facilities for disabled	3.116*	1.784	0.081	0.805
Additional services (ξ_5)	$\Leftarrow$	F22 (x_{22}) substitute services	0.920	0.887	0.299	0.222
Information (ξ_6)	$\Leftarrow$	F23 (x_{23}) information at stations	1	–	–	0.235

Table 4 continued

			RW	SE	P	St. RW
Information (ξ_6)	$\Leftarrow$	F24 (x_{24}) information on board	1.053	0.646	0.103	0.270
Information (ξ_6)	$\Leftarrow$	F25 (X_{25}) INFO TIMELINESS AT STATIONS	2.513**	1.002	0.012	0.641
Information (ξ_6)	$\Leftarrow$	F26 (x_{26}) info timeliness on board	1.008**	0.470	0.032	0.233
Information (ξ_6)	$\Leftarrow$	F27 (x_{27}) complaints	0.432	0.340	0.203	0.098
Information (ξ_6)	$\Leftarrow$	F28 (x_{28}) communication to office	0.673	0.455	0.139	0.130
Information (ξ_6)	$\Leftarrow$	F29 (x_{29}) info connections with PT	0.927**	0.433	0.032	0.205
Personnel (ξ_7)	$\Leftarrow$	F30 (x_{30}) kindness on Board	1	–	–	0.233
Personnel (ξ_7)	$\Leftarrow$	F31 (x_{31}) competence on board	1.887	0.441	0.191	0.420
Personnel (ξ_7)	$\Leftarrow$	F32 (x_{32}) ticket inspection	1.239*	0.652	0.058	0.313
Personnel (ξ_7)	$\Leftarrow$	F33 (x_{33}) kindness at station	2.183*	1.164	0.061	0.540
Observed endogenous variable		Latent endogenous variable				
Satisfaction (y_1)	$\Leftarrow$	Service quality (η_1)	1	–	–	0.437
Critical event (y_3)	$\Leftarrow$	Service quality (η_1)	−3.078**	0.114	0.000	−0.284

Formative approach

* Statistically significant at a level of 10%

** Statistically significant at a level of 5%

important for a service such as the rail service. Analogous conclusions can be reached by observing the results of the formative model, because also in this case comfort on board and air-conditioning are retained as more important than crowding on board, but air-conditioning results the most important attribute with a coefficient of 0.570, followed by comfort on board presenting a coefficient of 0.496.

If we observe the results of the reflective model, Information is prevalently explained by the timeliness in having the information both at stations and on board (having very similar coefficients equal to 0.886 and 0.867, respectively), and the quality of information about the services (0.876 at station and 0.806 on board). The other aspects linked to the complaints or the communication to the office are considered as less relevant. According to the formative approach, timeliness in having the information at stations is the factor mostly influencing service quality, showing a coefficient significantly higher than the others (0.641). However, also in this case timeliness in having the information on board and quality of information about the services on board and at station are considered as more important than the other secondary aspects, even if they show values of coefficients that are rather low as regards the timeliness in having the information at stations.

Concerning personnel characteristics we can observe that, according to a reflective approach, users consider as more important the factors linked to the personnel on board; in fact, competence and kindness of the personnel on board present coefficient values equal to 0.860 and 0.847, respectively. On the contrary, the formative model suggests that personnel is best explained by the kindness at stations, which shows the highest coefficient value (0.540).

Finally, also the analysis of the less important service aspects suggests interesting findings. Safety is mostly explained by the service factors concerning security; we obtained a coefficient of 0.828 for personal security on board and 0.816 for personal security at stations from the reflective model. These results are different from the results of the formative model according to which safety is mostly represented by travel safety, showing a coefficient of 0.642. additional services are moreover understood as services substitute of irregular services (0.792), besides as facilities for disabled (0.727), which is the most important additional service factor for the formative model (0.805).

6 Conclusions

Transit service quality has been widely analysed through structural equation models. An investigated issue concerning the SEM approach has been the contrast between formative and reflective models. In this paper we wanted to investigate the opportunity to consider both the reflective and the formative approach for analysing transit service quality, although in the literature the reflective approach has been largely adopted rather than the formative one.

By adopting a reflective approach we assumed that the latent constructs, which represent service macro-factors (safety, cleanliness, comfort, service, additional services, information, and personnel) exist independently from the various service factors describing each macro-factor. On the other hand, by adopting a formative approach, we assumed that the service macro-factors depend on the factors that form them.

From the application of the two approaches, we obtained interesting findings. Specifically, we obtained similar results from the two models for those service aspects which should be analysed better with a reflective approach, evidently because they represent concepts that are well defined a priori (examples are cleanliness and comfort). On the contrary, for the service aspects for which a formative approach would seem more convenient, the results of the two approaches are rather discordant or not similar; so in this case, a formative approach could be more convenient. Surely, the aspects more suitable for the formative approach are the service aspects which represent more characteristics being heterogeneous among them, such as the macro-factors "service" and "personnel". The results concerning the less important aspects (safety and additional services) show differences, but there are many variables that are less significant in the formative model, so we could avoid to consider them for our findings.

Definitively, we can conclude that the reflective approach is most suitable for analysing transit service quality because of the obtained values of goodness-of-fit, but both methods should be considered for analysing transit service quality, because there are some service aspects which can be more conveniently investigated through a reflective approach, but other service aspects follow very well a formative approach.

References

Anderson JC, Gerbing DW (1988) Structural equation modeling in practice: a review and recommended twostep approach. Psychol Bull 103(3):411–423

Andreassen TW (1995) (Dis)satisfaction with public services: the case of public transportation. J Serv Mark 9:30–41

Arbuckle JL, Wothke W (1995) AMOS 4.0 user's guide. SmallWaters Corporation, Chicago

Bamberg S, Schmidt P (1998) Changing travel-mode choice as rational choice: results from a longitudinal intervention study. Ration Soc 10(2):223–252

Bearden WO, Netmeyer RG (1999) Handbook of marketing scales. Sage, Thousand Oaks

Blalock HM (1964) Causal inferences in nonexperimental research. University of North Carolina Press, Chapel Hill

Boari G (2000) Uno sguardo ai modelli per la costruzione di indicatori nazionali di customer satisfaction. In: Valutazione della qualità e customer satisfaction: il ruolo della statistica. Vita e Pensiero, Milano, pp 317–336

Bollen K (1989) Structural equations with latent variables. Wiley, New York

Bollen K, Lennox R (1991) Conventional wisdom on measurement: a structural equation perspective. Psychol Bull 100:305–314

Borsboom D, Mellenbergh GJ, Heerden JV (2003) The theoretical status of latent variables. Psychol Rev 110(2):203–219

Borsboom D, Mellenbergh GJ, Heerden JV (2004) The concept of validity. Psychol Rev 111(4):1061–1071

Browne MW, Cudeck R (1993) Alternative ways of assessing model fit. In: Bollen KA, Long JS (eds) Testing structural equation models. Sage, Newbury Park, pp 136–162

Bruner GCB II, James KE, Hensel PJ (2001) Marketing scales handbook. American Marketing Association, Chicago

Byrne BM (1994) Structural equation modeling with EQS and EQS/WINDOWS: basic concepts, applications and programming, 3rd edn. Sage, Thousand Oaks

Carreira R, Patrício L Natal, Jorge R, Magee C (2014) Understanding the travel experience and its impact on attitudes, emotions and loyalty towards the transportation provider—a quantitative study with mid-distance bus trips. Transp Policy 31:35–46

Coltman T, Devinney TM, Midgley DF, Venaik S (2008) Formative versus reflective measurement models: two applications of formative measurement. J Bus Res 61(12):1250–1262

Comité Européen de Normalisation (CEN) (2002) European Standard EN 13816. Transportation - logistics and services - public passenger transport - service quality definition, targeting and measurement

de Abreu e Silva J, Goulias KG (2009) A structural equations model of land use patterns, location choice, and travel behavior in Seattle and comparison with Lisbon. In: Proceedings of the 88th annual transportation research board meeting, 11–15 Jan 2009, Washington DC

de Abreu e Silva J, Morency C, Goulias KG (2012) Using structural equations modeling to unravel the influence of land use patterns on travel behavior of workers in Montreal. Transp Res Part A 46:1252–1264

de Oña J, de Oña R, Eboli L, Mazzulla G (2013) Perceived service quality in bus transit service. A structural equation approach. Transp Policy 29:219–226

de Oña J, de Oña R, Eboli L, Mazzulla G (2016a) Index numbers for monitoring transit service quality. Transp Res Part A 84:18–30

de Oña J, de Oña R, Eboli L, Forciniti C, Mazzulla G (2016b) Transit passengers' behavioural intentions: the influence of service quality and customer satisfaction. Transportmetrica A 12(5):385–412

Diamantopoulos A, Siguaw JA (2006) Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. Br J Manag 17:263–282

Diamantopoulos A, Winklhofer HM (2001) Index construction with formative indicators: an alternative to scale development. J Mark Res 38(5):269–277

Eboli L, Mazzulla G (2007) Service quality attributes affecting customer satisfaction for bus transit. J Public Transp 10:21–34

Eboli L, Mazzulla G (2012) Structural equation modelling for analysing passengers' perceptions about railway services. Proc-Soc Behav Sci 54:96–106

Eboli L, Mazzulla G (2015) Relationships between rail passengers' satisfaction and service quality: a framework for identifying the key service factors. Public Transp Plan Oper 7(2):185–201

Eboli L, Forciniti C, Mazzulla G (2012) Exploring land use and transport interaction through structural equation modelling. Proc-Soc Behav Sci 54:107–116

Edwards JR, Bagozzi RP (2000) On the nature and direction of the relationship between constructs and measures. Psychol Methods 5:155–174

Eskildsen JK, Dahlgaard JJ (2000) A causal model for employee satisfaction. Total Qual Manag 11(8):1081–1094

Fillone AM, Montalbo CM, Tiglao NC (2005) Assessing urban travel: a structural equations modeling (SEM) approach. Proc East Asia Soc Transp Stud 5:1050–1064

Fornell C (1982) A second generation of multivariate analysis. Praeger, New York

Grace JB, Pugesek BH (1997) A structural equation model of plant species richness and its application to a coastal wetland. Am Nat 149(3):436–460

Hair JF, Anderson RE, Tatham RL, Black WC (2009) Multivariate data analysis. Prentice Hall, Englewood Cliffs

Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model: Multidiscip J 6(1):1–55

Irfan SM, Mui HKD, Shahbaz S (2011) Service quality in rail transport of Pakistan: a passenger perspective. In: Proceedings of 3rd SAICON: international conference on management, business ethics and economics (ICMBEE), 28–29 Dec 2011, Lahore, Pakistan

Joreskog KG (1973) Analysis of covariance structures. In: Krishnaiah PR (ed) Multivariate analysis-III. Academic Press, New York, pp 263–285

Karlaftis MG, Golias J, Papadimitriou E (2001) Transit quality as an integrated traffic management strategy: measuring perceived service. J Public Transp 4(1):27–44

Kline RB (1998) Principles and practice of structural equation modeling. Guilford Press, New York

MacCallum RC, Austin JT (2000) Applications of structural equation modelling in psychological research. Annu Rev Psychol 51:201–226

MacCallum RC, Browne MW, Sugawara HM (1996) Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1(2):130–149

MacLean S, Gray K (1998) Structural equation modelling in market research. J Aust Mark Res Soc 6(1):17–32

Manaresi A, Marzocchi G, Tassinari G (2000) La soddisfazione del cliente dei servizi di segreteria universitaria: un modello a equazioni strutturali. In: Valutazione della qualità e customer satisfaction: il ruolo della statistica. Vita e Pensiero, Milano, pp 291–316

Mitchell RJ (1992) Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. Funct Ecol 6:123–129

Muthén B, Kaplan D, Hollis M (2006) On structural equation modelling with data that are not missing completely at random. Psychometrika 52(3):431–462

Netmeyer RG, Bearden WO, Sharma S (2003) Scaling procedures: issues and applications. Sage, Thousand Oaks

Ngatia GJ, Okamura T, Nakamura F (2010) The structure of users' satisfaction on urban public transport service in developing country: the case of Nairobi. J East Asia Soc Transp Stud 8:1288–1300

Rossiter JR (2002) The C-OAR-SE procedure for scale development in marketing. Int J Res Mark 19(4):1–31

Schumacker RE, Lomax RG (2004) A beginner's guide to structural equation modeling, 2nd edn. Lawrence Erlbaum Associates, Mahwah

Simonetto A (2012) Formative and reflective models: state of the art. Electron J Appl Stat Anal 5(3):452–457

Spector PE (1992) Summated rating scale construction. Sage, Newbury Park

Steiger JH (1990) Structural model evaluation and modification: an interval estimation approach. Multivar Behav Res 25:173–180

Stuart KR, Mednick M, Bockman J (2000) Structural equation model of customer satisfaction for the New York City subway system. Transp Res Rec 1735:133–137

Tam Mei Ling, Tam Mei Lang, Lam WHK (2005) Analysis of airport access mode choice: a case study in Hong Kong. J East Asia Soc Transp Stud 6:708–723

Transportation Research Board (2003) Transit capacity and quality of service manual. TCRP report 100. National Academy Press, Washington, DC

Tschopp M, Axhausen KW (2007) Transport infrastructure and spatial development in Switzerland between 1950 and 2000. In: Proceedings of 86th annual meeting of the transportation research board, Jan 2007, Washington, DC

Ullman JB (2001) Structural equation modeling. In: Tabachnick BG, Fidell LS (eds) Using multivariate statistics, 4th edn. Allyn and Bacon, Needham Heights, pp 653–771

Van Acker V, Witlox F, Van Wee B (2007) The effects of the land use system on travel behavior: a structural equation modeling approach. Transp Plan Technol 30:331–353

Wiley DE (1973) The identification problem for structural equation models with unmeasured variables. In: Goldberger AS, Ducan OD (eds) Structural equation models in the social science. Seminar Press, New York, pp 69–83