

UNIVERSIDAD DE GRANADA

E.T.S. Ingenieros de Caminos, Canales y Puertos

Departamento de Ingeniería Civil



PROGRAMA DE DOCTORADO EN INGENIERÍA CIVIL (B23.56.1)

TESIS DOCTORAL

DESARROLLO Y VALIDACIÓN DE TÉCNICAS AVANZADAS PARA LA EVALUACIÓN DE LA CALIDAD DEL SERVICIO PERCIBIDA POR EL USUARIO EN TRANSPORTES PÚBLICOS

DEVELOPMENT AND VALIDATION OF ADVANCED TECHNIQUES TO EVALUATE THE QUALITY OF PUBLIC TRANSPORT SERVICE FROM PASSENGER'S PERSPECTIVE

Para la obtención del
GRADO DE DOCTOR POR LA UNIVERSIDAD DE GRANADA CON
MENCIÓN DE DOCTORADO INTERNACIONAL

AUTOR:

FRANCISCO JAVIER DÍEZ DE LOS RÍOS MESA

Granada, 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: Francisco Javier Díez De Los Ríos
ISBN: 978-84-1306-454-3
URI: <http://hdl.handle.net/10481/59911>

UNIVERSIDAD DE GRANADA
E.T.S. Ingenieros de Caminos, Canales y Puertos
Departamento de Ingeniería Civil



TESIS DOCTORAL

DESARROLLO Y VALIDACIÓN DE TÉCNICAS AVANZADAS PARA LA EVALUACIÓN DE LA CALIDAD DEL SERVICIO PERCIBIDA POR EL USUARIO EN TRANSPORTES PÚBLICOS

DEVELOPMENT AND VALIDATION OF ADVANCED TECHNIQUES TO EVALUATE THE QUALITY OF PUBLIC TRANSPORT SERVICE FROM PASSENGER'S PERSPECTIVE

AUTOR:

FRANCISCO JAVIER DÍEZ DE LOS RÍOS MESA

DIRECTORES:

JUAN DE OÑA LÓPEZ

Doctor Ingeniero de Caminos, Canales y Puertos

Universidad de Granada

ROCÍO DE OÑA LÓPEZ

Doctor Ingeniero de Caminos, Canales y Puertos

Universidad de Granada

“El placer más noble es el júbilo de comprender”

Leonardo da Vinci

AGRADECIMIENTOS

El hecho de que hoy pueda escribir estas palabras en esta página se debe principalmente a cinco personas. Por tanto, me gustaría comenzar agradeciéndoles a ellos su gran papel en la consecución de esta meta.

A mi padre Francisco Díez de los Ríos Torralvo. Él es mi modelo de persona a seguir, mi héroe sin capa, mi maestro de vida. Su forma de ser, sus palabras en momentos críticos y su lucha diaria por conseguir lo mejor para sus hijos, me ha llevado a culminar con éxito esta tesis, así como otras metas que nunca pensé que sería capaz de alcanzar.

A mi madre M^ª Teresa Mesa Algarra. Ella es mi acompañante en el camino, mi apoyo continuo, mi fe inquebrantable. Durante toda mi vida, y en especial durante la realización de esta tesis, ha sido quien me ha escuchado día a día, entendiendo cada uno de mis problemas, sufriendo y alegrándose conmigo en cada paso conseguido. Sin ella, el camino nunca hubiera llegado a su fin.

A mi hermana M^ª Teresa Díez de los Ríos Mesa. Ella, aparte de ser la mejor hermana que uno puede desear, es mi mejor amiga. He tenido la gran suerte de que haya sido mi compañera de piso durante los años de realización de la tesis, lo que me ha llevado a crear un vínculo de amor, confianza y amistad con ella que pocos hermanos pueden tener. Todos estos años de tesis han sido mucho menos duros por el simple hecho de estar a su lado.

A mi prometida Ana Isabel Cartas Reyes. Ella es la gran suerte de mi vida. Su aparición en el comienzo de esta aventura, ha sido crucial para mí en todos los sentidos. Es quien me proporciona cada día fuerzas para seguir adelante, es mi equilibrio, mi comprensión, mi apoyo, llegando a creer en mí incluso más de lo que ha hecho nadie. Es la alegría de mis días, el motor de mi vida.

Y a mi primo Fran “Grande”. Él sabe el porqué.

También me gustaría agradecer a todas las personas que he tenido el gusto de conocer durante esta aventura, con mención especial para Pedro, Ángel y Mercedes, mis tres grandes amigos con quienes he compartido grandes y felices momentos.

Igualmente, quisiera agradecer a mis directores y tutores de tesis Juan y Rocío de Oña. Quisiera agradecerles su tiempo, paciencia y cariño con el que me han enseñado, acompañado y ayudado en esta aventura. El hecho de que me dierais la oportunidad de realizar esta tesis y trabajar con vosotros es algo que nunca podre agradecer.

A las profesoras Silvia Salini y Bridgette Wessels, por acogerme y ayudarme tanto durante mis estancias en Italia e Inglaterra, respectivamente.

Al Ministerio de Economía y Competitividad del Gobierno de España, responsable de la beca de formación de personal investigador asociada al proyecto de investigación TRA2012-37823, de la que fui adjudicatario en el año 2013 y que me ha permitido realizar esta tesis doctoral, y de su apoyo a través del proyecto de investigación TRA2015-66235-R. Así mismo, agradecer por su apoyo financiero al FEDER de la Unión Europea via el proyecto “Mejora de la calidad del transporte público para fomentar la movilidad sostenible: Metro de Sevilla” del “Programa Operativo FEDER de Andalucía 2007-2013”.

A mi familia, que sin su gran apoyo al comienzo de esta aventura, y sin su constante fe en mi, no hubiera sido posible que pudiera realizar este doctorado.

Y por último, este párrafo de agradecimiento es para todas aquellas personas que no he nombrado, pero saben perfectamente que sin su granito de arena esto no hubiera sido posible.

Muchas gracias a todos de corazón.

RESUMEN

En los últimos años, el estudio de la calidad del servicio se ha convertido en uno de los grandes focos de atención por parte de los planificadores del transporte. Su principal objetivo con esta tarea es conocer hacia dónde deben dirigir sus esfuerzos e inversiones para mejorarla, puesto que se ha demostrado que un transporte público que opere con alta calidad puede llegar a competir directamente con el vehículo privado, atrayendo nuevos pasajeros y mejorando la satisfacción de los actuales. Lo que a su vez, repercute al mismo tiempo en una mejora inmediata de la movilidad dentro de las ciudades y su área metropolitana, así como de la economía de la zona y del medio ambiente en general.

Tal estudio se suele realizar desde el punto de vista del pasajero utilizando encuestas de satisfacción, pues se argumenta que es quién se beneficia o perjudica en primera instancia de dicha calidad. Dicho lo cual, se puede encontrar en la literatura un amplio rango de técnicas estadísticas y metodologías que son aplicadas en el ámbito científico, y que trabajan con esta tipología de datos (p.e., Modelo Probit o Logit, Redes Neuronales Artificiales, Árboles de decisión, etc.). Sin embargo, en el ámbito profesional de los proveedores del transporte público, se debe destacar que la mayoría de las técnicas de análisis utilizadas son más sencillas (p.e., análisis de frecuencias, medias, medianas o análisis bivariado).

Una de las técnicas estadísticas más famosas durante los últimos años, y que mejor consideración tiene entre la comunidad científica en el ámbito del transporte público, son los Modelos de Ecuaciones Estructurales (Structural Equation Modelling, SEM). Su capacidad para medir y determinar relaciones estructurales entre un gran número de variables, tanto de tipología latente como observada, hacen de ésta una técnica idónea para desarrollar y validar modelos.

No obstante, esta técnica presenta un gran hándicap, como es la necesidad de conocimiento experto previo para el desarrollo de dichos modelos. Una posible solución a dicha limitación sería su combinación con otra técnica estadística que permita realizar dicha extracción de conocimiento de forma automática y directa a partir de los datos disponibles, y sin necesidad de conocimiento previo.

En esta tesis doctoral se proponen para desempeñar dicho rol a las Redes Bayesianas (RB). Se trata de una técnica estadística que surge del campo de la Minería de Datos y que debido a su fácil aplicación a diferentes tipologías de datos y variables, a sus diferentes y bien estudiados algoritmos de aprendizaje (ya implementados en multitud de software), a que permite conseguir hipotéticos modelos directamente de los datos, y a su intuitiva interpretación de sus resultados, hacen de ésta la técnica estadística que mejor puede complementar a los SEM.

Dicho lo cual, se propone una metodología de dos etapas basada en la combinación de ambas técnicas, cuya finalidad sea el aprendizaje automático, desarrollo y validación de modelos en cualquier ámbito de conocimiento. Por estas razones, el objetivo principal de esta tesis doctoral es el desarrollo, exposición y validación de dicha metodología.

Con el fin de alcanzar dicho objetivo, se han realizado dos casos de estudio diferentes, en cada uno de los cuáles, se ha puesto énfasis en diferentes ventajas y aspectos de dicha metodología de dos etapas.

En el primer caso, se ha desarrollado un modelo sobre la calidad del servicio del Metro Ligero de Sevilla (España), donde se ha puesto el foco de atención en el proceso a seguir para su aplicación, así como en los parámetros a considerar para su evaluación. Los buenos resultados obtenidos y su coherencia con la realidad han permitido, por un lado, constatar la validez de dicha propuesta metodológica, y por otro lado, obtener un modelo que explica cómo los pasajeros perciben la calidad del servicio de Metro Ligero de Sevilla, muy conveniente desde el punto de vista de los operadores de este servicio.

Y en el segundo caso, con el objetivo de comprobar la utilidad de dicha metodología, no solo en el ámbito de la calidad del transporte, si no en cualquier otro ámbito, se ha desarrollado un modelo sobre actitudes hacia el servicio del Metro Ligero de Sevilla, cuyos resultados satisfactorios y afines a los conseguidos en otros trabajos en el campo de las actitudes que hay en la literatura, han permitido testear dicho objetivo.

ABSTRACT

In recent years, the study of service quality has become one of the main focus of attention by transport planners and operators. Their main objective with this task is to find out where they should direct their efforts and investments to improve it. This fact is due to a public transport which operates with high quality can directly compete with the private vehicle, attracting new passengers and improving satisfaction of the current ones. Besides, it has other consequences such as an immediate improvement of mobility within cities and metropolitan areas, as well as of the economy and the environment, in general.

Such a study is usually carried out from the passenger's point of view using satisfaction surveys, since it is argued that the passenger is who benefits or harms in the first instance, about this quality. A wide range of statistical techniques and methodologies can be found in the literature, which work with this kind of data in the scientific field (e.g., Probit or Logit Model, Artificial Neural Networks, Decision Trees, etc.). However, it should be noted that, in the professional field of public transport providers, most of the analysis techniques used are simpler than the mentioned previously (e.g., frequency, average, medium or bivariate analysis).

One of the most famous statistical techniques in recent years, which has a good consideration among the scientific community in the field of public transport, are the Structural Equation Modelling (SEM). SEM has the ability to measure and determine structural relationships between a large numbers of variables, both latent and observed, making it an ideal technique to develop and validate models.

However, this technique presents the great handicap of requiring prior expert knowledge in order to develop the models. A possible solution to this limitation would be its combination with another statistical technique, which allows to extract knowledge automatically and directly from the available data, and without the requirement of prior expert knowledge.

In this doctoral thesis, Bayesian Networks is proposed to play this role. It is a statistical technique that arises from the field of Data Mining and it has the capacity of extract hypothetical models directly from the data. Besides, Bayesian Networks show multiple characteristics which point it out as the best complement for the SEM. Such as, it can be easily applied to different kind of data and variables, there are a wide range of well-studied learning algorithms (already implemented in multitude of softwares), and its results are intuitive and easily interpretable.

In consequence, a two-stage methodology is proposed. It is based on the combination of both techniques (i.e. BN and SEM) and its purpose is the automatic learning, development and validation of models in any field of knowledge. For these reasons, the main objective of this doctoral thesis is the development, presentation and validation of this methodology.

In order to achieve this objective, two different case studies have been carried out. In each of which, emphasis has been placed on different advantages and aspects of this two-stage methodology.

In the first case, a service quality model of the Light Rail Transit of Seville (Spain) has been developed. The focus has been placed on the process followed for applying the two-stage methodology, as well as on the parameters which has been considered for its evaluation. The good results obtained and its coherence with reality have allowed, on the one hand, to verify the validity of this proposed methodology, and, on the other hand, to obtain a model which explains how passengers perceive the quality of the Light Rail Transit of Seville. This last point is very convenient for operators of this service.

Finally, in the second case, a model of attitudes towards the service of the Light Rail Transit of Seville (Spain) was employed to verify the usefulness of the two-stage methodology in other areas, not only in the field of service quality in transport. The satisfactory results of the model and their similarity to those achieved in other works in the literature related with this field of attitudes have allowed us to test this objective.

ÍNDICE GENERAL

AGRADECIMIENTOS.....	I
RESUMEN.....	III
ABSTRACT	V
ÍNDICE GENERAL.....	VII
ÍNDICE DE TABLAS	XI
ÍNDICE DE FIGURAS	XIII
TABLA DE ACRÓNIMOS.....	XV
1. INTRODUCCIÓN	1
1.1. PLANTEAMIENTO DEL PROBLEMA.....	1
1.2. ORGANIZACIÓN DE LA TESIS	6
1.3. PRINCIPALES APORTACIONES	7
2. ESTADO DEL ARTE.....	11
2.1. LA CALIDAD DEL SERVICIO EN EL TRANSPORTE PÚBLICO.....	12
2.1.1. DEFINICIÓN DEL CONCEPTO DE CALIDAD DEL SERVICIO	15
2.1.2. SINGULARIDADES DE LA CALIDAD DEL SERVICIO.....	17
2.1.2.1. SELECCIÓN DE ATRIBUTOS Y/O DIMENSIONES	17
2.1.2.2. CONOCIMIENTO TEÓRICO O EXPERTO PREVIO.....	20
2.1.2.3. EXPECTATIVAS	20
2.1.2.4. HETEROGENEIDAD.....	22
2.1.2.5. PARADIGMA CALIDAD DEL SERVICIO-SATISFACCIÓN-INTENCIÓN DE RECOMPRA Y OTRAS ACTITUDES HACIA EL TRANSPORTE PÚBLICO	23
2.1.3. METODOLOGÍAS PARA LA OBTENCIÓN DE INFORMACIÓN SOBRE LA CALIDAD DEL SERVICIO.....	25
2.1.3.1. ASPECTOS GENERALES DE LAS ENCUESTAS	28
2.2. MÉTODOS DE ANÁLISIS DE LA CALIDAD DEL SERVICIO	32
2.2.1. MÉTODOS ESTADÍSTICOS TRADICIONALES	32
2.2.1.1. TÉCNICAS SIMPLES	33
2.2.1.2. TÉCNICAS AVANZADAS	36
2.2.2. MÉTODOS DE ANÁLISIS CON ORIGEN EN LA MINERÍA DE DATOS.....	40

2.2.2.1. TÉCNICAS SUPERVISADAS.....	42
2.2.2.2. TÉCNICAS NO SUPERVISADAS.....	51
2.3. DISCUSIÓN Y CONCLUSIONES	56
3. OBJETIVOS.....	69
3.1. OBJETIVO PRINCIPAL.....	70
3.2. OBJETIVOS ESPECÍFICOS	70
3.3. HIPÓTESIS.....	71
4. METODOLOGÍA Y MATERIALES	75
4.1. METODOLOGÍA	75
4.1.1. ANÁLISIS FACTORIAL	76
4.1.1.1. ANÁLISIS FACTORIAL EXPLORATORIO (AFE).....	76
4.1.1.2. ANÁLISIS FACTORIAL CONFIRMATORIO (AFC)	78
4.1.2. REDES BAYESIANAS (RB).....	81
4.1.3.1. ALGORITMOS DE APRENDIZAJE DE LAS REDES BAYESIANAS	84
4.1.3.2. MÉTODO DE OBTENCIÓN DE LA RED BAYESIANA MÁS ROBUSTA.....	87
4.1.4. MODELOS DE ECUACIONES ESTRUCTURALES (SEM)	89
4.2. FASES DE LA INVESTIGACIÓN	98
4.2.1. CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA). 99	
4.2.2. CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO.	103
4.3. DATOS	103
4.3.1. EL METRO LIGERO DE SEVILLA	103
4.3.2. ENCUESTA DEL PROYECTO “MEJORA DE LA CALIDAD DEL TRANSPORTE PÚBLICO PARA FOMENTAR LA MOVILIDAD SOSTENIBLE: METRO DE SEVILLA (G-GI3002/IDIK)”	104
5. RESULTADOS Y DISCUSIÓN	117
5.1. CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA).....	118
5.1.1. PRE-PROCESAMIENTO DE LOS DATOS.	118
5.1.1.1. RE-CATEGORIZACIÓN DE LOS ATRIBUTOS.....	119
5.1.1.2. ANÁLISIS FACTORIAL EXPLORATORIO (AFE).....	119

5.1.1.3. ANÁLISIS FACTORIAL CONFIRMATORIO (AFC)	122
5.1.2. 1º ETAPA. EXTRACCIÓN DE CONOCIMIENTO DE LOS DATOS	128
5.1.3. 2ª ETAPA: VALIDACIÓN DE LAS RELACIONES	132
5.1.4. DISCUSIÓN SOBRE EL MODELO DE CALIDAD DEL SERVICIO.	138
5.2. CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO.	141
5.2.1. PRE-PROCESAMIENTO DE LOS DATOS.	142
5.2.1.1. RE-CATEGORIZACIÓN DE LOS ATRIBUTOS.....	142
5.2.1.2. ANÁLISIS FACTORIAL EXPLORATORIO	143
5.2.1.3. ANÁLISIS FACTORIAL CONFIRMATORIO	148
5.2.2. 1º ETAPA. EXTRACCIÓN DE CONOCIMIENTO DE LOS DATOS	153
5.2.3. 2ª ETAPA: VALIDACIÓN DE LAS RELACIONES	155
5.2.4. DISCUSIÓN SOBRE EL MODELO DE ACTITUDES HACIA EL SERVICIO CONSEGUIDO.	164
6. CONCLUSIONES	169
6. CONCLUSIONS	177
7. FUTURAS LÍNEAS DE INVESTIGACIÓN	183
8. REFERENCIAS BIBLIOGRÁFICAS	189
9. ANEXOS.....	219

ÍNDICE DE TABLAS

Tabla 1. Clasificación de los modelos de regresión (Adaptada de Lopez, 2014).	37
Tabla 2. Muestra de trabajos que estudian la calidad del servicio en el transporte público.	56
Tabla 3. Resumen de los Índices de Ajuste seleccionados para modelos SEM y sus umbrales recomendados.	97
Tabla 4. Esquema de las fases de investigación.	102
Tabla 5. Resumen de los resultados de la Encuesta de Actitudes: Parte A.	108
Tabla 6. Resumen de los resultados de la Encuesta de Actitudes: Parte B.	110
Tabla 7. Resumen de los resultados de la Encuesta de Actitudes: Parte C.	111
Tabla 8. Resumen de los resultados de la Encuesta de Actitudes: Parte D.	113
Tabla 9. Resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 1.	121
Tabla 10. Resultados del Análisis Factorial Exploratorio tras la depuración y mejora de las dimensiones, caso de estudio 1.	123
Tabla 11. Resultados del Análisis Factorial Confirmatorio, caso de estudio 1.	125
Tabla 12. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 1.	127
Tabla 13. Ratios de missclassification de las RB con el mismo número de arcos con valores mayores al umbral establecido para el caso de estudio 1.	129
Tabla 14. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 1.	131
Tabla 15. Pesos de carga y Errores del modelo de medición para el SEM del caso de estudio 1.	133
Tabla 16. Pesos de carga y Errores del modelo estructural para el SEM del caso de estudio 1.	134
Tabla 17. Efectos directos, indirectos y totales estandarizados sobre la dimensión “Calidad del Servicio” (CS) del caso de estudio 1.	136
Tabla 18. Medidas de bondad de ajuste del modelo SEM del caso de estudio 1.	137
Tabla 19. Primeros resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 2.	143
Tabla 20. Segundos resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 2.	147
Tabla 21. Resultados del Análisis Factorial Exploratorio tras la depuración y mejora de las dimensiones, caso de estudio 2.	149
Tabla 22. Resultados del Análisis Factorial Confirmatorio, caso de estudio 2.	150
Tabla 23. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 2.	152

Tabla 24. Ratios de missclassification de las RB con el mismo número de arcos con valores mayores al umbral establecido para el caso de estudio 2.....	154
Tabla 25. Pesos de carga y Errores del modelo de medición para el SEM del caso de estudio 2.	157
Tabla 26. Pesos de carga y Errores del modelo estructural para el SEM del caso de estudio 2.	157
Tabla 27. Efectos directos estandarizados en el modelo de Actitudes.	159
Tabla 28. Efectos indirectos estandarizados en el modelo de Actitudes.	159
Tabla 29. Efectos totales estandarizados en el modelo de Actitudes.	159
Tabla 30. Medidas de bondad de ajuste del modelo SEM del caso de estudio 2.....	161

ÍNDICE DE FIGURAS

Figura 1. Resumen de los métodos de análisis con origen fuera de la Minería de Datos.	33
Figura 2. Matriz IPA (Adaptado de Oña, 2014).	35
Figura 3. Estructural general de un SEM.	40
Figura 4. Resumen de los métodos de análisis con origen en la Minería de Datos aplicados a estudios de calidad del servicio del transporte público (Adaptada de Rokach and Maimon, 2008).....	41
Figura 5. Ejemplo de estructura de una RNA para analizar la calidad del servicio en el transporte público (Adaptada de Garrido et al., 2014).....	43
Figura 6. Estructura de un ADD (Adaptado de Kashany and Mohaymany, 2011).....	45
Figura 7. Ejemplo de una RB (DAG y Tablas de probabilidad condicional para cada nodo).....	48
Figura 8. Ejemplo AC con Algoritmo Jerárquico.....	52
Figura 9. Ejemplo AC con Algoritmo No Jerárquico.	52
Figura 10. Esquema de los 5 pasos para realizar un AFE (Adaptado Hair et al., 2010). 78	
Figura 11. Esquema de pasos para realizar un AFC (Adaptado y completado de Hair et al., 2010).....	81
Figura 12. Tipos de variables en un modelo SEM.	90
Figura 13. Pasos básicos recomendados del SEM (Kline, 2011).	91
Figura 14. Tarjetas divulgativas de la Encuesta de Actitudes del Metro de Sevilla.	106
Figura 15. RB más robusta del caso de estudio 1.....	130
Figura 16. RB más robusta del caso de estudio 1, con la proporción de ocurrencia de cada arco en el proceso de re-muestreo iterativo.	131
Figura 17. Modelo Estructural del SEM del caso estudio 1, con relaciones significativas.	134
Figura 18. RB más robusta para el caso de estudio 2.	154
Figura 19. RB más robusta para el caso de estudio 2, con la proporción de ocurrencia de cada arco en el proceso de re-muestreo iterativo.	155
Figura 20 Modelo estructural del SEM del caso de estudio 2, con relaciones significativas.	158
Figura 21 Modelo estructural del SEM de Machado et al. (2018).....	162

TABLA DE ACRÓNIMOS

<i>Análisis Cluster</i>	AC
<i>Análisis de Importancia-Rendimiento (Importance-Performance Analysis)</i>	IPA
<i>Análisis Factorial Exploratorio</i>	AFE
<i>Análisis Factorial Confirmatorio</i>	AFC
<i>Árboles de Decisión</i>	ADD
<i>Criterio de Información de Akaike (Akaike Information Criterion)</i>	AIC
<i>Criterio de Información Bayesiano (Bayesian Information Criterion)</i>	BIC
<i>Distribución de Probabilidad Conjunta (Joint Probability Distribution)</i>	JPD
<i>Fiabilidad del Constructo (Construct Reliability)</i>	CR
<i>Grafo Dirigido Acíclico (Directed Acyclic Graph)</i>	DAG
<i>Índice de Bondad de Ajuste (Goodness of Fit Index)</i>	GFI
<i>Índice de Bondad de Ajuste Comparativo (Comparative Fit Index)</i>	CFI
<i>Índice de Bondad de Ajuste Corregido (Adjusted Goodness of Fit Index)</i>	AGFI
<i>Índice de Bondad de Ajuste Parsimonioso (Parsimonious Goodness of Fit Index)</i>	PGFI
<i>Índice de Ajuste Normalizado (Normed Fit Index)</i>	NFI
<i>Índice de Ajuste Normalizado Parsimonioso (Parsimonious Normed Fit Index)</i>	PNFI
<i>Influencia de la Distancia-Ponderada (Distance-Weighted Influence)</i>	DWI
<i>Media de la Varianza Extraída (Average Variance Extracted)</i>	AVE
<i>Medida de Adecuación de la Muestra (Measure of Sampling Adequacy)</i>	MSA
<i>Método Máxima Verosimilitud (Maximun Likelihood)</i>	ML
<i>Modelo Chi-Cuadrado</i>	χ^2
<i>Modelos de Ecuaciones Estructurales (Structural Equation Modelling)</i>	SEM
<i>SERVice QUALity</i>	SERVQUAL
<i>SERVice PERFormance</i>	SERVPERF
<i>Raíz Cuadrada del Promedio de los Residuos (Root Mean Square Residual)</i>	RMR
<i>Raíz Cuadrada del Promedio de los Residuos de Aproximación (Root Mean Square Error of Aproximation)</i>	RMSEA
<i>Redes Bayesianas</i>	RB
<i>Redes Neuronales Artificiales</i>	RNA



CAPÍTULO 1.

INTRODUCCIÓN

Capítulo 1.

INTRODUCCIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA

En la sociedad actual la calidad es considerada un aspecto clave en cualquier tipo de servicio o negocio siendo, por tanto, objeto de continuo estudio y mejora. Empresarios, productores, o incluso gobiernos, se encuentran focalizados en proveer servicios de alto nivel de calidad ya que se ha demostrado que la misma produce, no solo un aumento del nivel de satisfacción del cliente en general, sino también un aumento del uso del servicio o sistema, atrayendo nuevos clientes y mejorando la lealtad y la imagen pública (Constantine et al., 2006; Ennew and Binks, 1996; Transportation Research Board, 1999).

Si nos centramos en el campo del transporte público, se puede observar que esta situación se repite de forma similar, ya que los operadores, gobiernos y planificadores del transporte, están centrando gran parte de sus esfuerzos en estudiar y analizar la calidad del servicio. Su objetivo, aparte de los relacionados con la satisfacción y lealtad de los pasajeros que ya han sido previamente mencionados, es lograr un medio de transporte público sostenible y de alta calidad, que sea capaz de competir en igualdad de condiciones con el vehículo privado. Hecho realmente primordial hoy en día, puesto que el excesivo uso de este último medio en ciudades y área metropolitanas es el principal causante de la mayoría de problemas que se derivan de la movilidad (p.e., congestión, contaminación, ruidos, accidentes, etc.).

El concepto de calidad del servicio en el transporte público puede ser definido como el resultado de una serie de relaciones entre las diferentes dimensiones que lo conforman (p.e., accesibilidad, confort, seguridad, etc.). Dichas dimensiones, que son elementos latentes, es decir, que no son observables o medibles de forma directa, se componen de una agrupación de diferentes atributos del servicio relativamente similares. Estos son prestados por los operadores, y al mismo tiempo, son percibidos y usados por los pasajeros (p.e., iluminación, frecuencia, funcionamiento de ascensores, etc.). Como consecuencia de este hecho, se puede afirmar que la calidad del servicio puede ser medida desde estos dos puntos de vista que son totalmente diferentes. Por un lado, los operadores realizan una evaluación de la calidad basándose en los costes de eficiencia y efectividad del servicio de transporte prestado. Por otro lado, los pasajeros valoran la calidad basándose en sus propias percepciones del servicio utilizado y en el cumplimiento de sus expectativas sobre el mismo.

Como se observa, ambos puntos de vista son importantes y aportan información relevante. Sin embargo, debido a que los pasajeros son quienes van a disfrutar y/o sufrir los beneficios y/o problemas del servicio, muchos autores consideran que la calidad del servicio debe ser medida desde el punto de vista de los pasajeros (p.e., Berry et al., 1990; Das and Pandit, 2014; Joewono and Kubota, 2007, Parkan, 2002; Rietveld, 2005; etc.).

La principal metodología que se utiliza para la recolección de las percepciones y/o expectativas de los pasajeros son las llamadas encuestas de satisfacción. Se observa que, dado el creciente interés en analizar y estudiar la calidad del servicio por parte de los agentes del transporte, se están llevando a cabo, cada vez más, diversas campañas de encuestas de satisfacción en diferentes servicios de transporte público y con una cierta periodicidad (p.e., cada tres meses, seis meses, un año, etc.). Este hecho está permitiendo la obtención de una gran cantidad y variedad de información útil, que se está usando en una amplia gama de tipologías de estudios (p.e., sobre calidad, de marketing, de evolución de la demanda, etc.).

De acuerdo al tipo de encuesta de satisfacción empleada, la información lograda será más o menos completa, pudiendo contener desde simples opiniones o percepciones de los pasajeros sobre los diferentes atributos del servicio, hasta un amplio rango de información sobre características socioeconómicas del pasajero, hábitos de viaje, actitudes hacia el transporte público, ranking de importancias declaradas, etc. Por tanto, de acuerdo a la información disponible, se puede afirmar que quedarán habilitadas para su aplicación diferentes técnicas estadísticas, que al mismo tiempo, estarán sujetas a la tipología de estudio que se quiera realizar o a los resultados que se deseen lograr.

Centrando la atención en el análisis de la calidad del servicio, se puede considerar la existencia en la literatura de un amplio abanico de técnicas estadísticas, como por ejemplo, el Análisis de Importancia-Rendimiento (Importance-Performance Analysis, IPA), los modelos Logit o Probit, el Análisis Factorial, la Regresión Múltiple, etc. Sin embargo, los SEM se erigen como la técnica estadística más popular y cuya aplicación, especialmente en el ámbito del transporte público, ha sufrido un gran crecimiento en los últimos años. La principal causa de este hecho es su capacidad para tratar fácilmente con un gran número de variables, ya sean de tipo exógeno o endógeno. Así mismo, otra causa de su auge es que SEM es considerado uno de los mejores métodos estratégicos integrados para la medición y cálculo de relaciones estructurales de variables de tipo latente (no observables), lo cual, es realizado mediante una combinación lineal de las variables observadas del modelo.

Por tanto, dadas las características intrínsecas de la calidad del servicio en el transporte público que la convierten en un concepto complejo, abstracto y difuso, y que depende o se explica mediante una serie de variables observables y latentes, lo apropiado de la aplicación de SEM en su análisis y estudio está más que justificado.

No obstante, se debe destacar que el fin último de dicha técnica en cualquier campo es la confirmación de relaciones entre variables y validación de modelos hipotetizados, siendo necesario para su aplicación un conocimiento científico o experto previo que indique cuál es la estructura de dicho modelo y las relaciones entre las diferentes variables.

Debido a esta necesidad, ciertas relaciones importantes entre las variables podrían no haber sido consideradas. Hecho que habría repercutido de forma directa sobre los resultados y modelos obtenidos, al no ser éstos totalmente acordes a la realidad. Se debe destacar que esta circunstancia ocurre de forma totalmente involuntaria por parte del investigador, ya que su origen puede ser el desconocimiento de dichas relaciones, o que se traten de características propias y singulares del fenómeno estudiado.

Esta situación se ve especialmente agravada en los casos de análisis y estudio de la calidad del servicio en el transporte público, donde la comunidad científica simplemente confirma que existe una relación e influencia directa sobre la percepción global de la calidad del servicio por parte de todas las dimensiones que la caracterizan. Es decir, que no hay constancia de ningún tipo de acuerdo sobre las posibles relaciones entre dichas dimensiones a pesar de las diferentes evidencias recogidas en la literatura, que confirman la existencia de las mismas y su repercusión en los resultados finales.

Resumiendo, se hace patente la necesidad de recurrir a otra técnica estadística que sea complementaria al SEM, y que solvente este problema de necesidad de conocimiento científico o experto previo sobre el fenómeno que se está modelizando, extrayéndolo directamente de los datos iniciales.

Con tal fin, si se realiza una revisión más profunda de las diferentes técnicas con origen en la Minería de Datos, se pueden encontrar las RB (Redes Bayesianas) que permiten solventar esta limitación, ya que se trata de una técnica de tipo exploratorio que aprende relaciones entre variables y desarrolla hipotéticos modelos, directamente de los datos suministrados, y sin necesidad de ningún tipo de información previa o hipótesis, presentándolos en una forma muy fácil e intuitiva de interpretar.

Sin embargo, dado que las RB por sí solas no son capaces de establecer un modelo preciso sobre la realidad, puesto que, puede darse el caso de que ciertas relaciones sean establecidas debido a propiedades puramente estadísticas, alejadas de lo que realmente ocurre, se requiere la aplicación de una técnica que las valide.

En consecuencia, se puede afirmar que estas características la convierten en una técnica complementaria perfecta del SEM.

Considerando estos hechos, en esta tesis doctoral, se propone una metodología de dos etapas basada en el uso combinado de las RB y del SEM. Se trata de un novedoso proceso metodológico en el campo de los transportes públicos que permite, en una primera etapa exploratoria, obtener ciertas hipótesis sobre las relaciones de las variables que componen el fenómeno analizado mediante el uso de las RB, y en una segunda etapa confirmatoria, validar esas relaciones y el modelo global mediante el uso de SEM.

No obstante, se debe destacar que esta metodología de dos etapas, no es simplemente aplicable en el específico ámbito de la calidad, sino que puede ser extrapolable a otros campos de conocimiento, tal y como ha sido ya probada en el campo de la sanidad (p.e., Duarte et al., 2011; Scheines et al., 1999; Trentini et al., 2015) y en el de los juegos online (Jo and Park, 2017).

Con tal fin, se considera otro aspecto del transporte público que es muy interesante para los operadores y planificadores del transporte, en cuanto a que su conocimiento y comprensión permite posicionarlos en una situación de ventaja competitiva con respecto al resto de transportes, así como ayudar en el diseño efectivo de estrategias que cumplan las necesidades de los pasajeros. Se trata del estudio y análisis de las actitudes de los pasajeros hacia el servicio de transporte público.

Dicho estudio persigue el establecer un modelo que describa cuáles son los principales constructos que gobiernan dichas actitudes, y cómo se relacionan entre ellas. Al igual que ocurría en el estudio y análisis de la calidad del servicio, dichas actitudes que conforman este ámbito, se encuentran agrupadas en diferentes constructos latentes, que a priori, presentan una serie de relaciones. Es cierto que algunas de ellas son aceptadas por la comunidad científica, como es el caso de la relación entre los constructos de Calidad del servicio-Satisfacción-Lealtad. Sin embargo, otras son totalmente desconocidas e interpretadas de diferente forma, tal y como se puede observar en la literatura.

Además, se debe destacar que la mayoría de los modelos que se han establecido mediante SEM sobre este ámbito, están desarrollados a partir de conocimiento científico o experto previo.

En consecuencia, se puede observar que se dan similares circunstancias que las descritas en el análisis de la calidad del servicio, lo que lo convierten en un ámbito idóneo para mostrar la versatilidad y aplicabilidad de esta metodología de dos etapas en cualquier ámbito que demande la extracción automática de conocimiento directamente de los datos, y su posterior validación.

Esta tesis doctoral está orientada a desarrollar, implementar y validar esta metodología de dos etapas mediante dos casos de estudio diferentes, centrados en el servicio de Metro Ligero de Sevilla (España).

El primer caso de estudio se encuentra orientado a estudiar la calidad de dicho servicio usando esta metodología de dos etapas, enfatizando los diferentes pasos a seguir para su correcto funcionamiento, y analizándose en profundidad los resultados obtenidos.

Y el segundo caso de estudio se centra en las actitudes hacia el servicio de Metro Ligero de Sevilla (España), destacando la capacidad de esta metodología de dos etapas de poder aplicarse a otros ámbitos diferentes de la calidad.

1.2. ORGANIZACIÓN DE LA TESIS

La tesis está organizada en los siguientes siete capítulos.

Capítulo 1: Incluye una introducción a la tesis, la estructura del documento y las principales contribuciones de esta investigación.

Capítulo 2: Se constituye por el estado del arte, que presenta una descripción de las características generales y singularidades de la calidad del servicio en el sector del transporte público, las diferentes tipologías de encuestas y sus características, y las principales técnicas estadísticas para el análisis de dicha calidad, ya tengan o no, su origen en la Minería de Datos. Se presta mayor grado de atención a las técnicas de RB y SEM.

Capítulo 3: Muestra los diferentes objetivos e hipótesis de este trabajo de investigación.

Capítulo 4: Presenta las fases de investigación, la metodología y la base de datos con la que se trabaja en esta tesis. Dicho capítulo se encuentra estructurado en tres partes bien diferenciadas. En la primera parte, se establecen las fases de investigación de cada uno de los dos casos de estudio que se han realizado. En la segunda parte, es definida la metodología de dos etapas propuesta en esta tesis para aplicarla al campo del transporte público, los algoritmos de cada una de las técnicas empleadas en dicha metodología, las técnicas estadísticas empleadas para reducir la heterogeneidad y agrupar los atributos en dimensiones, los parámetros de ajuste utilizados en cada una de ellas y los criterios para evaluar los diferentes modelos. Y, por último, en la tercera parte se lleva a cabo una descripción detallada de la base de datos, su procedencia, la metodología y tipología de encuesta empleada para su obtención, y un análisis descriptivo de los datos obtenidos.

Capítulo 5: Muestra los resultados obtenidos de cada uno de los dos casos de estudio realizados. En el primer caso se aplica únicamente la metodología de dos etapas para analizar y estudiar la calidad del servicio de Metro Ligero de Sevilla (España). Y en el segundo caso, se vuelve a utilizar la metodología de dos etapas, pero esta vez, para analizar y estudiar las actitudes hacia el servicio de Metro Ligero de Sevilla (España).

Capítulo 6: Presenta las principales conclusiones de este trabajo.

Capítulo 7: Expone las futuras líneas de investigación.

Finalmente, **Capítulo 8** incluye todas las referencias usadas en esta tesis.

1.3. PRINCIPALES APORTACIONES

Las principales aportaciones a las que ha dado lugar esta tesis son dos artículos publicados en revistas internacionales de alto impacto (Anexo 4 y 5):

- **Artículo 1.** Diez-Mesa, F., De Oña, R., De Oña, J. (2018). Bayesian Networks and Structural Equations Modelling to Develop Service Quality Models: Metro of Seville Case Study. *Transportation Research Part A: Policy and Practice*, 118, 1-13.
- **Artículo 2.** De Oña, R., Diez-Mesa, F., De Oña, J. (under review) Bayesian Network and Structural Equation Model for Analysing Passenger Behavioural Intentions, Perceptions, Satisfaction and Involvement. *Transportation Research Board 99th Annual Meeting/Transportation Research Record*.



CAPÍTULO 2.

ESTADO DEL ARTE

Capítulo 2.

ESTADO DEL ARTE

El análisis de la calidad del servicio en el campo del transporte público se trata de un tema bastante conocido y desarrollado en la literatura. En este capítulo se muestra una revisión del estado del arte del mismo, desarrollando el concepto propio de calidad del servicio, exponiendo las diferentes singularidades que existen en el campo del transporte público, y presentando las principales vías de recolección de datos e información, así como las diferentes técnicas estadísticas que se utilizan para analizar dicha calidad del servicio en el transporte público, ya tengan su origen, o no, en la Minería de Datos. Se destaca que se ha estudiado de forma más detallada los casos de las técnicas de RB y SEM.

Finalmente, se concluye con una breve descripción y resumen de la literatura actual sobre trabajos que versan sobre calidad del servicio en el transporte público, poniendo el foco de atención en el modo de transporte y país, objetivos perseguidos, atributos utilizados, tipología de encuesta empleada, periodicidad de aplicación, número de encuestas realizadas, escalas de medida utilizadas, tecnología de toma de datos y metodología de análisis. Dicha revisión motivará y pondrá de relieve la novedad de dicho proceso metodológico de dos etapas propuesto en esta tesis doctoral, así como su necesidad de aplicación para resolver ciertas limitaciones que se presentan, actualmente, no solo en el análisis de la calidad del servicio en el campo del transporte público, sino también en cualquier ámbito de conocimiento.

2.1. LA CALIDAD DEL SERVICIO EN EL TRANSPORTE PÚBLICO

Desde hace tiempo, las sociedades sufren graves problemas asociados con la movilidad debido al excesivo uso del vehículo privado (p.e., contaminación, atascos, accidentes, ruido, etc.) (Greene and Wegener, 1997; Sari et al., 2017; Steg, 2003; Woods and Masthoff, 2017), y a la alta dependencia que la mayoría de las personas tienen del mismo (Anable, 2005). Dicha situación, hace pensar en la necesidad de una alternativa sostenible y de calidad, que produzca un cambio de mentalidad y reduzca el número de personas que cada día usan su vehículo privado para ir al trabajo, a la Universidad, a realizar una actividad de ocio, etc. Debido a esta razón, gobiernos, consorcios y operadores del transporte han puesto su foco de atención en potenciar el transporte público, de forma que sea capaz de asumir este rol tan demandado hoy en día (Holmgren, 2007).

De acuerdo a diversos autores (p.e., Beirão and Cabral, 2007; dell’Olio et al., 2011; Eboli and Mazzulla, 2007; Nathanail, 2008; etc.), una forma de alcanzar dicho objetivo es a través de la evaluación y mejora de la calidad del servicio. Aunque la forma de la relación “Calidad-Satisfacción”, en ocasiones, haya sido cuestionada y discutida (Friman, 2004; Friman and Fellesson, 2009; Friman et al., 2001; Mattsson, 1992), un amplio abanico de trabajos pueden ser encontrados en la literatura (p.e., Bolton and Drew, 1991; Eboli and Mazzulla, 2007; 2008; Lai and Chen, 2011; Nandan, 2010; Nilsson et al., 2001; Simester et al., 2000; etc.) que demuestran que un servicio que se preste con un alto nivel de calidad produce un alto grado de satisfacción en los pasajeros, lo que deriva a su vez en un aumento del uso de dicho servicio, la atracción de nuevos usuarios y una mejora de su imagen pública (Constantine et al., 2006; Ennew and Binks, 1996; Transportation Research Board, 1999).

En este contexto, donde es reconocido ampliamente que la calidad es un factor determinante sobre el comportamiento de los viajeros (Cascetta and Carteni, 2014), se comprende que la Unión Europea haya puesto su foco de atención en la calidad del servicio en el transporte público, convirtiéndose en una de sus principales prioridades el desarrollo de políticas de transporte sostenible, dado que anima a los viajeros a seleccionar modos de transporte más eficientes en energía y espacio (European Commission, 2007).

En el año 2002, el Comité de Estandarización de la Unión Europea, promulgó la norma EN 13816 (2003), donde se establecen los dos principales puntos de vista con los que se puede llevar a cabo la evaluación de la calidad del servicio.

Por un lado, se encuentra el punto de vista de los proveedores del servicio u operadores. Se trata del enfoque más tradicional, y para su evaluación se basa en el cálculo de los costes de eficiencia y efectividad del servicio de transporte público y sus operaciones (Carter and Lomax, 1992; Chu et al., 1992; Fielding, 1992; Fielding et al., 1985, 1978; Hensher and Daniels, 1995; Karlaftis, 2004; Miller, 1980; Ozment and Morash, 1998; Perry and Babitsky, 1986; Pullen, 1993; Talley and Becker, 1982; Wipper, 1993).

Y por el otro lado, se haya el punto de vista de los pasajeros, que ha cobrado mucha importancia durante los últimos años, entre los investigadores, proveedores del servicio y operadores (Barabino et al., 2012; Chang and Yeh, 2002; de Oña et al., 2012; dell’Olio et al., 2010; Hensher et al., 2003; Hu and Jen, 2006; Eboli and Mazzulla, 2008; Isikli et al., 2017; Nathanail, 2008; Too and Earl, 2010; Zakaria et al., 2010). Desde este enfoque, los mismos pasajeros evalúan la calidad del servicio basándose en sus propias percepciones y expectativas. Dicho de otra forma, se basan en un rango de mediciones desagregadas sobre el funcionamiento del servicio, que pueden ser usadas para medir la habilidad de los operadores en ofrecer servicios que cumplan sus expectativas (Transportation Research Board, 1999).

A pesar de que algunos autores han desarrollado diferentes metodologías para combinar ambos enfoques (de Oña et al., 2016a; Eboli and Mazzulla, 2011; Nathanail, 2008; Hassan et al., 2013; Tyrinopoulos and Aifadopoulou, 2008), el punto de vista más utilizado desde la década de los 80, es el de los pasajeros (Grönroos, 1990), ya que, como Berry et al. (1990) argumenta, y apoya Das and Pandit (2014), “los consumidores son los únicos jueces de la calidad del servicio”.

Como consecuencia, se puede afirmar que la evaluación de la calidad del servicio desde el punto de vista de los pasajeros, permite la identificación de los diferentes atributos que afectan al conjunto de la calidad, así como las posibles relaciones entre los mismos, y su mayor o menor influencia sobre la calidad global.

Dicha información constituye la base para el desarrollo de posibles estrategias de marketing, señalando, al mismo tiempo, dónde los operadores, planificadores y gobiernos, deben localizar sus esfuerzos e inversiones para implementar posibles mejoras.

Actualmente, una de las principales y más populares herramientas que se usan para la recolección de las percepciones y/o expectativas de los pasajeros sobre la calidad del servicio, son las denominadas “Encuestas de Satisfacción” (de Oña and de Oña, 2014; Marcucci and Gatta, 2007; Transportation Research Board, 1998, 1999). No obstante, se debe destacar que a través de estas encuestas se puede recoger diferentes tipologías de datos con respuestas y/o escalas de medida específicas, y que, de acuerdo a estos datos, se puede aplicar unas técnicas estadísticas u otras, así como obtener diferentes clases de información.

Además, se debe considerar el hecho de que los datos recogidos mediante dichas encuestas van a reflejar una serie de singularidades y características relevantes del propio concepto de calidad del servicio, que si no son consideradas o tratadas, pueden influir negativamente en los resultados que se extraigan. Como ejemplos de dichas singularidades y/o características, se puede mencionar la heterogeneidad de los datos (Beirão and Cabral, 2008), el tipo y número de atributos considerados (Parasuraman et al., 1985), la composición de las dimensiones que explican la calidad del servicio (de Oña et al., 2017), la consideración de percepciones, expectativas y/o importancias, así como el significado de las expectativas (Landrum and Prybutok, 2004; Nathanail, 2008), la paradoja de la relación Calidad-Satisfacción-Lealtad/Intención de Recompra (Friman, 2004), etc.

En conclusión, se pone de manifiesto que, a la hora de estudiar y analizar la calidad del servicio en el transporte público desde el punto de vista del pasajero, es necesario realizar una serie de consideraciones, a priori, sobre la tipología de la encuesta, basadas en los objetivos e información a conseguir, así como un tratamiento adecuado de los datos extraídos de las mismas. Esto permitirá en un análisis posterior, la modelización de la calidad de forma adecuada, precisa y generalizable, en la medida de lo posible.

Dicho lo cual, esta revisión del estado del arte pone su foco de atención en analizar el concepto de calidad del servicio, explorando sus diferentes singularidades y características propias, los tipos de encuestas de satisfacción que se utilizan actualmente, y las diferentes metodologías y/o técnicas estadísticas que se aplican para su análisis y modelización.

2.1.1. DEFINICIÓN DEL CONCEPTO DE CALIDAD DEL SERVICIO

Gran cantidad de esfuerzo ha sido puesto en la difícil tarea de definir el concepto de calidad del servicio. De acuerdo a lo recogido en el diccionario, que un servicio sea de calidad significa que es “Sin defectos – Realizado correctamente desde el comienzo”.

Si se analiza la literatura, se observa un amplio abanico de posibles definiciones, como por ejemplo, Crosby (1979) define la calidad como algo que es “conforme a los requisitos”. Marcucci and Gatta (2007) la definen en función de la utilidad, de forma que a mayor nivel de calidad prestado, mayor utilidad provista para los pasajeros. Y una visión más objetiva de su definición puede ser encontrada en Garvin (1983), que realiza un balance entre los fallos en el producto antes de salir de su lugar de creación (“Fallos internos”) y los cometidos siendo usado (“Fallos externos”), de forma que la diferencia entre éstos es la calidad del producto.

Una definición, que ha sido ampliamente contrastada en la literatura (Czepiel, 1990; Grönroos, 1988; Joewono and Kubota, 2007; Lai and Chen, 2011; Lehtinen and Lehtinen, 1982; Lewis and Booms, 1983; Liljander and Strandvik, 1993; Sachdew and Verma, 2004; Sasser et al., 1978; Transportation Research Board, 1999; 2013), es la expresada por Parasuraman et al. (1985), el cual plantea la idea de que una empresa presta un servicio de calidad si éste cumple las expectativas de los consumidores considerando una base consistente.

Como se puede observar en las diferentes definiciones de calidad, las percepciones y las expectativas tienen un rol relevante. No obstante, este hecho no es generalizado, ya que otros autores, solo consideran las percepciones de los pasajeros o de los operadores para definir la calidad del servicio, haciendo caso omiso a las expectativas (Cronin and Taylor, 1992; Eboli and Mazzulla, 2011; Nathanail, 2008; Tyrinopoulos and Aifadopoulou, 2008).

En consecuencia, ante esta dificultad para su definición, se puede establecer que la calidad del servicio se trata de un concepto complejo, difuso y abstracto, así como latente, es decir, que no puede ser medido directamente, sino a través de sus diferentes atributos y/o dimensiones. Además, para su completo entendimiento, es necesario conocer sus tres propiedades inherentes: Inseparabilidad, Intangibilidad y Heterogeneidad (Carman, 1990; Parasuraman et al., 1985; Zeithaml, 1988).

Se puede observar que cada una de ellas se cumple en el campo del transporte público de la siguiente manera:

- **Inseparabilidad:** El transporte se trata de un servicio que es producido y consumido en el mismo acto. No se trata de un objeto físico que se produce en la fábrica, se almacena y se envía, posteriormente, al destino final en perfecto estado. El transporte público se produce, se vende y se consume una sola vez y al mismo tiempo.
- **Intangibilidad:** Los resultados de prestar el servicio de transporte es algo que no puede ser medido, contado, inventariado, almacenado, etc., en términos de atributos físicos, ya que se trata de funcionamiento de un servicio más que de un objeto, y para los pasajeros se trata de una experiencia.
- **Heterogeneidad:** Debido que se trata de un servicio dirigido a una gran cantidad de diferentes tipologías de pasajeros, que poseen sus propias características y contextos socioeconómicos, actitudes, hábitos, etc., es lógico afirmar que cada individuo que consume este servicio puede percibir el servicio de forma totalmente única y diferente, aunque el servicio sea el mismo. De la misma forma, no todos los servicios prestados son iguales. Estos pueden variar tanto en un mismo día (horas), como a lo largo del tiempo (días, meses, años, etc.).

Por tanto, y a modo de conclusión, se puede considerar que la calidad del servicio se trata de un concepto de difícil definición, así como latente, ya que su evaluación se puede realizar únicamente a través de una determinada cantidad y variedad de atributos y/o dimensiones. En el caso específico del transporte público se acepta, de forma más o menos generalizada, que dichos atributos y/o dimensiones son evaluados a través de las percepciones de los pasajeros o de una comparación de éstas con las expectativas de los mismos. Este hecho, pone de relieve la importancia de cómo el pasajero recibe la calidad, más que de la calidad en sí misma como resultado (Grönroos, 1984; Parasuraman et al., 1985; 1988). Similarmente, se destaca que el concepto de calidad del servicio posee una naturaleza puramente subjetiva (Carman, 1990).

Todas estas características intrínsecas de la calidad del servicio hacen que su estudio presente una serie de singularidades y características, que deben ser consideradas en todas las fases del mismo y/o ser tratadas, en la medida de lo posible, mediante diferentes técnicas estadísticas, de forma que no repercutan en los resultados finales y conclusiones.

2.1.2. SINGULARIDADES DE LA CALIDAD DEL SERVICIO

Diversas consideraciones y/o aplicaciones de técnicas estadísticas deben ser llevadas a cabo, si se desea evitar que los resultados obtenidos de un análisis de la calidad de un determinado transporte público se encuentren sesgados o no reflejen la realidad.

Para tal fin, previo análisis, es necesario identificar y conocer las posibles singularidades de las que se derivan estas posibles influencias y sesgos.

2.1.2.1. SELECCIÓN DE ATRIBUTOS Y/O DIMENSIONES

La primera singularidad que se presenta en estas tipologías de análisis, dada la inmediatez con la que hay que lidiar con ella, es la elección de los atributos para la evaluación de la calidad del servicio. O dicho con otras palabras, la selección de aquellos atributos y/o dimensiones¹ que componen dicha calidad del servicio.

Se trata de un punto de gran controversia, ya que actualmente no existe un consenso sobre qué atributos y/o dimensiones deben ser seleccionados, ni el número de ellos (Brady and Cronin, 2001). Sin embargo, es de aceptación general que la calidad del servicio es multidimensional o una composición de multi-constructos de atributos (Cronin and Taylor, 1992; Grönroos, 1982; Parasuraman et al., 1985, 1988), y multinivel o jerárquico (Dabholkar et al., 1996; Jen et al., 2011).

Diferentes investigaciones han tratado de establecer diferentes categorías o “listas” de atributos y dimensiones que componen o influyen en la calidad del servicio (Andreassen, 1995; Eboli and Mazzulla, 2008; Haywood-Farmer, 1988; Hensher et al., 2003; Joewono and Kubota, 2007; Philip and Hazlett, 1997; Redman et al., 2013; Tripp and Drea, 2002; Tyrinopoulos and Antonious, 2008).

Uno de los primeros trabajos sobre esta singularidad, y que más reconocimiento y aceptación tiene, es el de Parasuraman et al. (1985, 1988), que está basado en la desconfirmación del modelo de expectativas (Oliver, 1980). Este trabajo propone una lista de 22 atributos, que se pueden agrupar en 5 dimensiones de la calidad del servicio (Equipamiento tangible, Fiabilidad, Capacidad de respuesta, Cortesía y Empatía) y conforman la escala SERVQUAL (“SERVice QUALity”).

¹Las dimensiones son agrupaciones de un conjunto de atributos en base a una temática común subyacente.

Gran variedad de trabajos en la literatura se han basado en esta escala para evaluar la calidad del servicio en el transporte público (Abdlla et al., 2007; Kiatcharoenpol and Laosirihongthong, 2006; Randheer and Al-Motawa, 2011; Park et al., 2004; Too and Earl, 2010), la calidad del servicio en otros campos (Behara et al., 2002; Farooq, 2016; Lai, 2006) e incluso han propuesto variantes de esta escala SERVQUAL (Awasthi et al., 2011; Barabino et al., 2012; Cavana et al., 2007; Cronin and Taylor, 1992).

Su gran uso, tanto en el campo del transporte público como en otros campos, es debido a que el desarrollo de dicha lista fue realizado en base a una serie de entrevistas con diferentes profesionales de varios ámbitos, con el fin de que pudiera ser utilizada para cualquier tipo de servicio (Parasuraman et al., 1985; 1988)

Hu and Jen (2006) propusieron una escala de calidad del servicio de autobús con 20 atributos, agrupados en 4 dimensiones: “Interacción con los pasajeros”, “Equipamiento del Servicio Tangible”, “Conveniencia del servicio” y “Soporte Operativo de Gestión”.

La primera se refiere al respeto y cuidado que los pasajeros sienten cuando interactúan con los operarios y cómo éstos solucionan los problemas. La segunda se relaciona con los niveles de confort de los equipamientos e instalaciones. La tercera concierne a temas de accesibilidad, información y conveniencia de uso del servicio. Y la cuarta trata sobre elementos de horarios de servicio, periodos de funcionamiento, empleados, etc.

Redman et al. (2013), tras una revisión bibliográfica de los diferentes atributos utilizados en el estudio de la calidad del servicio en el transporte público, obtuvieron una lista de los 12 atributos más empleados. Estos se pueden categorizar en “Físicos” (Fiabilidad, Frecuencia, Velocidad, Accesibilidad, Tarifa, Información, Facilidad de transferencia o intercambio y Condiciones del vehículo) y “Percibidos” (Confort, Seguridad, Conveniencia y Estética).

Los primeros son medidos sin tener que involucrar a los pasajeros, y se realizan hipótesis sobre cómo dichos atributos inciden sobre los pasajeros. Para los segundos, por su parte, es necesaria la respuesta de los pasajeros, de forma que sus valores pueden ser obtenidos directamente (preguntas o encuestas a los pasajeros) o indirectamente (derivadas a partir de otros aspectos medibles por los pasajeros).

Incluso diferentes normas o manuales han propuesto sus propias recomendaciones sobre los atributos a considerar. La norma UNE-EN 13186 (2003) propone 117 atributos agrupados en 30 sub-dimensiones y 8 dimensiones (disponibilidad, accesibilidad, información, tiempo, atención al cliente, confort, seguridad e impacto ambiental), y clasifica los diferentes atributos del servicio en tres categorías (Básicos, Proporcionales y Atractivos), de acuerdo a los diferentes efectos sobre los pasajeros que puede tener su cumplimiento o no.

Por su parte, el Transit Capacity and Quality of Service Manual (Transportation Research Board, 2013) agrupa los atributos en dos clases: “Atributos de disponibilidad” y “Atributos de confort y comodidad”. En este caso, la primera clase lo conforman los atributos mínimos necesarios para que se realice el servicio, cuya no disponibilidad produce que el resto de atributos no revistan importancia alguna (p.e., Disponibilidad Espacial y Temporal, Información y Capacidad). La segunda clase la conforman el resto de atributos, que son menos importantes para los pasajeros (p.e., Seguridad, Fiabilidad, Coste, Apariencia, Confort, Ocupación y Puntualidad).

Se puede comprobar en la literatura que existe una amplia variedad de propuestas sobre este punto. Sin embargo, dichas “listas” genéricas han sido puestas en entredicho en más de una ocasión por ser inapropiadas (Cronin and Taylor, 1992), ya que cada caso de estudio presenta unas características y contextos diferentes, que influyen en la selección de atributos a emplear (Carrillat et al., 2007). Este hecho ha derivado en que dichas listas hayan sido criticadas por la comunidad científica (Babakus and Boller, 1992; Carman, 1990; Carrillat et al., 2007; Cronin and Taylor, 1992; Finn and Lamb, 1991; Gilbert and Wong, 2003; Holmbergh et al., 1991; Jabnoun and Khalifa, 2005; Lai, 2006; Teas, 1993; Triplett et al., 1994; Van Dyke et al., 1997).

Por tanto, la mayoría de autores han llegado a la conclusión de que, para cada caso de estudio, será necesario realizar un análisis exhaustivo de los atributos y/o dimensiones más importantes en términos de calidad del servicio en el transporte público, seleccionando aquellos que más se ajusten al mismo (Babakus and Boller, 1992; Carman, 1990; Cavana et al., 2007; de Oña and de Oña, 2014).

Dicho lo cual, diferentes técnicas han sido usadas para identificar estos atributos, como por ejemplo, “focus group” (Alonso et al., 2018; Cunningham et al., 2000; dell’Olio et al., 2011; Hu and Jen, 2006), revisión de la literatura (Cunningham et al., 2000; Geetika and Nandan, 2010; Redman, 2013; Shiftan et al., 2008), encuestas previas a operadores (Andreassen, 1995; Chen and Chang, 2005; Hensher et al., 2003), encuestas pilotos a los pasajeros (Liou et al., 2011), paradigma de Churchill’s (Brown et al., 1993; Churchill, 1979; Hu and Jen, 2006), test estadísticos como análisis factorial exploratorio (AFE) y confirmatorio (AFC), etc., o incluso en algunos casos, una combinación de varias de estas técnicas (Abenoza et al., 2017; dell’Olio et al., 2010; 2011; Hensher and Prioni, 2002; Hensher et al., 2003; Hu and Jen, 2006; Kuo and Tang, 2013; Marcucci and Gatta, 2007; Sanchez-Perez et al., 2007b; Tyrinopoulos and Aifadopoulou, 2008).

2.1.2.2. CONOCIMIENTO TEÓRICO O EXPERTO PREVIO

Una vez seleccionados los atributos y/o dimensiones que van a conformar el análisis de la calidad del servicio, es necesario establecer las posibles relaciones entre los mismos, lo cual, se realiza generalmente recurriendo a la literatura o a conocimiento experto (Bagozzi, 1994; Eboli and Mazzulla, 2012; Fillone et al., 2005; Kamaruddin et al., 2012; etc.). Sin embargo, esta forma de proceder puede conllevar a que ciertas relaciones importantes para explicar la calidad del servicio sean obviadas o se realicen por otros caminos a los establecidos.

Focalizando en los casos específicos de análisis de la calidad del servicio en el transporte público, se observa que la única pauta en la literatura para esta fase, es que la comunidad científica simplemente está de acuerdo en el hecho de que la calidad del servicio está relacionada e influenciada directamente por todos sus atributos y/o dimensiones. No obstante, algunos trabajos recientes ponen de manifiesto que la relación de dichos atributos y/o dimensiones con la calidad del servicio no siempre tiene que hacerse de forma directa, sino que también puede existir influencia de forma indirecta (Chou and Kim, 2009; de Oña et al., 2017; Rahman et al., 2016; etc.).

Un ejemplo claro puede ser encontrado en de Oña et al. (2017), donde se agrupan los atributos del servicio en dos dimensiones latentes, nombrados como atributos primarios (Factores del Servicio de Transportes), y secundarios (Factores de Confort y Conveniencia). En dicho trabajo, se evidencia que los atributos secundarios ejercen un efecto sobre los primarios, y su relación con la satisfacción global del pasajero no se ejerce en un camino directo, sino indirecto a través de los primarios.

Estos trabajos refuerzan la idea anteriormente expuesta, por la que en ciertas ocasiones no es suficiente con recurrir a la literatura o conocimiento experto para establecer las relaciones entre los atributos y/o dimensiones con la calidad del servicio, ya que ciertas relaciones importantes pueden ser obviadas.

2.1.2.3. EXPECTATIVAS

A la hora de estudiar la calidad del servicio en el transporte público se debe realizar una consideración importante acerca del concepto de expectativas, ya que, lo que son y lo que realmente miden, presentan diversas interpretaciones en la literatura.

De acuerdo a Olson and Dover (1979), las expectativas se pueden definir como una serie de creencias previas sobre un determinado servicio, sirven como estándares o puntos de referencia, y mediante su comparación con las percepciones se puede establecer el nivel de calidad del servicio.

No obstante, esta forma de entender las expectativas plantea cuestiones sobre el rol que realmente desempeñan. Hay autores que han utilizado y entendido estas expectativas de diferentes formas (Barabino and Di Francesco, 2016), como por ejemplo, han sido interpretadas como “Estándar de Predicción”, es decir, como las predicciones que hacen los pasajeros sobre qué es probable que ocurra utilizando el servicio (Oliver, 1981; Prakash, 1984; Swan and Trawick, 1980). Otros como “Estándar Ideal” (Mattsson, 1992; Miller, 1977; Sanchez-Perez et al., 2007b), como “Calidad Deseada” (Anable, 2005; Cavana et al., 2007; Filipovic et al., 2009); o incluso como “Calidad Adecuada” o “Tolerable” (Fonseca et a., 2010; Hu and Jen, 2006; Liljander and Strandvik, 1993).

Si se atiende a la norma UNE-EN 13816 (2003), el concepto de calidad del servicio es definido mediante 4 visiones diferentes de esta: la calidad esperada por el pasajero, identificable con sus expectativas, la calidad objetivo, que es la que el titular y el operador del servicio pretenden proporcionar a los pasajeros, la calidad producida/entregada, que es el nivel de calidad alcanzado, y la calidad percibida, que es la interpretada por los pasajeros durante el trayecto. De acuerdo a esta norma el proveedor del servicio deberá de minimizar las discrepancias entre la calidad del servicio esperada, objetivo, entregada y percibida, siendo de vital importancia que los proveedores entiendan las necesidades de los pasajeros.

Teas (1993, 1994) establece que las expectativas podrían ser interpretadas como medidas de importancia de los atributos. Dicho concepto ha sido trasladado y utilizado en el ámbito del transporte público, donde varios investigadores (Chen and Chang, 2005; Eboli and Mazzulla, 2010; 2015; Filipovic et al., 2009; Gilbert and Wong, 2003; Hill et al., 2003; Sultan and Simpson, 2000) han sustituido las medidas de importancia por expectativas. No obstante, Landrum and Prybutok (2004) encontraron que la importancia y expectativas no son el mismo concepto, por lo que no existiría una base teórica para esto.

Similarmente, en la literatura se pueden encontrar diversos autores que no toman en cuenta las expectativas a la hora de definir dicha calidad, sino que solamente son consideradas las percepciones, ya sean de los pasajeros o de los mismos operadores y gobiernos (Barabino et al., 2012; Cronin and Taylor, 1992; Eboli and Mazzulla, 2011; Eboli et al., 2016; Hassan et al., 2013; Kim et al., 2005; Nathanail, 2008; Randheer and Al-Motawa, 2011).

En consecuencia, la explicación sobre qué son las expectativas y cuál va a ser su significado o rol, debe ser establecido previo comienzo del estudio, consiguiendo de esta forma que los resultados obtenidos sobre la calidad percibida por el pasajero, sean explicados en base a estas consideraciones.

2.1.2.4. HETEROGENEIDAD

Un problema importante, nacido de estudiar la calidad del transporte público desde el punto de vista del pasajero, es la heterogeneidad, que se muestra como una característica latente propia de la subjetividad de las percepciones de los pasajeros, reflejadas en los datos.

Este hecho se debe principalmente a que, si consideramos la definición de calidad del servicio como una comparación entre la percepciones y expectativas de cada pasajero, se puede observar que la propia situación, cultura y contexto sociodemográfico de cada uno de ellos tendrá influencia en las percepciones y en las expectativas sobre el servicio (Anable 2005; Burguillo et al., 2017; dell’Olio et al., 2011; Furrer et al., 2000; Morton et al., 2016), pudiendo variar éstas, incluso a lo largo del tiempo (Zeithaml et al., 1993).

Este hecho se muestra como realmente importante, puesto que se ha comprobado que la eficacia de las diferentes estrategias de marketing en el transporte público, dependen de si esta singularidad ha sido tenida en cuenta y, en consecuencia, las medidas de mejora implantadas se adaptan a las diferentes tipologías de pasajeros (Bamberg et al., 2011; Daniels and Mulley, 2013; dell’Olio et al., 2010).

Existen diferentes métodos para capturar esta singularidad (de Oña et al., 2016c), como por ejemplo, la utilización de modelos de elección discreta con parámetros aleatorios (Hensher et al., 2010), donde la variación en la percepción de los pasajeros es incorporada en los parámetros del modelo; o la segmentación de la muestra de usuarios en categorías de pasajeros con opiniones del servicio más homogéneas.

En el ámbito del transporte público, dos tipologías de segmentación en categorías se erigen como las más populares: la primera es la “Segmentación Tradicional”, que está basada en utilizar cierta característica socioeconómica, actitud o hábito de viaje para tal fin (Andreassen, 1995; Daniels and Mulley, 2013; dell’Olio et al., 2010; de Oña et al., 2015a; Hine and Scott, 2000; Sanchez-Perez et al., 2007a). Y la segunda es el Análisis Cluster, que se trata de una técnica estadística con origen en la Minería de Datos, y cuya principal aplicación es segmentar los datos en sub-grupos (denominados clusters) mediante la utilización de varias características o variables de forma conjunta, de manera que se presenten observaciones muy homogéneas dentro de cada cluster, y al mismo tiempo, éstas sean lo más heterogéneas posibles entre diferentes clusters (Abenoza et al., 2017; Chowdhury et al., 2018; de Oña et al., 2016c; Nesheli et al., 2017; Vicente and Reis, 2016; Wen et al., 2008).

Por tanto, dicha singularidad debe ser tratada adecuadamente, en caso contrario, al obtener un modelo general sobre la calidad del servicio, la fiabilidad de los resultados puede verse reducida, por lo que las posibles estrategias de marketing o mejoras indicadas para elevar la calidad del servicio puede que no produzcan el efecto deseado o simplemente lo hagan a una escala menor de la prevista. Esto resultaría bastante perjudicial para los operadores y gobiernos, ya que los esfuerzos e inversiones aplicados para conseguir los objetivos marcados sólo se conseguirían de forma parcial. Igualmente, la segmentación de los usuarios en grupos más homogéneos permitirá alcanzar uno de los actuales y más ambiciosos objetivos, que cada día, más y más empresas se fijan como meta, que es el logro del marketing personalizado.

2.1.2.5. PARADIGMA CALIDAD DEL SERVICIO-SATISFACCIÓN-INTENCIÓN DE RECOMPRA Y OTRAS ACTITUDES HACIA EL TRANSPORTE PÚBLICO

Comprender las intenciones de comportamiento de los pasajeros (entendido como Intención de recompra o Lealtad) después de haber utilizado un determinado servicio de transporte público es fundamental para cualquier gestor del transporte. Esta información les permite desarrollar estrategias efectivas que cumplan con las necesidades de los pasajeros, para retener e incrementar el uso de los ya existentes y atraer a nuevos pasajeros de otros modos de transporte (Lai and Chen, 2011). El estudio de estas Intenciones de recompra está basado en una serie de constructos latentes actitudinales (de Oña et al., 2016b). Se podrían destacar, por ser los constructos más analizados en la literatura, los siguientes: calidad del servicio, satisfacción, costes y beneficios percibidos, alternativas atractivas y actitud general hacia el transporte público (de Oña et al., 2016b; Jen and Hu, 2003; Lai and Chen, 2011; Machado-Leon et al., 2016).

Si se analizan en detalle los conceptos detrás de dichos constructos, se debe realizar una mención especial para los constructos de satisfacción y calidad del servicio, ya que presentan gran similitud en su naturaleza debido a que ambas derivan de la teoría de la disconformidad (Parasuraman et al., 1988). Este hecho ha llevado a que, en más de una ocasión, sus significados se hayan intercambiado en la literatura (p.e., Cavana et al., 2007; Howat et al., 1996; Oliver, 2010, etc.), así como la forma en la que se relacionan. Debido a esta razón, simplemente se puede afirmar que la mayoría de los autores aceptan que la calidad del servicio es un factor determinante de la satisfacción del pasajero, aunque no consigue explicarla en su totalidad (Bansal and Taylor, 2015; Bolton and Drew, 1991; Chen, 2008; Chou and Kim, 2009; Cronin and Taylor, 1992; Eboli and Mazzulla, 2015).

Si se analiza más en profundidad la forma en que se relacionan, es posible encontrar ciertos trabajos que consideran a la calidad del servicio como una consecuencia de la satisfacción, es decir, que la influencia es ejercida directamente por la satisfacción sobre la calidad del servicio (p.e., Andreassen and Lindestad, 1998; Bitner, 1990; Bolton and Drew, 1991; etc.). Sin embargo, se debe destacar que la mayoría de los trabajos defienden la idea de que la calidad del servicio se relaciona como un antecedente de la satisfacción (p.e., Cronin and Taylor, 1992; Farooq et al., 2018; Li et al., 2018; McDougall and Levesque, 2000; Mugion et al., 2018; Park et al., 2004; etc.).

Entre estos trabajos se puede destacar el caso de Han et al. (2008), que investigan el rol de la calidad del servicio percibida como antecedente de la satisfacción en diferentes industrias (p.e., sectores bancarios, hospitales, tecnologías de la información, sectores de la educación, salones de belleza y compañías de aerolíneas). Otro trabajo que apoya esta forma de pensar es Oliver (2010), que define la calidad del servicio como una opinión cognitiva (pensamiento/opinión), mientras define la satisfacción como una opinión afectiva (gusto/placer), puramente experimental. Por tanto, la percepción cognitiva de la calidad del servicio precede a la opinión afectiva de la satisfacción.

Aceptado este punto, se puede afirmar que el conflicto comienza o se agranda con la consideración del constructo de Intención de recompra, que es definido por Oliver (1999) como el fuerte compromiso para volver a utilizar o re-comprar un determinado producto o servicio preferido en el futuro. Por tanto, se crea el llamado paradigma “Calidad del Servicio–Satisfacción–Intención de Recompra”, que sugiere que la satisfacción es un enlace entre la calidad del servicio y la intención de recompra (Chiou and Chen; 2012; Dabholkar et al., 2000; Jen et al., 2011). Sin embargo, existen evidencias en la literatura que establecen que la calidad del servicio podría tener influencia directa sobre la intención de recompra (Chou et al., 2014; de Oña et al., 2015b; Lai and Chen, 2011; Machado-Leon et al., 2016; Minser and Webb, 2010).

Así mismo, varios autores han tratado de explicar las Intenciones de recompra de los pasajeros a través de, no sólo la calidad del servicio y la satisfacción, si no también de otros constructos latentes, tal y como se ha comentado anteriormente (de Oña et al., 2016b).

Cada uno de estos constructos cubre diferentes ambitos que influyen en ese comportamiento de los viajeros. Los costes percibidos pueden ser definidos como lo que se renuncia o se sacrifica para utilizar un servicio (Zeithaml, 1988), incluyendo tanto aquello que tiene valor monetario como lo que no (Choi et al., 2004; Wang et al., 2004).

Por su parte, los beneficios percibidos son las principales ventajas que se obtiene del uso del servicio (Zeithaml, 1988). Las alternativas atractivas cubre las preferencias sobre otros modos de transporte alternativos al investigado (De Oña et al., 2016b). Y la actitud general hacia el transporte público indica la pre-disposición y opinión de los pasajeros hacia el transporte público en general y hacia la gente que los utiliza (Lai and Chen, 2011).

Similar a lo que ocurre con la calidad del servicio y sus dimensiones, no existe acuerdo en cómo se relacionan dichos constructos para explicar la Intención de recompra, ya que éstos se conforman en base a valoraciones subjetivas respecto a diferentes items, pueden ser estas valoraciones diferentes para cada caso de estudio. En consecuencia, la solución a dicha limitación ha sido la de recurrir a la literatura en busca de conocimiento o pautas que sirvan como guía.

Por tanto, se establece que las “hipotéticas” relaciones entre los constructos mencionados deberían ser consideradas y testeadas específicamente para cada caso de estudio. Por ello, una solución ideal a dicha singularidad es la extracción, mediante alguna técnica estadística, de las relaciones existentes entre dichos constructos directamente de los datos en cada caso de estudio, evitando tener que proponer hipótesis de partida basadas en la literatura y conocimientos expertos previos, que pueden no representar la realidad de un caso de estudio determinado.

2.1.3. METODOLOGÍAS PARA LA OBTENCIÓN DE INFORMACIÓN SOBRE LA CALIDAD DEL SERVICIO

Bajo este contexto de dificultad y singularidad que presenta el concepto de calidad del servicio, diversos autores (Berry et al., 1990; Das and Pandit, 2014) y políticas de la UE (European Commission, 2001; 2007) han adoptado una visión de la calidad del servicio en el transporte público, orientada hacia el pasajero, es decir, centrándose en las necesidades y expectativas de los mismos. Esta situación demanda, de forma casi obligatoria, la obtención de dichas percepciones en la forma más actualizada posible. Siendo una situación ideal aquella en la que el operador, en base a su conocimiento del sistema, pudiese conocer la calidad percibida por los pasajeros en tiempo real. Sin embargo, por el momento, esto no es posible.

De acuerdo a la UNE-EN 13816 (2003) los métodos más útiles para la obtención de información sobre la calidad del servicio son los siguientes:

1. Panel de usuarios (focus groups), grupos de discusión y entrevistas en profundidad.

Los focus groups se tratan de grupos de pasajeros conocedores suficientes del servicio, que son seleccionados con el objetivo de obtener cada cierto tiempo sus sugerencias y opiniones partiendo desde sus propias percepciones y experiencias.

Los grupos de discusión reúnen personas (tanto pasajeros como no) que, de manera puntual, debaten sobre una guía previamente acordada con el operador con el fin de obtener datos cualitativos.

Por último, las entrevistas en profundidad se basan en charlas individuales entre un profesional y un pasajero con el objetivo de explorar exhaustivamente las expectativas de los pasajeros de un servicio.

2. Informes del personal en contacto con los clientes (grupos de diagnóstico).

Se basan en la obtención de información directamente de los empleados del servicio que tratan con los pasajeros. Dicha obtención puede ser realizada a través de encuestas, focus group, grupos de discusión, etc.

3. Investigaciones de mercado.

Tratan de averiguar oportunidades y amenazas a través de la evaluación de aspectos del servicio que realizan la totalidad de pasajeros potenciales del servicio. Una forma de realizar dicha investigación es a través de una entrevista a una muestra grande de personas con una frecuencia fija.

4. Cliente oculto o Control de pasajeros (Mystery Shopping).

Un equipo de observadores formados (o una simple persona) al efecto y que se comporta como verdaderos pasajeros que viajan en el sistema evaluado, realizan la medición de la calidad del servicio, de acuerdo a unos criterios específicos y aportando estimaciones objetivas.

5. Seguimiento de las muestras de insatisfacción.

Se realiza un estudio a través del buzón de sugerencias, de las quejas realizadas o de las “hojas de reclamaciones” interpuestas. Proporcionarán una información válida, al ser posible identificar las necesidades más destacables e importantes de los pasajeros en general.

6. Análisis de indicadores operativos internos.

Se trata de realizar la medición de la calidad a través de indicadores u otros métodos de seguimiento y medición internos de los operadores, de aspectos relacionados con la misma.

7. Medidas directas de la prestación.

Se basan en mediciones en tiempo real de alguna prestación real del servicio, ya sea mediante registro de las operaciones, o bien utilizando las observaciones de una muestra representativa, lo que la convierte en mediciones altamente precisas. Esta forma de proceder permite el seguimiento y la apreciación de resultados según las escalas definidas. Como ejemplo de este tipo de medidas, se encuentra los aforos o conteo de los pasajeros.

8. Encuestas de Satisfacción de clientes (Customer Satisfaction Surveys, CSS).

Se obtiene la información mediante cuestionarios a los pasajeros, que pueden haber sido seleccionados previamente mediante criterios. Se debe destacar que se trata de la herramienta más ampliamente usada para poder recolectar dicha información sobre la calidad del servicio, de una forma periódica, económica, altamente adaptable y de fácil comprensión para los participantes.

De hecho, estas encuestas de satisfacción gozan de una gran popularidad en el estudio de la calidad del servicio en el transporte público entre los trabajos de la literatura (Birago et al., 2017; de Oña et al., 2016c; Nandan, 2010; Nesheli et al., 2017; Suman et al., 2017; Yang et al., 2015; etc.). Además, la norma UNE-EN13816 (2003) y el "Transit Capacity and Quality of Service Manual" (Transportation Research Board, 1999) recomiendan su uso con este fin.

Se tratan de cuestionarios donde se les pregunta a los encuestados por un valor de satisfacción o de percepción del funcionamiento de los principales atributos del servicio. Estas encuestas suelen llevar asociados otros tipos de preguntas, principalmente, debido a los subsiguientes análisis y con motivo de que sirvan para analizar los datos (de Oña and de Oña, 2014). Entre las más comunes están las preguntas sobre la importancia de cada atributo, sobre la satisfacción global del servicio (Cavana et al., 2007; dell'Olio et al., 2010) o sobre las expectativas de cada atributo (Hu and Jen, 2006; Wang et al., 2010). Similarmente, también requieren una mención especial, aquellas preguntas que se emplean para obtener información sobre diferentes aspectos actitudinales que explican las intenciones de comportamiento de los pasajeros hacia el servicio de transporte público, como pueden ser preguntas relacionadas con el coste percibido, los beneficios percibidos, la existencia de alternativas de transporte atractivas, etc.

Numerosos autores han utilizado este tipo de preguntas para examinar de forma global este fenómeno (p.e., Andreassen, 1995; Christopher et al., 1999; Foote et al., 2001; Fu and Juan, 2017; Jen and Hu, 2003; etc.)

Por tanto, el tipo de información que puede ser abarcado con estas encuestas de satisfacción es bastante amplio y variado, ya que pueden incluir cuestiones sobre percepciones y expectativas (Hu and Jen 2006; Wang et al., 2010), importancias y grado de satisfacción (Cavana et al., 2007; Dell’Olio et al., 2010), actitudes e intenciones de comportamiento hacia el servicio (Chou et al., 2014; Park et al., 2006), etc.

Según la información lograda, se podrá conformar una determinada base, para que operadores, planificadores del transporte y gobiernos puedan llevar a cabo diferentes estudios y acciones sobre sus servicios. Como ejemplos se puede mencionar, la identificación de qué atributos de la calidad del servicio son considerados los más importantes y mejor percibidos por los pasajeros (Eboli and Mazzulla, 2011), la priorización de iniciativas de mejora de la calidad del servicio, midiendo el grado de éxito de anteriores, así como monitorear cambios en la calidad del servicio a lo largo del tiempo (Transportation Research Board, 2013), la modelización de comportamientos (Van Lierop and El-Geneidy, 2016), etc.

Por último, se debe mencionar que, generalmente, casi todas las CSS poseen una sección donde se le pregunta al pasajero por sus características o contexto socioeconómico, hábitos de viaje, etc. Esta parte provee la información necesaria para caracterizar la muestra sobre la que se está haciendo el análisis. Además, si las condiciones de representatividad son las adecuadas, se podrán realizar generalizaciones sobre los resultados obtenidos o, en caso de que se quiera realizar ciertos estudios condicionados a variables (p.e., edad, género, nivel de renta, etc.), esta sección proveerá la información necesaria para tales casos.

2.1.3.1. ASPECTOS GENERALES DE LAS ENCUESTAS

Todas las encuestas, independientemente de la tipología de preguntas que posean, presentan unos determinados aspectos a considerar, que deberían ser establecidos previa configuración de la versión definitiva de la misma y en concordancia a las necesidades de los operadores y gestores del servicio.

No obstante, se debe mencionar que esta recomendación no ocurre normalmente, sino que se obvian bastantes aspectos que pueden resultar relevantes. Es decir, se suelen emplear “diseños estándar” de encuestas, sin adaptar al contexto en el que se van utilizar, ni a las técnicas para la recogida de información u objetivos a conseguir.

Periodicidad.

El primer aspecto que se debe considerar es la periodicidad con la que se va a realizar dicha encuesta. El hecho de tener una periodicidad relativamente baja, permitirá un seguimiento más preciso y de una forma bastante continua de la evolución de la percepción de la calidad mediante la aplicación de diferentes técnicas estadísticas, como por ejemplo, el “Control Chart” (Kenett and Salini, 2011). Esta técnica se trata de un gráfico estadístico que permite realizar un seguimiento de la evolución de la percepción de la calidad del servicio o el análisis de una posible incidencia en dicha percepción que pueda tener una mejora implementada o una campaña de marketing. Sin embargo, se debe considerar que cuanto más baja sea la periodicidad, mayor coste y repercusión tendrá en la partida de la empresa destinada a este cometido.

En la literatura se puede encontrar una amplia diversidad de periodicidades, siendo las más comunes: anuales (Abenzoza et al., 2017; Ostrowski et al., 1993; Pedersen et al., 2011) y bianuales (Christopher et al., 1999; Figler et al., 2011; Minser and Webb, 2010). En menor grado, cada 5 años (Lopez-Bonilla and Lopez-Bonilla, 2008) y trimestrales (Stuart et al., 2000).

Tecnología de toma de datos.

En la actualidad, existen diferentes tecnologías para la realización de las encuestas, que ya han sido utilizadas en diversos estudios en la literatura, como por ejemplo, en papel con apoyo de encuestador (Aydin, 2017; Isikli et al., 2017; Joewono and Kubota, 2007), en papel sin apoyo de encuestador, con devolución al bajarse del transporte o envío postal (Friman, 2004; Liou et al., 2011; Murray et al., 2010), telefónica (Abenzoza et al., 2017; Nurul-Habib et al., 2011), on-line (Nesheli et al., 2017; Yang et al., 2012).

Su elección va a depender de diversos factores:

- Tipología de preguntas. Las encuestas con preguntas que requieren respuestas detalladas o buscan obtener una información específica es mejor realizarlas con encuestador, y demandan, por norma general, gran cantidad de tiempo su realización. Por otro lado, las encuestas de valoración numérica de aspectos del servicio son más flexibles.
- Lugar donde se realiza. Pueden llevarse a cabo durante el desplazamiento, previamente en la parada, en casa o al bajarse del transporte en cuestión. Normalmente, cuando se realizan en el transcurso de utilización del transporte público, suelen hacerse a papel o con Tablet, mientras que si es en casa se suelen realizar por teléfono o modalidad on-line.

- Recursos económicos disponibles. Los costes varían según el medio seleccionado para su realización, siendo las “on-line” las que menor coste presentan (Cobanoglu et al., 2001; Duffy and Smith, 2005). Como dato, se sabe que el coste por e-mail de encuesta representa entre un 5% y un 20% del coste de la encuesta en papel (Sheehan and Hoy, 1999; Weible and Wallace, 1998).

No obstante, se debe destacar que en ambos casos no es igual la ratio de respuesta, el tiempo de respuesta, la representatividad de ciertos grupos de pasajeros (p.e., personas mayores o sin acceso a internet), etc. (Kaplowitz et al., 2004; Sheehan, 2001).

- Extensión (número de preguntas) de la encuesta. Si la encuesta es muy extensa o el trayecto es corto se recomienda la realización on-line en casa, cuando el pasajero dispone de tiempo para poder contestarla con tranquilidad.
- Fiabilidad de los datos. Las encuestas postales u on-line poseen mayor grado de anonimato y evita la posible influencia del encuestador, consiguiendo mayor fiabilidad de determinadas preguntas (p.e., ingresos familiares). Además, las encuestas on-line o table-based evitan tener que transferir las posibles respuestas recogidas a papel, evitando los errores de transcripción (Duffy and Smith, 2005). Por otro lado, las encuestas con apoyo del encuestador, suelen asegurar un buen entendimiento por parte del encuestado (Cardamone et al., 2014).
- Contexto territorial. País de realización, su cultura, tamaño de la ciudad, etc. Por ejemplo, en países subdesarrollados no se aconseja la realización on-line de la encuesta, debido a que el alcance de dicho recurso puede estar limitado.

Escala de medida.

Así mismo, otro aspecto a tener en cuenta son las diferentes escalas, con que los pasajeros se expresarán. Este aspecto es fundamental cuando se está midiendo la calidad, donde habitualmente, los encuestados le dan un valor a cierto atributo de acuerdo a sus propias percepciones. Los datos que se pueden recoger en una encuesta se encuentran agrupados en 4 niveles o escalas de medición (Allen and Seaman, 2007):

- Nominales: El nivel más “débil” de medición en cuanto a que representa categorías sin numeración.
- Ordinales: En este nivel se puede realizar un ranking de categorías, pero no medir la distancia entre ellas.

- Intervalos: Se trata de datos enteros que pueden ser ordenados en un ranking, medir la distancia entre ellos y las unidades de medida son uniformes. En esta escala, el cero no indica ausencia de atributo (p.e., cuando se mide la temperatura en grados, el cero es relativo).
- Ratio: Se pueden ordenar, medir distancias, decimales y fracciones entre las variables. La única diferencia con la escala de intervalos es que el cero significa ausencia de atributo (p.e., cuando se mide el peso, el cero es absoluto).

Desde el punto de vista del encuestado, las escalas ordinales, en especial las escalas Likert, son las más sencillas e intuitivas, de ahí que sean las más usadas en las encuestas sobre calidad del servicio en el transporte público (Echaniz et al., 2018; Lai and Chen, 2011; Eboli and Mazzulla, 2009; Grönroos, 1984).

Las escalas Likert fueron desarrolladas en 1932 como una familia bipolar de 5 puntos (p.e., Muy importante, Importante, Neutral, Sin importancia, Nada Importante). A lo largo del tiempo han ido variando e introduciendo nuevos niveles, como es el caso de las escalas de 7 puntos Likert y de otros muchos ejemplos, tal y como se puede observar en la literatura (de Oña et al., 2013a; Stuart et al., 2000; Tsai et al., 2011). Además, se debe destacar que dichas escalas pueden ser tanto numéricas como lingüísticas, siendo las primeras más ampliamente usadas y variando de 3 a 11 niveles. En el caso de las lingüísticas varían de 3 a 7 niveles (de Oña and de Oña, 2014).

Allan and Seaman (2007) sugieren que lo mejor es usar escalas lo más amplias posibles, ya que siempre es posible agrupar a posteriori en un número menor las categorías. No obstante, ante los diferentes casos analizados, se puede concluir que cada encuesta o cada sección de la misma, debería establecer una escala de medición de acuerdo a los objetivos que se quieran conseguir con dichos datos.

Tamaño muestral.

Como último aspecto relevante a considerar, se encuentra la determinación de cuál es el tamaño de muestra adecuado para obtener una buena representatividad de la población de análisis, es decir, el número de pasajeros que hay que encuestar en el servicio de transporte público a evaluar. Este aspecto depende del flujo de viajeros del servicio, del margen de error que se quiera asumir, del nivel de confianza establecido, etc., pero sobre todo de los recursos económicos disponibles para tal fin.

Como se observa, la gran cantidad de aspectos a considerar cuando se configura una encuesta, y que, además, pueden estar condicionados por los sub-siguientes análisis y/u objetivos fijados, hacen que exista un amplio abanico de posibilidades. Este hecho muestra como paso fundamental el analizar las necesidades de los operadores del servicio, las características del servicio prestado y los objetivos que se quieran alcanzar, previo paso a la configuración de la encuesta a aplicar.

2.2. MÉTODOS DE ANÁLISIS DE LA CALIDAD DEL SERVICIO

Actualmente, el rango de diferentes métodos para analizar la calidad del servicio en el campo del transporte público es muy amplia y variada (Alonso et al., 2018). La elección de una u otra para analizar la información obtenida con las encuestas, viene condicionada por los objetivos que tenga el operador del servicio de transporte, así como por los datos obtenidos según la encuesta utilizada.

Se debe destacar que existen diferencias significativas entre los métodos utilizados en los ámbitos académicos y las prácticas habituales de los operadores del transporte público. Por un lado, los operadores emplean sencillos métodos de análisis de frecuencias o análisis de cuadrantes para evaluar los servicios prestados (Christopher et al., 1999; Figler et al., 2011). Por otro lado, los académicos utilizan métodos y técnicas más avanzadas que se ajustan más a los objetivos, a los datos disponibles y a las necesidades de estudio, por lo que los beneficios de utilizarlas inciden directamente en el desarrollo de estrategias de mejora del servicio más eficientes y, consecuentemente, una distribución de recursos más efectiva.

Dentro de la amplia gama de métodos de análisis que se emplean en la literatura para el estudio de la calidad del servicio, se puede observar que, cada vez más, se están utilizando técnicas cuyo origen está en la Minería de Datos, es decir, en el campo de las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos (Maimon and Rokach, 2010). Por tanto, una primera clasificación básica de los mismos podría basarse en este hecho.

En los siguientes apartados se verá una revisión bibliográfica de los principales métodos para la medición de la calidad del servicio, ya tengan su origen o no en la Minería de Datos, siendo el principal foco de atención los SEM, el Análisis Factorial y las RB, por ser técnicas que han sido utilizadas en esta tesis doctoral. Una profunda revisión de las metodologías más aplicadas e importantes se puede encontrar en de Oña and de Oña (2014) y Barabino and Di Francesco (2016).

2.2.1. MÉTODOS ESTADÍSTICOS TRADICIONALES

Se tratan de análisis basados en técnicas estadísticas que no recurren a la inteligencia artificial y/o aprendizaje automático, sino que están basadas en cálculos estadísticos tradicionales (p.e., test, comparaciones, estadísticos simples, etc.), y que han servido como base para el desarrollo de los métodos más modernos. Entre sus principales aplicaciones se pueden destacar la caracterización de bases de datos, la reducción o agrupamiento de un número de atributos en un número menor de dimensiones, la realización de comparaciones simples entre diferentes tipos de características, la verificación de modelos, etc.

Indagando en la literatura, se puede apreciar que la gran mayoría de los trabajos existentes de estudios de calidad del servicio en el transporte público han sido realizados con este tipo de métodos (p.e., Aksoy et al., 2003; Castillo and Benitez, 2013; de Oña et al., 2013a; dell’Olio et al., 2010; Eboli and Mazzulla, 2008; Echaniz et al., 2018; Hensher and Prioni, 2002; Hu, 2010; Hu and Jen, 2006; Tyrinopoulos and Aifadopoulou, 2008; Zeithaml et al., 1993; etc.). Se debe destacar que, en la mayoría de los casos, estos modelos presentan ciertas limitaciones, ya que tienen sus propias hipótesis y relaciones preestablecidas entre la variable dependiente y las variables independientes, de forma que, en el caso de que estas hipótesis no se cumplan, se podrían realizar estimaciones erróneas (De Oña et al., 2012).

Si se analizan las técnicas estadísticas empleadas por estos métodos, se puede realizar una sub-división de las mismas en simples o avanzadas, de acuerdo al grado de complejidad de los cálculos que realizan (Figura 1). De esta forma, se aprecia por un lado que, el sub-grupo de técnicas estadísticas simples cubriría el SERVQUAL, SERVPERF, IPA, Correlación Bivariada y otras técnicas estadísticas sencillas (p.e., ANOVA, TOPSIS, t-test, etc.). Y, por el otro lado, el sub-grupo de técnicas estadísticas complejas cubriría los SEM, Análisis de Regresión y Análisis Factorial, entre otros. A continuación, se realiza una breve introducción de cada una de ellas.

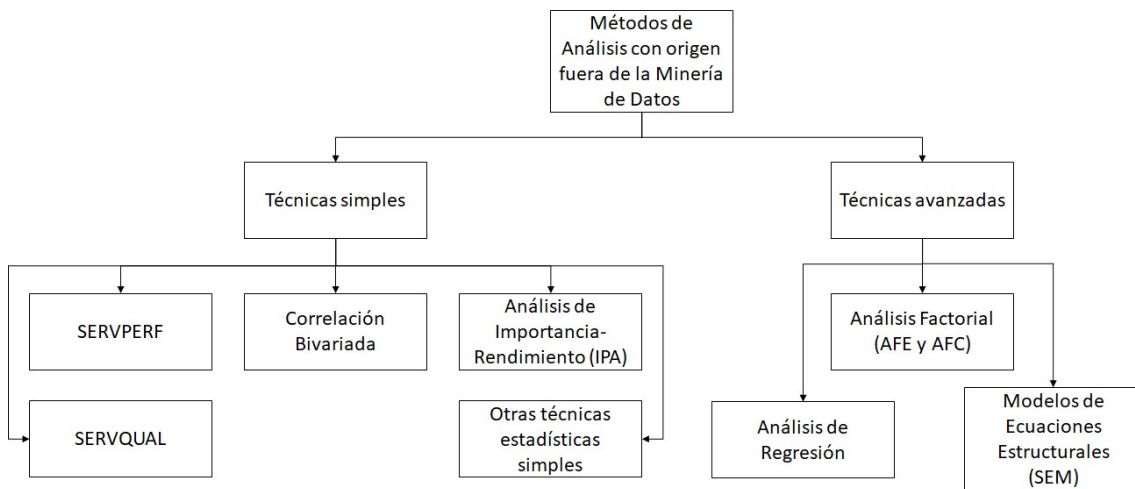


Figura 1. Resumen de los métodos de análisis con origen fuera de la Minería de Datos.

2.2.1.1. TÉCNICAS SIMPLES

Las técnicas cubiertas por este apartado fueron las primeras y más básicas que se han usado para analizar la calidad, así como para describirla y caracterizarla. Muchas de estas técnicas se siguen utilizando actualmente, ya sean como base para desarrollar estudios más complejos o como primer acercamiento a las bases de datos. En este sub-grupo destacan los modelos SERVQUAL y SERVPERF, la Correlación Bivariadas, así como el método IPA, dado su gran relevancia en el estudio de la calidad del servicio.

SERVQUAL.

Este sub-grupo se debe comenzar con uno de los primeros y más importantes modelos utilizados para la medición de la calidad del servicio, como es el SERVQUAL (SERVice QUALity). Se propuso en el trabajo de Parasuraman et al. (1988) y está basado en que la calidad del servicio es una función de la diferencia entre las expectativas y percepciones de los pasajeros. Sus resultados proporcionan un indicador de calidad que permite analizar el servicio en el tiempo y comparar diferentes tipos de servicios (p.e., diferente alcance territorial, proveedores, etc.). Se trata de una de las metodologías más ampliamente usadas (p.e., Awasthi et al., 2011; Cavana et al., 2007; Wen et al., 2005; etc.).

Debido a su popularidad se pueden encontrar en la literatura una amplia gama de variantes del SERVQUAL, como por ejemplo, el Índice de Satisfacción del Cliente (Customer Satisfaction Index, CSI) de Hill et al. (2003), que pondera los atributos para construir dicho índice en función de su importancia. Otro ejemplo puede ser el Índice de Satisfacción del Cliente Heterogéneo (Heterogeneous Customer Satisfaction Index, HCSI) de Eboli and Mazzulla (2009), quienes, basándose en el CSI, corrigieron las importancias y los valores de satisfacción en base a sus dispersiones.

SERVPERF.

Otro modelo que se debe destacar en este sub-grupo es el SERVPERF (SERVice PERFormance), que fue propuesto por Cronin and Taylor (1992) y nació como una crítica al SERVQUAL. Se trata de una variante del mismo, ya que utiliza los mismos atributos y dimensiones que el SERVQUAL, pero elimina las expectativas de los pasajeros para medir la calidad del servicio.

Se puede considerar como una de las metodologías fundamentales para la medición de la calidad del servicio, pues, al igual que el SERVQUAL, ha sido ampliamente aplicada en diversos estudios (p.e., Awasthi et al., 2011; Kuo, 2011; Yeh et al., 2000, etc.) y se han propuesto variantes de la misma (Saha and Theingi, 2009; Yao and Ding, 2011). Un ejemplo resaltable puede ser el de Sanchez-Perez et al. (2007a), que adaptó la escala y la ponderó mediante la importancia de cada uno de los atributos considerados.

Correlación Bivariada.

Como mención especial, se debe destacar la Correlación Bivariada, siendo una de las tipologías más usada en la literatura (Barabino et al., 2012; Figler et al., 2011; Weinstein, 2000) para establecer un ranking de importancia de los diferentes atributos que componen la calidad del servicio o la satisfacción.

No obstante, se debe tener cuidado en su lectura e interpretación de los coeficientes, debido a que esta técnica no tiene en cuenta la correlación entre atributos a pesar de que existen evidencias de la colinearidad entre los mismos (de Oña and de Oña, 2014).

Análisis de Importancia-Rendimiento (IPA).

El método IPA (Martilla and James, 1977) se trata de un análisis de cuadrantes, donde las coordenadas de cada atributo indican su importancia (eje vertical) y su rendimiento/valoración (eje horizontal) (Figura 1). Éste goza de gran popularidad dada su facilidad para trabajar con ella y a la intuitiva interpretación gráfica de sus resultados (p.e., Burns, 1986; Sampson and Showalter, 1999; Wei-Jaw, 2008; Weijawa, 2007; etc.).

Dicha gráfica se puede dividir en cuatro cuadrantes (I, II, III y IV) (Shen et al., 2016). Los operadores y planificadores del transporte pueden interpretar y localizar los atributos en función a su posición en los cuadrantes. Es decir, en el cuadrante I están localizados los atributos con alto rendimiento/valoración e importancia, por lo que mantener su estado sería lo deseable. En el cuadrante II están los atributos que muestran una satisfacción baja e importancia relativamente alta, por lo que la atención y esfuerzos deben ser concentrados en éstos. En el cuadrante III se representan aquellos atributos que no poseen mucha importancia y su satisfacción también es baja, por lo que éstos deberían tener una prioridad baja en cuanto a su atención, y una mejora gradual podría ser beneficiosa para obtener una mejor atención del público. Por último, en el cuadrante IV se encuentran aquellos atributos con alta satisfacción pero baja importancia, por tanto, se identifican como atributos sobrevalorados para los cuales se recomienda mantener su estado de funcionamiento.

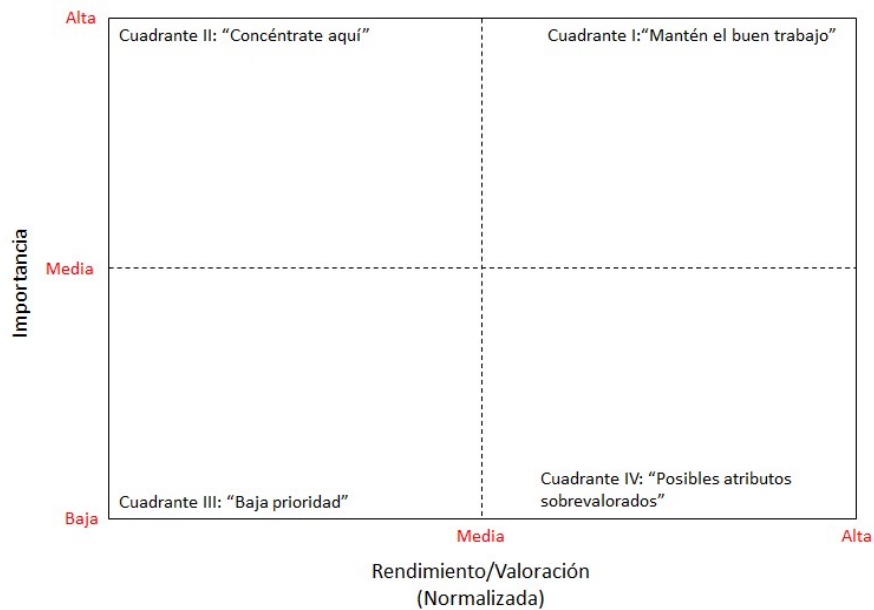


Figura 2. Matriz IPA (Adaptado de Oña, 2014).

Otras técnicas estadísticas simples.

Por último, se han aplicado otras técnicas estadísticas simples o estándar (p.e., medias, varianza, ANOVA, MANOVA, t-test, etc.), cuyos principales objetivos han sido verificar hipótesis, comparación de modelos y/o atributos, establecimiento de relaciones, etc. (Aksoy et al., 2003; Pedersen et al., 2011; Tyrinopoulos and Antoniou, 2008).

Como ejemplo, se puede mencionar el caso de Filipovic et al. (2009) que utiliza las diferentes medias y varianzas de los atributos, así como sus importancias para poder comparar su evolución a lo largo de tres años. Aydin (2017) utiliza esta tipología de análisis simple para caracterizar la muestra (p.e., distribución de los diferentes atributos) y, posteriormente, aplicar técnicas más avanzadas, como la metodología TOPSIS difusa. Birago et al. (2017) calculó las medias y las varianzas, así como la distribución de cada categoría de cada atributo, en base a que junto a una segmentación tradicional basada en la frecuencia de uso del servicio, realizó un estudio comparativo y descriptivo. Además, utilizó la técnica estadística ANOVA para destacar las diferencias existentes en la percepción de los diferentes atributos entre los tres grupos segmentados.

2.2.1.2. TÉCNICAS AVANZADAS

Se tratan de técnicas estadísticas cuyo cálculo de resultados, así como interpretación de los mismos, es más complejo. No obstante, su origen sigue estando fuera de la Minería de Datos. Entre las técnicas que se pueden destacar en este sub-grupo destacan el Análisis de Regresión, el Análisis Factorial y los SEM.

Análisis de Regresión.

Se trata de una técnica empleada para el cálculo de la importancia relativa de cada variable independiente respecto a una dependiente, así como para comprobar la capacidad explicadora general del conjunto de variables independientes (Fox, 1997). Por lo que se trata de una ecuación que expresa la variable dependiente como combinación de las independientes.

Esta técnica aplicada al ámbito de la calidad del servicio en el transporte público, toma como variables independientes los atributos y/o dimensiones de la calidad del servicio, y como dependiente la percepción general de la misma.

Existen diferentes modelos de análisis de regresión en función de la naturaleza y número de categorías de la variable dependiente, así como del tipo de función utilizada para estimarla. Un resumen de los principales modelos existentes se puede ver en la Tabla 1.

Tabla 1. Clasificación de los modelos de regresión (Adaptada de Lopez, 2014).

Nº de categorías de la variable dependiente	Naturaleza de las categorías	Tipo de función	Modelos
Dicotómico (2 categorías)	Complementarias	Lineal	<i>Linear Generalizado</i>
		Logística	<i>Logit</i>
		Normal Tipificado	<i>Probit</i>
Respuesta múltiple (Más de 2 categorías)	No ordenadas	Logística	<i>Logit Multinomial</i>
		Normal Tipificado	<i>Probit Multinomial</i>
	Ordenadas	Logística	<i>Logit Ordenado</i>
		Normal Tipificado	<i>Probit Ordenado</i>

Diversos trabajos en el campo del transporte público utilizan estos modelos. Por ejemplo, Ubillos and Sainz (2004) utilizaron un modelo Logit jerárquico para averiguar la elasticidad tanto del tiempo de viaje como el coste, al igual que para encontrar los efectos potenciales de las diferentes políticas de promoción del uso del transporte público en su cometido de atraer nuevos pasajeros desincentivando el vehículo privado.

Por su parte, Jiang and Zhang (2016) utilizaron un modelo Probit para investigar la calidad del servicio en cuatro principales aerolíneas del mercado doméstico Chino, así como para explorar las relaciones entre la calidad del servicio y la satisfacción de clientes y las condiciones en las que las aerolíneas pueden conservar sus cantidades de pasajeros.

En Eboli and Mazzulla (2008) usaron un modelo logit multinomial para identificar la importancia que tenían los atributos de calidad de un servicio de transporte público en autobús sobre la satisfacción global del pasajero, y calcularon un índice de calidad del servicio.

En relación a los modelos probit multinomiales se observa en la literatura que su aplicación es escasa en el campo del transporte público, destacando el trabajo de Horowitz (1991) donde se realiza una reflexión sobre los ventajas y desventajas de utilizar un modelo probit multinomial en modelos de demanda.

Rojo et al., (2008; 2011) utilizaron un modelo Logit ordenado y un modelo Probit ordenado para analizar la satisfacción de los pasajeros en servicios de bus interurbano. Y en Tyrinopoulos and Antoniou (2008) se estudió, a través de un modelo Logit Ordenado, como las percepciones de los pasajeros sobre la calidad variaban a través de diferentes operadores.

Por su parte, Alonso et al. (2018) analizó la calidad del servicio de taxis percibida por los pasajeros utilizando dos tipos de modelos Probit Ordenados. El primero, mostraba los atributos y sus influencias en la calidad percibida cuando no se suministraba información previa sobre la disponibilidad del servicio. Y el segundo, mostraba lo mismo, pero considerando el cambio de percepción de la calidad después de haber sido informados sobre la disponibilidad del servicio.

Otro ejemplo destacable es el trabajo de Barabino et al. (2012), donde se utilizó un Análisis de Regresión para proveer una herramienta útil de evaluación de la calidad, utilizable por los operadores del transporte que estuvieran dispuestos a certificar el servicio ofrecido.

Análisis Factorial.

Se trata de un conjunto de técnicas estadísticas multivariadas cuyo principal objetivo es investigar si un número de variables de interés (atributos) se encuentran linealmente relacionadas con un pequeño número de factores latentes o no observables (dimensiones) (Spearman, 1904).

Dentro de estas técnicas, las más populares son: el Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC).

El AFE es normalmente usado para descubrir un pequeño número de factores latentes (dimensiones o constructos) a partir de un gran número de variables observadas (atributos), mientras que el AFC es usado para confirmar una estructura factorial predeterminada basada en la teoría o en una investigación previa (Cudeck, 2000).

Por lo que se observa que juntos conforman una buena combinación para poder reducir un número de atributos en dimensiones, confirmar su validez y hacer más manejable los subsiguientes análisis. Esta técnica ha sido usada como paso previo en gran cantidad de trabajos en la literatura (p.e., Aksoy et al., 2003; Beirão and Cabral, 2008; Eboli and Mazzulla, 2012; Weinstein, 2000; etc.).

Modelos de Ecuaciones Estructurales (SEM).

Las técnicas tradicionales mencionadas hasta ahora pueden ser de gran ayuda para caracterizar una muestra, reducir o agrupar su número de atributos en un número menor de dimensiones, o para realizar comparaciones simples entre diferentes tipos de pasajeros, lugares, antes y después de llevar a cabo una mejora, etc. Sin embargo, sus capacidades son limitadas en cuanto a modelización de la calidad del servicio se refiere.

Si se pone el foco de atención en el estudio de la calidad del servicio en el transporte público, se puede observar que la popularidad de los SEM ha crecido de forma exponencial en los últimos años (p.e., Farooq et al., 2018; Farzana et al., 2016; Hapsari et al., 2017; Yilmaz and Ari, 2017, Wan et al., 2016; Yilmaz and Ari, 2017 etc.), siendo especialmente importante en los casos en que se han utilizado datos que provienen de encuestas de actitudes (p.e., Fu et al., 2018; Irtema et al., 2018; Li et al., 2018; Mugion et al., 2018; etc.). Y se puede destacar que su principal causa es debido a su capacidad para modelizar un fenómeno completo de una forma fácil y tratando con un gran número de variables (tanto observadas como latentes) (Golob, 2003).

La generalización de dicha técnica fue llevada a cabo por Jöreskog (1973) y Wiley (1973). Se trata de una técnica estadística multivariada, que combina el análisis de varianza y/o covarianza, regresiones múltiples, análisis factorial, análisis de caminos (Path analysis), modelos econométricos de ecuaciones simultáneas y no recursivos, y modelización multinivel (Bowen and Guo, 2011; Ullman and Bentler, 2003). Por lo que, aunque hay numerosas definiciones del SEM que se han ido formulando a lo largo del tiempo, entre ellas se destaca la de Byrne (1998): “Metodología estadística que utiliza un enfoque confirmatorio del análisis multivalente aplicado a una teoría estructural relacionada con un fenómeno determinado”.

Entre sus principales funciones, dos representan las más utilizadas en el ámbito de la modelización de la calidad del servicio, especialmente, en el campo del transporte público:

- Testear hipótesis sobre cómo un conjunto de variables observadas (atributos), explican un conjunto de factores latentes (dimensiones).
- Testear la fuerza y dirección de las hipotéticas relaciones existentes entre los diferentes factores latentes que componen el modelo.

En consecuencia, si se toma como ejemplo básico de un SEM la Figura 3, se pueden diferenciar dos partes principales en los SEM: por un lado, las relaciones entre los factores latentes (Modelo Estructural). Y por el otro lado, las relaciones entre los diferentes factores latentes con las variables observadas (Modelo de Medición).

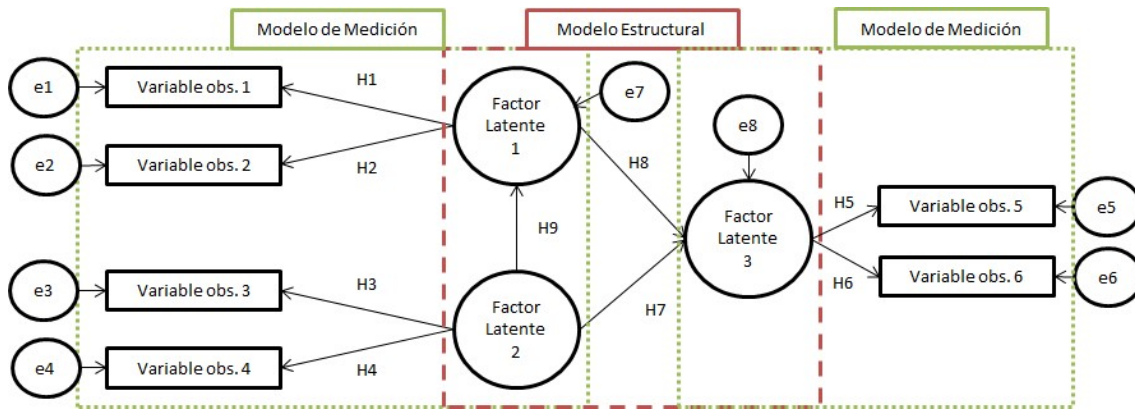


Figura 3. Estructural general de un SEM.

Como caso especial de esta técnica podría ser destacado el Análisis de Caminos (Path Analysis), que solamente posee modelo de medición (sin modelo estructural), es decir, que trabaja únicamente con variables que han sido observadas directamente sin error. Su aplicación en el estudio de la calidad del servicio en el transporte público es igualmente amplia (p.e., Forgas et al., 2010; Jen and Hu; 2003; Joewono and Kubota, 2007; Lin et al., 2008; etc.)

A modo de resumen, se puede destacar que la aplicación del SEM tiene por finalidad estimar los parámetros desconocidos de un modelo establecido a partir de conocimiento previo y/o experto, para después contrastarlo estadísticamente. Esta consideración destaca su naturaleza confirmatoria, es decir, que se trata de una técnica, que basándose en un conocimiento previo y/o experto sobre el fenómeno que se está modelizando, valida las hipótesis presentadas y su posible generalización. Dicho punto se puede considerar su mayor desventaja o limitación.

2.2.2. MÉTODOS DE ANÁLISIS CON ORIGEN EN LA MINERÍA DE DATOS

Se tratan de aquellas técnicas estadísticas que provienen de la rama de la Minería de Datos, siendo definidas como herramientas que, aplicando algoritmos a grandes volúmenes de datos (Maimon and Rokach, 2010), son capaces de extraer información y/o patrones de los mismos. Es decir, se tratan de técnicas de análisis que tienen como objetivo extraer conclusiones y conocimiento implícito, manifestado de diferentes formas (p.e., relaciones, patrones, asociaciones, reglas inferidas, etc.), para desarrollar de forma automatizada modelos previamente desconocidos a partir de una base de datos inicial (Piatetsky-Shapiro and Frawley, 1991).

Según Hernandez et al. (2004) se pueden distinguir dos tipos de modelos:

- Predictivos. Su objetivo es estimar mediante diferentes variables, los valores desconocidos de una variable de interés.

- Descriptivos. Su fin es discernir propiedades de los datos examinados mediante identificación de patrones en los mismos.

La gran mayoría de estas técnicas trabajan segmentando la base de datos en dos submuestras: una de las cuales es utilizada para desarrollar el modelo y aprender los parámetros (base de Entrenamiento o Aprendizaje), y otra para chequear la validez, la ratio de error y el poder explicativo de dicho modelo (base de Test) (Lopez, 2014).

Si a continuación se analizan las técnicas que hay disponibles en la Minería de Datos, se pueden distinguir dos grandes grupos de acuerdo al criterio propuesto por Weiss and Indurkaya (1998), que las agrupa según el proceso utilizado para la extracción de conocimiento:

- Supervisadas. En este tipo de técnicas, los datos de aprendizaje llevan información proporcionada por un experto y un atributo común, denominado clase, que especifica si un caso pertenece o no a un cierto grupo. Por lo que se puede establecer que el objetivo de dichas técnicas será el aprendizaje de este último hecho.
- No supervisadas. En ellas no existe una agrupación de datos según un atributo clase, sino que se realizan descripciones, hipótesis o teorías a partir de un conjunto de datos mediante un aprendizaje por observación.

Diferentes técnicas de minería de datos están empezando a ser usadas en el estudio y análisis de la calidad del transporte público, aunque aún son escasos los trabajos en la literatura. No obstante, se pueden encontrar ejemplos interesantes y con resultados relevantes, como de Oña et al., 2016c; Islam et al., 2016; Machado-Leon et al., 2017; Shen and Li, 2014, etc.

Una clasificación esquemática de las principales técnicas aplicadas en el estudio de la calidad del servicio en el transporte público puede ser encontrada en la Figura 4.

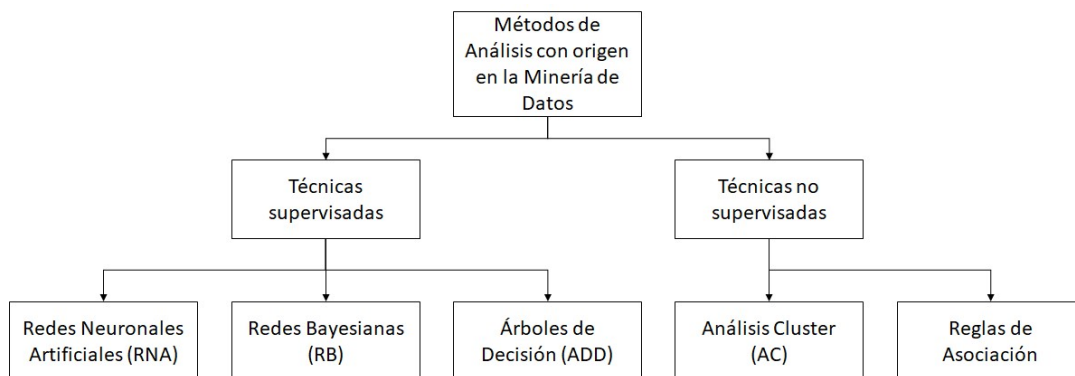


Figura 4. Resumen de los métodos de análisis con origen en la Minería de Datos aplicados a estudios de calidad del servicio del transporte público (Adaptada de Rokach and Maimon, 2008).

A continuación, se explican sus principales características de cada una de ellas, así como se destacan los principales trabajos encontrados en la literatura.

2.2.2.1. TÉCNICAS SUPERVISADAS

De acuerdo a la definición de este tipo de técnicas, se encuentran en este grupo aquellas que obtienen a partir de los datos iniciales un modelo que relaciona el valor de un atributo y los valores de otros atributos. De esta forma, el modelo obtenido se podrá utilizar, posteriormente, para realizar predicciones con datos nuevos. Se pueden destacar como técnicas de este tipo las Redes Neuronales Artificiales (RNA), las Redes Bayesianas (RB) y los Árboles de Decisión (ADD).

Redes Neuronales Artificiales (RNA).

Se trata de un sistema de procesamiento de información basado en el comportamiento del cerebro humano, siendo su principal ventaja su alta capacidad de clasificación y predicción (Martin del Bio and Sanz, 2006).

Su estructura interna está compuesta por unidades de procesamiento de información elemental, llamadas neuronas, organizadas en diferentes capas e interconectadas mediante pesos, que representan la intensidad entre cada par de neuronas (Martin del Bio and Sanz, 2006; Garrido et al., 2014).

Si se interpretan como un grafo dirigido (Figura 5), se pueden identificar las neuronas con los nodos. Éstos llevan asociados una función matemática, cuya principal función es transformar la información de entrada en otra de salida para que pueda ser utilizada por la siguiente neurona.

Las relaciones y sus parámetros, es decir, los pesos, son obtenidos en la fase de entrenamiento, que se trata de una fase clave, ya que esta técnica permite capturar automáticamente relaciones altamente complejas y no lineales (Mahammadipour and Alavi, 2009).

Por tanto, la entrada de la función de cada neurona es la suma de todas las señales de entrada multiplicada por el peso asociado a la conexión de entrada de la señal. Así mismo, la función de transferencia es la relación entre la señal de salida y la de entrada. Esta forma de operar permite afirmar que la información “fluye” desde las neuronas de entrada a las de salida, sin existencia de retroalimentación (Garrido et al., 2014).

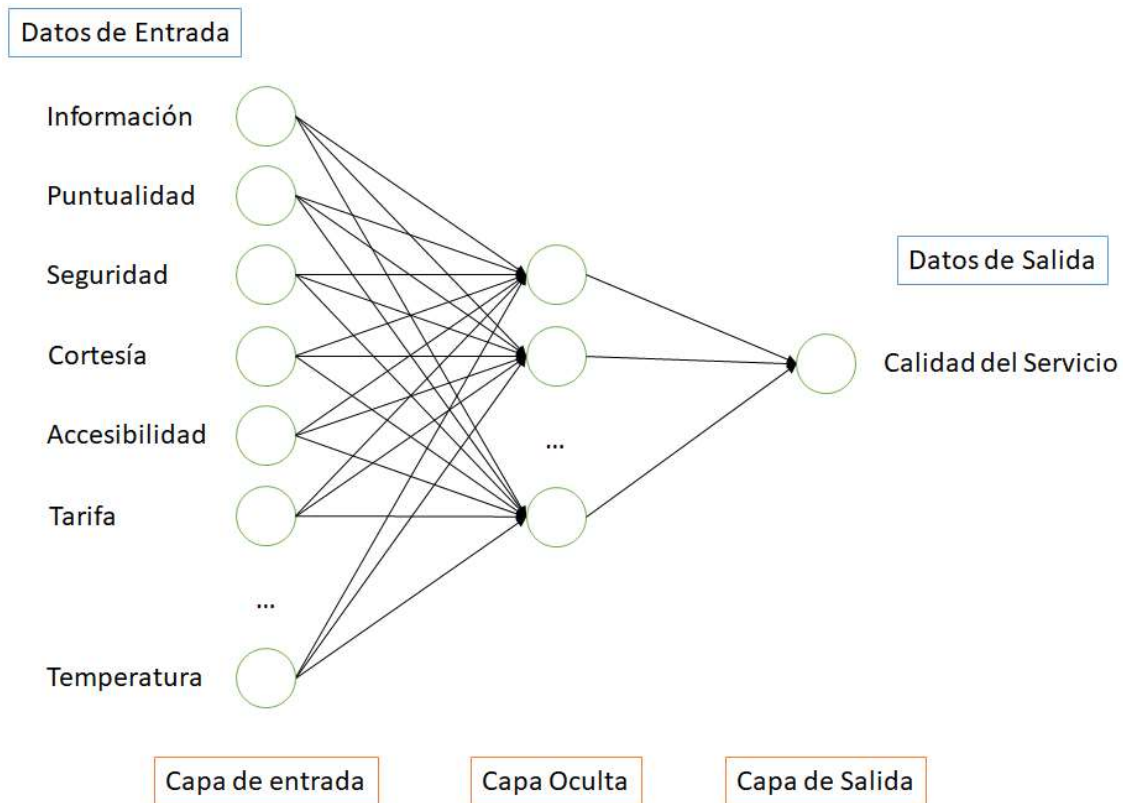


Figura 5. Ejemplo de estructura de una RNA para analizar la calidad del servicio en el transporte público (Adaptada de Garrido et al., 2014).

A pesar de su gran poder clasificatorio y predictivo, las RNA presentan diversas desventajas. La primera de ellas es el llamado efecto “caja negra”, es decir, que proporcionados los datos de entrada, es altamente difícil y casi imposible conocer de forma explícita las relaciones establecidas entre las variables explicativas y las dependientes. Por tanto, no podría ser utilizada en conjunto con otras técnicas que necesiten el conocimiento de dichas relaciones para poder confirmar el modelo que se ha aprendido (Cortez and Embrechts, 2013).

Otro inconveniente, son los pesos, que impiden la inteligibilidad de las asignaciones a cada clase que se realizan. Estos pesos son ocultos y no pueden ser modificados por el investigador (Tullis and Jensen, 2003).

Como último inconveniente, y que se encuentra relacionado con la desventaja anterior, es que el modelo aprendido es difícilmente comprensible y requiere gran cantidad de datos para su entrenamiento (Lopez, 2014).

En cuanto a trabajos en la literatura que hayan utilizado las RNA en el ámbito del transporte público, se debe mencionar su escasez (Pekel and Kara, 2017). Si se pone el foco de atención en el estudio de la calidad del servicio en el transporte público con las RNA, destaca el trabajo de Garrido et al., 2014, quienes se postulan como los primeros en utilizar dicha técnica para analizar el servicio metropolitano de transporte público en Granada (España) con resultados muy favorables. Se utilizó una RNA para analizar dicha calidad del servicio. Posteriormente, mediante tres métodos distintos (Connection Weights, Profile y Perturb) se estableció la importancia relativa de los atributos; y por último, mediante un análisis estadístico sobre los resultados de dichos métodos, se identificaron los grupos de atributos con diferencias significativas en sus importancias relativas.

Behara et al (2002) utilizó las RNA para analizar la calidad del servicio de una red de concesionarios automotrices en Holanda, basándose para ello en un diagrama sistemático SERVQUAL de calidad del servicio. Sus resultados fueron bastante relevantes de cara a políticas de mejora de la calidad en el sistema estudiado, manifestando que estos eran aún mejores cuando se combinaban con Modelos de Expectativas – Percepciones. Además, en relación a su RNA, consiguieron que fuera capaz de predecir con un 75% de precisión la calidad global del servicio utilizando datos de test.

Sa (2010) usó las RNA para evaluar el grado de satisfacción de los pasajeros en el transporte por ferrocarril en China, y cuyos resultados muestran el potencial de esta técnica para entender mejor la demanda de los pasajeros y poder mejorar la calidad del servicio.

Leong et al. (2015) usó SERVPERF con las técnicas SEM y RNA para examinar la calidad del servicio de una aerolínea, así como la satisfacción de los pasajeros y la lealtad de los mismos. Los principales resultados obtenidos mostraron influencias significativas de las dimensiones SERVPERF sobre la satisfacción y sobre la lealtad/intención de recompra. Su principal contribución sería la de ayudar a la localización de los recursos y, al mismo tiempo, a la mejora de la satisfacción y de la lealtad/intención de recompra.

Posteriormente, Islam et al. (2016) realizó un estudio comparativo sobre la capacidad de predicción de tres RNA de la calidad del servicio de autobuses en la ciudad de Dhaka (Bangladesh), encontrando que las dimensiones de puntualidad, fiabilidad, frecuencia, disponibilidad de asientos y experiencia del viaje, eran las más influyentes y significativas en relación con la calidad del servicio.

Árboles de Decisión (ADD).

Los ADD es una técnica para la clasificación y predicción de una variable clase, es decir, representa el conocimiento mediante un modelo que ha sido obtenido mediante el proceso de clasificación, y describe y distingue las clases de los datos. De esta forma, también puede ser usado para predecir clases de objetos (Han and Kamber, 2006).

Una definición más gráfica (Figura 6) y matemática de dicha técnica sería la de una representación de conocimiento en forma de árbol, que está formado por nodos y arcos dirigidos (ramas), y que representan un estado de la variable utilizada como divisora.

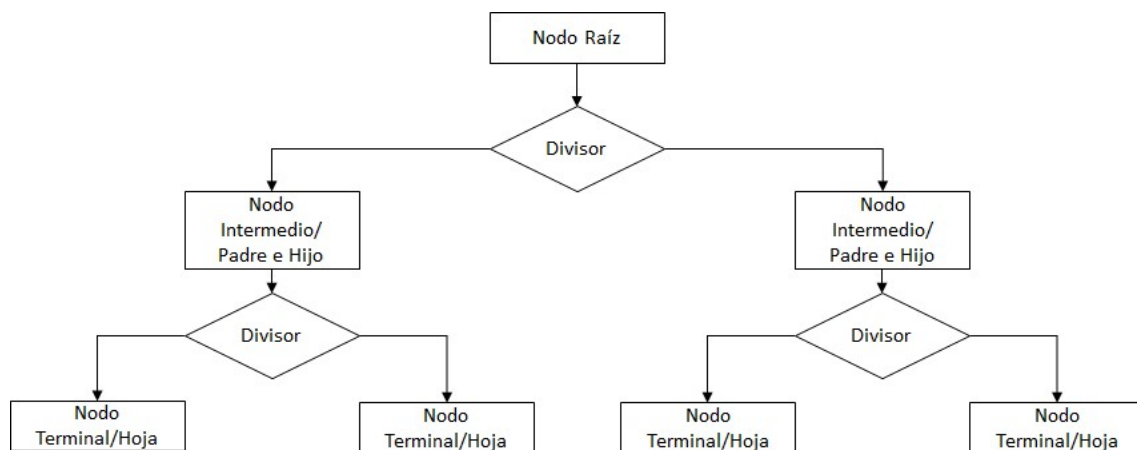


Figura 6. Estructura de un ADD (Adaptado de Kashany and Mohaymany, 2011).

Dicho lo cual, se pueden encontrar diferentes nodos en función de su posición en el árbol:

- **Nodo Raíz.** Localizado en lo más alto del árbol, solo tiene ramas salientes y concentra o representa todos los datos de estudio.
- **Nodos Intermedios.** Se tratan de nodos que involucran el testeo de un atributo específico para determinar cuál es el próximo nodo. Los nodos que poseen ramas salientes se denominan también nodos padres y los nodos que poseen ramas entrantes, se les conocen como nodos hijos. De acuerdo a esta definición, un mismo nodo puede ser al mismo tiempo nodo hijo y padre.
- **Nodos Terminales (Nodos Hoja).** Localizados al final del árbol y no poseen hijos. Se les asigna una clasificación o predicción a todas las instancias que los alcanzan.

Las nuevas instancias son clasificadas a través de una serie de preguntas sobre los valores de sus atributos. Para ello, el proceso se comienza desde lo alto del árbol hasta la hoja, testeando en cada nodo intermedio los valores de los atributos de dicha instancia, y cuando se alcanza la hoja, se clasifica la instancia o se predice un valor de acuerdo a la clase o valor asignado a la hoja (Witten and Frank, 2005).

Al igual que ocurre con las RNA, los trabajos en la literatura que utilizan los ADDs para estudiar la calidad del servicio en el transporte público son bastante recientes (p.e., de Oña et al., 2012; 2016c; 2017; Hernandez et al., 2016; Machado-Leon et al., 2017; Nesheli et al., 2016; Tsami and Nathanail, 2017; Wong and Chung, 2007; etc.).

Como ejemplos, se puede comenzar por Wong and Chung (2007), quienes llevaron a cabo un proceso multi-etapa de descubrimiento de conocimiento en bases de datos de aerolíneas Taiwanesas con el objetivo de aumentar y retener más pasajeros. Entre sus múltiples etapas, aplicaron los ADDs sobre datos de pasajeros de vuelos nacionales que previamente habían sido identificados como relevantes, identificando los aspectos más críticos de la calidad para ellos, así como se extrajo información sobre sus comportamientos y características demográficas.

Posteriormente, de Oña et al. (2012) aplicaron los ADDs para identificar los factores claves que afectan a la calidad del servicio metropolitano de autobús de Granada (España). Para ello, compararon mediante dos modelos los atributos identificados como clave, antes y después de que los pasajeros reflexionaran sobre los principales aspectos del servicio. Uno de sus resultados más destacable fue que dicha percepción sobre la calidad del servicio cambiaba una vez que habían realizado esta reflexión sobre las características del servicio.

Nesheli et al. (2016) realizaron un estudio cualitativo y cuantitativo de los pasajeros del transporte público en Auckland (Nueva Zelanda) y en Lyon (Francia) para un mejor entendimiento de las actitudes de los pasajeros, así como para explorar las percepciones de los mismos en tiempo real ante operaciones de mejora en la calidad del servicio. Para ello, utilizaron un Modelo de Regresión Logístico y ADDs en el análisis de dos aspectos principales: 1) Evaluaron los efectos de los retrasos en la percepción de los pasajeros y su decisión para cambiar de ruta o modo de transporte; 2) Evaluaron las decisiones de los pasajeros ante varias operaciones sobre el transporte en tiempo real. Se consiguieron resultados interesantes, como que el 60% de los pasajeros ante diferentes situaciones de retraso u otras operaciones deciden cambiar sus decisiones de viaje (p.e. cambio de ruta, de modo, etc.).

Hernandez et al. (2016) propusieron una metodología para identificar las fortalezas y debilidades del intercambiador de transporte urbano de Moncloa (Madrid, España), y así gestionar recursos más eficientemente. Dicha metodología estaba basada en dos etapas, una que consistía en aplicar ADDs para deducir la importancia de cada atributo, y la otra en un IPA para identificar qué atributos necesitaban ser mejorados con mayor urgencia.

Por su parte, Tsami and Nathanail (2017) utilizan los ADDs para vincular las percepciones y expectativas de los pasajeros del sistema de transportes en Grecia, con la calidad global del servicio, obteniendo información sobre qué atributos de la calidad son los más importantes. De esta forma, se suministraba una guía para operadores y/o planificadores del transporte sobre cómo incrementar la calidad de sus servicios y al mismo tiempo mejorar su eficiencia y rentabilidad.

Redes Bayesianas (RB).

Las RB se trata de una estructura para el razonamiento bajo incertidumbre y es ampliamente utilizada para la representación de conocimiento incierto (Pearl, 1988). Puede ser definida en términos de una componente cualitativa y otra componente cuantitativa (Kennett and Salini, 2011) (Figura 7). La componente cualitativa consiste en un Grafo Dirigido Acíclico (Directed Acyclic Graph, DAG), en el que son representadas las interacciones entre el conjunto de variables consideradas. En dicha estructura gráfica se pueden encontrar dos elementos bien diferenciados: por un lado, los nodos que representan a las variables; y por otro lado, los arcos dirigidos (flechas) que representan las interacciones (dependencias probabilísticas directas) entre los nodos.

La componente cuantitativa consiste en las diferentes tablas de probabilidad condicional representadas en cada nodo para cada una de sus categorías o estados. Es decir, cada nodo contiene la probabilidad de que se obtenga una categoría específica de la variable representada.

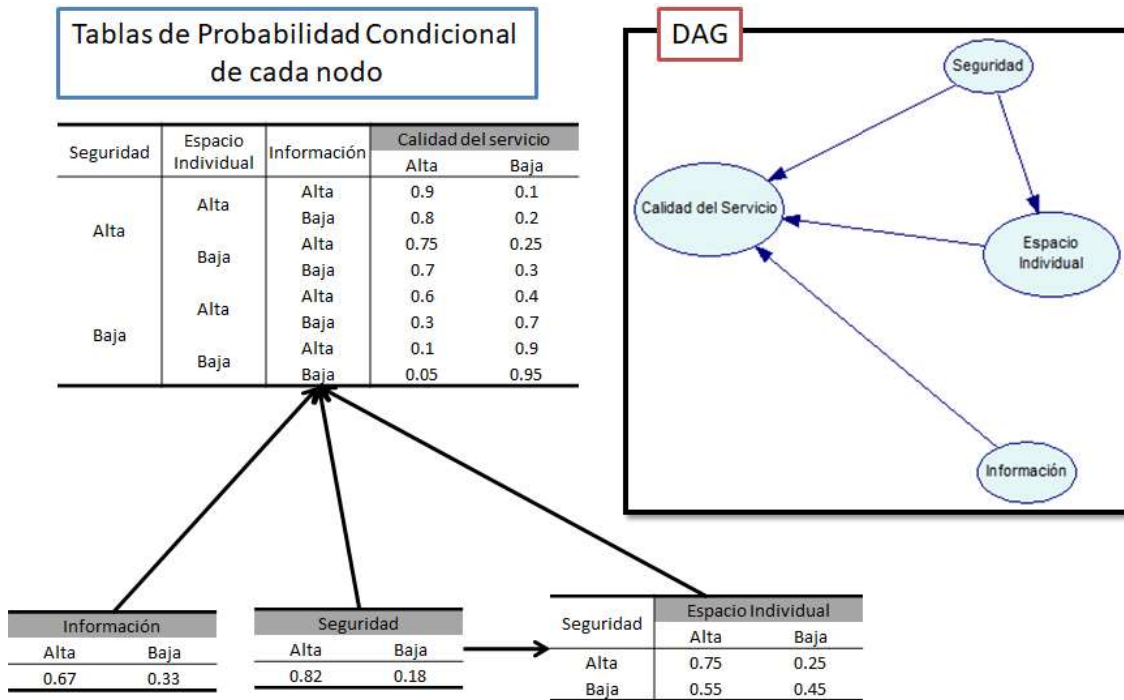


Figura 7. Ejemplo de una RB (DAG y Tablas de probabilidad condicional para cada nodo).

Tal y como se observa en la Figura 7, dos diferentes tipologías de relaciones pueden ser observadas. Por un lado, si los pares de nodos no están conectados, indica independencia (condicional) entre ellos. Por otro lado, si se encuentran conectados, indica dependencia (condicional) entre los mismos. Estas relaciones y nodos son los componentes básicos del DAG, y las configuraciones de dependencia de los mismos poseen una nomenclatura particular: nodos con arcos dirigidos a ellos, son llamados “hijos”, mientras que los nodos desde los cuales se originan dichos arcos, son llamados “padres”.

Las RB necesitan ser aprendidas, tanto su estructura (DAG) como sus parámetros, a través de diferentes metodologías. Este hecho es llamado el problema de aprendizaje de las RB (Kendall and Salini, 2011), que puede ser enunciado informalmente de la siguiente forma:

“Dado un conjunto de datos de entrenamiento e información previa (p.e., conocimiento experto sobre las relaciones causales), se trata de estimar el gráfico topológico (DAG) y los parámetros que conforman las distribuciones de probabilidad conjunta en la RB.”

Para el caso del aprendizaje de la estructura de la red se pueden utilizar las siguientes metodologías:

- Tradicional. La estructura de la RB es suministrada por un experto o basándose en la literatura. Esta metodología es la más fácil de aplicar ya que la mayoría de las aplicaciones son ejemplos donde las relaciones representan causa-efecto. Sin embargo, presenta el inconveniente de que necesita basarse en conocimiento experto previo o en hipótesis cuyo origen está en la literatura. De igual forma, se debe destacar la problemática de que ciertas relaciones, que pueden ser fundamentales para explicar el fenómeno que se está modelando, pueden ser obviadas.
- Aprendizaje automático. La estructura de la RB viene dada por un algoritmo que aprende la estructura a partir de la base de datos suministrada. En dicha metodología no se hace necesario el conocimiento experto previo o hipótesis para el aprendizaje. No obstante, sí que se recomienda la supervisión de la misma por parte del investigador, ya que al aprender libremente puede dar lugar a relaciones que no son acordes a la realidad y, por tanto, el modelo que se obtiene no es fiel a la misma.

Se pueden observar dentro de esta metodología tres tipos de enfoques (de Oña et al., 2011):

- Basado en restricciones (Constraint Based). También conocido como aprendedor de independencias condicionales. En este enfoque la estructura es aprendida a base de realizar test de independencia condicional en los datos y, de esta forma, detectar el Manto de Markov de las variables. Es decir, encontrar el conjunto de nodos de los que depende un nodo específico. Finalmente, la red seleccionada debe ser consistente con las observaciones de dependencia e independencia.
- Basado en puntuaciones (Scored Based). Se tratan de algoritmos de optimización heurísticos de propósito general, que construyen un ranking de posibles estructuras de una RB con respecto a una puntuación de ajuste. En consecuencia, para dicho enfoque se debe definir un criterio de ajuste en el que se basará dicha puntuación.
- Algoritmos híbridos. Esta tipología combina aspectos de los dos casos anteriores, ya que utilizan test de independencia condicional, normalmente usados para reducir el espacio de búsqueda, y puntúan las diferentes redes para encontrar la red óptima en el espacio reducido.

- Mixto. En esta metodología se utiliza una combinación de ambas metodologías anteriormente descritas. Es decir, se introduce en la RB conocimiento experto previo (p.e., imposición sobre ciertas relaciones causales entre las variables) y, posteriormente, se aprende el resto de la estructura de forma automática supervisada a través de algún algoritmo y utilizando cualquier enfoque (Constraint Based, Scored Based o Algoritmos híbridos).

Una vez conocida la estructura de la RB, el objetivo es calcular los valores de los parámetros que conforman la red a través de la maximización de verosimilitud (o logaritmo de la verosimilitud) del conjunto de entrenamiento, es decir, la parte cuantitativa de las RB.

Entre las diferentes ventajas que presenta esta técnica respecto a otras se puede destacar su habilidad para extraer conocimiento de forma automática y sin necesidad de hipótesis previas, su habilidad para integrar en sus modelos conocimiento experto, su capacidad de trabajar con datos perdidos, de evitar sobre-aprendizaje, y de obtener relaciones causales en representaciones gráficas fácilmente interpretables (Heckerman, 1995; Uusitalo, 2007). No obstante, también presenta ciertas desventajas, como es su alto número de parámetros probabilísticos en modelos relativamente simples o la necesidad de discretizar las variables utilizadas (Chen and Pollino, 2012; Uusitalo, 2007).

Esta técnica ha llegado a sufrir un gran crecimiento durante los últimos años, en los diferentes campos donde se hace necesario la representación de conocimiento experto incierto, como por ejemplo, medicina (Bae and Lee, 2018), informática (Fazil and Abulaish, 2018), ingeniería ambiental (Garcia-Prats et al., 2018), etc. Sin embargo, se debe destacar que, como ocurre con el resto de técnicas pertenecientes al ámbito de la Minería de Datos, muestra escasez de aplicación en el estudio de la calidad del servicio en el transporte público (Cugnata et al. 2016; Perucca and Salini, 2014; Wu et al., 2016, Yanik et al., 2017).

Cugnata et al. (2016) utilizaron los datos de una encuesta de satisfacción realizada en el sector de las aerolíneas para desarrollar una metodología que permitiera descubrir una RB que fuera robusta, y calcular la influencia de cada dimensión sobre la variable objetivo (basado en el trabajo de Albrecht et al., 2014), que en este caso era la valoración global de la experiencia del pasajero. Posteriormente, a través de un análisis IPA, pudieron determinar que dimensiones del servicio requerían una mayor urgencia de mejora o cuáles se encontraban en el buen camino.

Destaca el trabajo de Perucca and Salini (2014), quienes llevaron a cabo un análisis de la satisfacción de los clientes con el servicio ferroviario en 14 países europeos, obteniendo una RB que permitía predecir la satisfacción de los pasajeros de manera muy adecuada, siendo relevante para la implantación de mejoras en dichos servicios. Similarmente, en este trabajo se realiza una comparación de las RBs con la regresión logística ordenada en términos de capacidad predictiva, obteniéndose mejores resultados con las RBs.

Wu et al. (2016) empleó las RB para explorar y cuantificar la influencia de diferentes atributos del servicio, relacionados con las dimensiones de seguridad, confort, conveniencia de uso, fiabilidad y tarifa, en la satisfacción general de los pasajeros con la calidad del servicio de autobuses en Nanjin (China). De forma similar, se llevó a cabo un análisis de sensibilidad basado en evidencias, que demostró la utilidad de esta técnica para poder establecer estrategias de mejora y evaluar posibles escenarios futuros, así como se realizó un análisis de reducción de la entropía para destacar la importancia de los atributos considerados.

Por último, Yanik et al. (2017) investigaron en una red multimodal en Estambul (Turquía), mediante la aplicación de las RBs, las relaciones existentes entre la satisfacción de los pasajeros, la calidad percibida de los diferentes aspectos del servicio, y las características socioeconómicas y de viaje de los pasajeros

2.2.2.2. TÉCNICAS NO SUPERVISADAS

Las técnicas no supervisadas, también son conocidas por el nombre de técnicas de descubrimiento del conocimiento, se utilizan para la detección de patrones ocultos en las bases de datos. Estos patrones representan por sí mismos información útil que puede ser utilizada directamente en la toma de decisiones. Dentro de esta tipología de técnicas, se destaca el AC como la más empleada en el ámbito de la calidad del servicio en el transporte público.

Análisis Cluster (AC).

Kaufman and Rousseeuw (1990) definen el AC como la clasificación de observaciones similares en grupos o cluster², en los que tanto el número como sus parámetros (medias, varianzas y covarianzas) son desconocidos. Para tal fin, se utilizan diferentes técnicas heurísticas que tratan de maximizar la similitud entre las observaciones dentro de un cluster (homogeneidad dentro del mismo cluster), y de conseguir la máxima diferencia entre las observaciones de distintos clusters (heterogeneidad entre clusters) (Hair et al., 2010).

² En la terminología del Análisis Cluster, a los grupos de observaciones en los que se segmentan los datos se les conocen también como “clusters”.

Diferentes tipos de algoritmos pueden ser utilizados para llevar a cabo dichos análisis, pudiendo ser agrupados en dos clases (Chakrapani, 2004). Por un lado, estarían los que utilizan algoritmos de tipo jerárquico (Figura 8), es decir, los que agrupan las observaciones en forma de estructuras de árbol y se seleccionan los clusters mediante una función de distancia. Esta tipología de algoritmos no se recomienda cuando las bases de datos son grandes debido a que incrementa considerablemente el coste computacional y la dificultad de interpretación de los resultados.

Por el otro lado, estarían los algoritmos de tipo no jerárquicos (Figura 9), lo cuales, dividen las observaciones en K clusters, cuyo número y la metodología que determina la asignación, pueden ser seleccionados a priori. Son apropiados para grandes bases de datos.

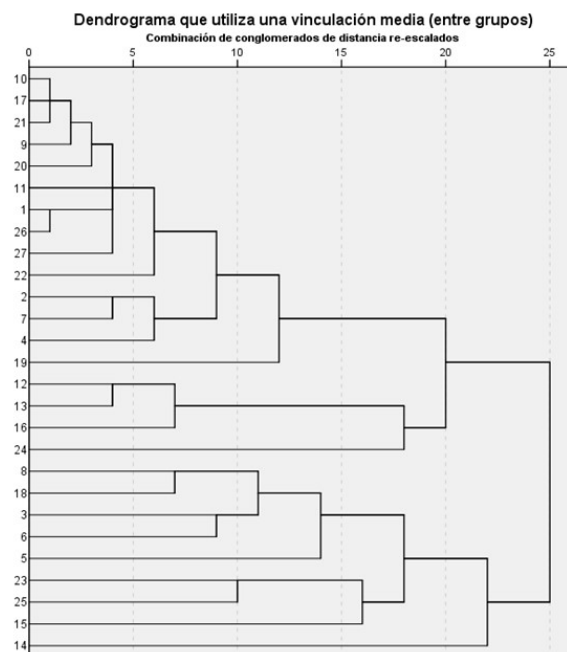


Figura 8. Ejemplo AC con Algoritmo Jerárquico.

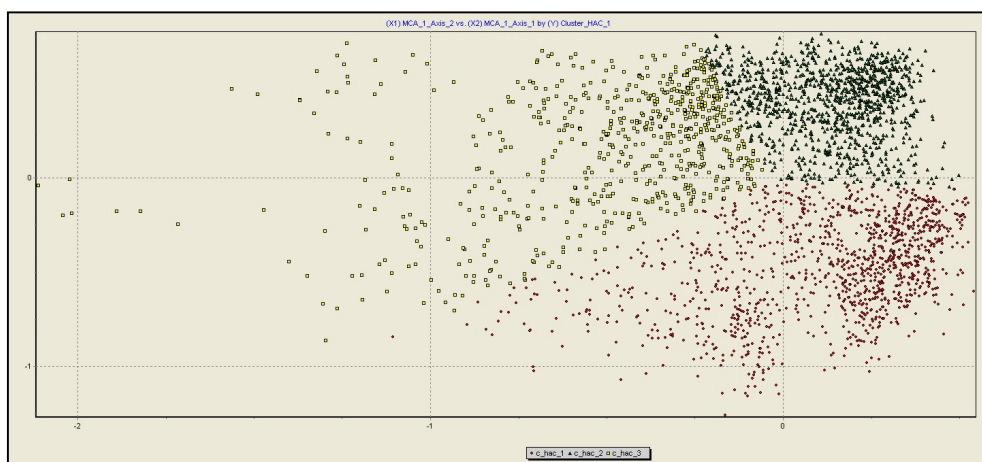


Figura 9. Ejemplo AC con Algoritmo No Jerárquico.

La principal aplicación que podría tener esta técnica para el análisis de la calidad del transporte público, tal y como confirma su propia definición, sería la de solventar de manera eficiente y efectiva el problema de la heterogeneidad de los datos. La principal razón es que se ha demostrado que existen pocas diferencias en los modelos de calidad cuando los datos son segmentados de manera tradicional, siguiendo únicamente una característica socioeconómica (Anable, 2005). Si además se considera la facilidad que presenta esta técnica para combinar un gran número de diferentes tipologías de variables para la agrupación de los datos, ésta se postula como la solución ideal.

A pesar de este hecho, su aplicación en el campo del estudio de la calidad del servicio o las actitudes hacia el transporte público ha sido muy reciente, encontrándose los principales trabajos durante la última década (Abenoza et al., 2017; d'Ovidio et al., 2014; de Oña et al., 2016c; Nesheli et al., 2017; Shiftan et al., 2008; Vicente and Reis, 2016; Wen et al., 2008).

Existe un método particular del AC, llamado LCC, que podría ser categorizado como un caso especial de algoritmo no jerárquico. En dicho análisis, las observaciones se asumen que pertenecen a un conjunto de K clases latentes, con un número y tamaño de los mismos desconocidos a priori. Por su parte, las observaciones pertenecientes a la misma clase son similares con respecto a las variables observadas, en el sentido de que sus distribuciones de categorías se asumen que poseen la misma distribución de probabilidad, y cuyos parámetros a priori son, sin embargo, cantidades desconocidas para ser estimadas (Vermunt and Magidson, 2002).

El LCC tiene múltiples ventajas sobre otras tipologías de AC, como son (Alarcon-del-Amo et al., 2011; de Oña et al., 2013b; Hair et al., 2010; Magidson and Vermunt, 2002; Vermunt and Magidson, 2002; 2005):

- Es posible considerar diferentes tipos de variables (p.e., ordinales, categóricas, métricas o incluso una combinación de ellas) sin la necesidad de realizar a priori una estandarización que podría influir en los resultados.
- LCC permite la clasificación de probabilidades a través del uso de la probabilidad de pertenencia de cada objeto, que ha sido previamente clasificado usando el criterio de máxima verosimilitud. Es decir, LCC está basado en un enfoque clustering probabilístico.
- LCC usa medidas no basadas en la distancia entre objetos, de forma que la estandarización de los datos no tiene efecto en los Clusters finales.
- No se demanda grandes cantidades de memoria de computación para su aplicación, permitiendo la construcción de modelos con grandes cantidades de datos.

- Los modelos, normalmente, pueden incorporar variables independientes, conocidas como covariables o agrupamientos de variables, que son usadas para describir las clases latentes en vez de para definir las.

Sus aplicaciones en otros campos han sido muy satisfactorias y abundantes (p.e., Alarcon-del-Amo et al., 2011), pero en el estudio de la calidad del transporte público son escasas. Por ejemplo, de Oña et al. (2016b) aplicó una LCC para segmentar la muestra de pasajeros del servicio metropolitano de autobuses de Granada (España), obteniendo 4 clusters diferentes, e identificando a posteriori para cada uno de ellos los atributos de la calidad del servicio que más influyen en su satisfacción general.

Otro caso es el de Wen and Lai (2010), quienes segmentaron mediante un LCC datos de pasajeros de diferentes aerolíneas que viajaban de Taipéi a Tokio y de Taipéi a Hong Kong, obteniendo para cada ruta dos clusters (en total 4). Con cada uno de ellos, se analizaron la disposición a pagar más de acuerdo al tipo de mejora en la calidad del servicio realizada (p.e., frecuencia, puntualidad, espacio, etc.) y la elasticidad de la demanda. Los resultados pusieron de manifiesto la necesidad de explorar mediante este tipo de segmentaciones, el comportamiento de los pasajeros de aerolíneas para una mejora en las campañas de marketing y de implantación de mejoras.

Reglas de Asociación.

Al igual que el AC, se trata de un modelo de aprendizaje no supervisado cuya función principal es la extracción de conocimiento directamente de los datos y sin saber a priori el resultado de interés. Dichas reglas han sido tradicionalmente utilizadas en bases de datos transaccionales definidas estas como “bases de datos que tienen como objetivo la recepción y envío de datos a gran velocidad “(Lopez, 2014).

De una manera más formal, una regla de asociación puede definirse como una implicación del tipo “ $X \rightarrow Z$ ”, donde X y Z son ítems y son conocidos como antecedente y consecuente, respectivamente.

El método principal de extracción de reglas de asociación es la aplicación de algoritmos sobre las bases de datos, siendo el algoritmo de Agrawal et al., (1993) el más utilizado. Su funcionamiento está basado en la extracción de conjuntos de ítems que aparecen juntos y con más frecuencia en la base de datos.

Una vez extraída la regla de asociación, se debe evaluar su importancia, para ello se utilizan 3 parámetros:

- Support. Número de veces que aparecen conjuntamente los ítems de dicha regla en la base de datos.
- Confidence. Número de veces en que los ítems del consecuente aparecen en una instancia donde están los ítems del antecedente.

- Lift. Probabilidad de encontrar el consecuente en una búsqueda en aquellos conjuntos de ítems donde se encuentra en antecedente.

Como ventaja se puede destacar su facilidad de interpretación y como desventaja que cuando se aplican modelos de generación de muchas reglas, suelen aparecer un gran número de reglas de asociación no importantes o redundantes. Hasta donde se conoce no se han encontrado estudios que apliquen las reglas de asociación en el estudio de la calidad del transporte público. Su principal aplicación ha sido para análisis de compras, donde se identifican productos que se compran juntos frecuentemente (p.e., Chen et al., 2005; Kaur and Kang, 2016; Rezende and Ladeira, 2019; etc.). Sí que ha sido usado ampliamente en el análisis de accidentes de transportes, como por ejemplo, para diferenciar patrones de accidentes dentro y fuera de carreteras peligrosas (puntos o áreas negras) (Geurts et al., 2005), para analizar la gravedad de accidentes donde hay al menos un automóvil involucrado (Lopez, 2014), o para analizar las características que influyen en los accidentes de peatones (Montella et al., 2012).

2.3. DISCUSIÓN Y CONCLUSIONES

La Tabla 2 resume una revisión de la literatura actual sobre trabajos que versan sobre calidad del servicio en el transporte público.

Tabla 2. Muestra de trabajos que estudian la calidad del servicio en el transporte público.

AUTOR	MODO DE TRANSPORTE	PAÍS (CIUDAD)	OBJETIVOS	TIPOLOGÍA DE ENCUESTA	ESTRUCTURA	PERIODICIDAD	Nº DE ENCUESTAS VÁLIDAS (REALIZADAS)	ESCALA	TECNOLOGÍA DE TOMA DE DATOS	METODOLOGÍA DE ANÁLISIS
Transporte Público Urbano y Metropolitano										
<i>Barabino et al., 2012</i>	Autobús	Italia (Cagliari)	Establecer una herramienta para estudiar la calidad del servicio	Encuesta de Satisfacción	1. SERVQUAL adaptado 2. Características Socioeconómicas	-	2.611 (2.849)	11 puntos Likert	A papel con supervisión	1. Análisis de Regresión 2. Otras técnicas estadísticas simples 3. Correlación Bivariada
<i>Celik et al., 2013</i>	Autobús	Turquía (Estambul)	Establecer una metodología para estudiar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	2.006	5 puntos Likert	A papel	Multicriterio de toma de decisiones (Multi Criteria Decision Making, MCDM) difuso tipo 2
<i>Chica-Olmo et al., 2018</i>	Autobús	España (Granada)	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características del Servicio 3. Características Socioeconómicas	-	1484 (1720)	11 puntos Likert	A papel con supervisión	1. Análisis Factorial 2. Modelo Logit Multinivel
<i>Fu and Juan, 2017</i>	Autobús	China (Keqiao)	Analizar las actitudes de los pasajeros	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Hábitos de viaje 4. Características Socioeconómicas	-	2.265 (3.900)	5 puntos Likert	A papel sin supervisión en casa	1. Análisis Factorial 2. SEM

<i>Garrido et al., 2014</i>	Autobús	Granada (España)	Analizar la influencia de las características de los pasajeros sobre la calidad global y conocer la importancia de los atributos	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características del Servicio	-	858	11 puntos Likert	A papel con supervisión	RNA
<i>Imam, 2014</i>	Autobús	Jordania (Amán)	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	191 (450)	10 puntos Likert	A papel con supervisión	Otras técnicas estadísticas simples
<i>Islam et al., 2016</i>	Autobús	Daca (Bangladés)	Comparación sobre el poder de predicción de la calidad del servicio de tres RNs	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	655	5 puntos Likert	A papel con supervisión	RNA
<i>Wu et al., 2016</i>	Autobús	Nanjing (China)	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	609 (745)	4 puntos Likert	A papel con supervisión	RB
<i>Bordagaray et al., 2012</i>	Bicicleta	España (Santander)	Modelización de la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	195	4 puntos Likert	-	1. Modelo Probit Ordenado
<i>Awasthi et al., 2011</i>	Metro	Canadá (Montreal)	Establecer una metodología para estudiar la calidad del servicio	Encuesta de Satisfacción	SERVQUAL	-	60 (60)	5 puntos Likert	-	1. TOPSIS difuso 2. Otras técnicas estadísticas simples

<i>Birago et al., 2017</i>	Metro	Ghana (Accra)	Analizar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características Socioeconómicas	-	134 (150)	5 puntos Likert	A papel y con/sin supervisión	Otras técnicas estadísticas simples
<i>Celik et al., 2014</i>	Servicios ferroviarios urbanos	Turquía (Estambul)	Establecer una metodología para estudiar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. SERVQUAL 2. Hábitos de viaje 3. Características Socioeconómicas	-	4.966	5 puntos Likert	A papel con supervisión mediante teléfono	VIKOR difuso
<i>Andreassen, 1995</i>	Servicios de Transporte Público	Noruega	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad	-	1.000	7 puntos Likert	Telefónica	SEM
<i>Beirão and Cabral, 2007</i>	Servicios de Transporte Público	Portugal (Oporto)	Estudio cualitativo del comportamiento de pasajeros del transporte público y pasajeros de vehículo privado.	Encuesta de Preferencias Reveladas	1. Modos de transporte y comparación 2. Características Socioeconómicas	-	24	-	Entrevista Semi-Estructurada	-
<i>Beirão and Cabral, 2008</i>	Servicios de Transporte Público	Portugal (Oporto)	Analizar la intención de recompra entre segmentos de pasajeros, en especial, hombres y mujeres	Encuesta de Actitudes	1. Actitudes hacia el Servicio 2. Hábitos de viaje 3. Características Socioeconómicas	-	2.812 (3.009)	11 puntos Likert	Telefónica	1. Análisis Factorial 2. AC

<i>Diana, 2012</i>	Servicios de Transporte Público	Italia	Establecer una metodología para mejorar el estudio de los datos de satisfacción	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características Socioeconómicas	Anual	4.123 (48.253)	4 puntos Likert	-	Otras técnicas estadísticas simples
<i>Fiorio et al., 2013</i>	Servicios de Transporte Público	Varios Países de Europa	Analizar la satisfacción en relación a la existencia de alternativas en el transporte público	Encuesta de Satisfacción	Características del Servicio	Anual	500 en cada ciudad	4 puntos Likert	Telefónica	Modelo Probit
<i>Hernandez et al., 2016</i>	Servicios de Transporte Público	Madrid (España)	Proposición de una metodología para analizar la calidad del transporte	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Hábitos de viaje 4. Características Socioeconómicas	-	865 (4000)	5 puntos Likert	On-line	1. ADD 2. IPA
<i>Nesheli et al., 2016</i>	Servicios de Transporte Público	Nueva Zelanda (Australia) y Lyon (Francia)	Analizar la reacción de los pasajeros ante situaciones de incertidumbre en el transporte	Encuesta de Preferencias Reveladas	1. Características socioeconómicas 2. Uso de las aplicaciones de Smartphone 3. Elección de acciones mediante opciones 4. Valor del tiempo en diferentes escenarios 5. Importancias	-	290 en Nueva Zelanda y 321 en Lyon	5 puntos Likert (Importancias)	A papel y con supervisión	1. Otras técnicas estadísticas simples 2. Modelo de regresión 3. ADD

<i>Shifftan et al., 2008</i>	Servicios de Transporte Público	EE.UU. (Utah)	Identificar pasajeros potenciales y desarrollar estrategias para incrementar el uso del transporte público	Encuesta de Actitudes	1. Actitudes hacia el Servicio 2. Características Socioeconómicas	-	522	11 puntos Likert	A papel y con supervisión	1. Análisis Factorial 2. SEM 3. AC
<i>Tsami and Nathanail (2017)</i>	Servicios de Transporte Público	Grecia	Analizar el GAP entre las expectativas y percepciones de los pasajeros vinculándolos con la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	211	5 puntos Likert	-	ADD
Aerolíneas										
<i>Abdlla et al., 2007</i>	Avión	Egipto	a) Desarrollar un instrumento para gestionar la calidad del servicio, llamada TNE-Matriz. b) Medir la calidad del servicio.	Encuesta de Satisfacción	SERVQUAL	-	474 (600)	9 puntos Likert	A papel y con supervisión	Otras técnicas estadísticas simples
<i>Aksoy et al., 2003</i>	Avión	Turquía	Identificar diferencias entre pasajeros de diversas nacionalidades	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características Socioeconómicas	-	1014 (1350)	7 puntos Likert	A papel	1. Otras técnicas estadísticas simples 2. Análisis Factorial

<i>Chang and Yeh, 2002</i>	Avión	Taiwán (Taipei) e Inglaterra (Londres)	Analizar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje	-	363 (390)	11 puntos Likert	A papel con supervisión	Otras técnicas estadísticas simples con variantes difusas
<i>Chen, 2008</i>	Avión	Taiwán	Analizar la relación entre calidad, valor percibidos, satisfacción y lealtad	Encuesta de Actitudes	1. SERVQUAL adaptado 2. Actitudes hacia el Servicio 3. Características Socioeconómicas	-	245 (300)	5 y 7 puntos Likert	A papel	1. Análisis Factorial 2. SEM
<i>Chen and Chang, 2005</i>	Avión	Taiwán	Analizar la diferencia entre las percepciones y expectativas de los pasajeros y los empleados	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	470 (623) pasajeros 322 empleados	5 puntos Likert	A papel con supervisión	1. Otras técnicas estadísticas simples 2. IPA
<i>Chiou and Chen, 2012</i>	Avión	China	Analizar la calidad del servicio y examinar la relación entre dos etapas de servicio consecutivas desde una perspectiva de cadena	Encuesta de Actitudes	1. SERVQUAL adaptado 2. Actitudes hacia el Servicio 3. Hábitos de viaje 4. Características Socioeconómicas	-	968	5 puntos Likert	A papel con retorno por correo (para los que viajaban por primera vez)	1. Análisis Factorial 2. SEM
<i>Leong et al., 2015</i>	Avión	Kuala Lumpur (Malasia)	Analizar el impacto de las dimensiones del SERVPERF en la satisfacción hacia la lealtad	Encuesta de Actitudes	1. SERVPERF 2. Características Socioeconómicas	-	300 (350)	7 puntos Likert	A papel sin supervisión en casa	1. Otras técnicas estadísticas simples 2. Análisis Factorial 3. SEM 4. RNA

Transporte Público Interurbano										
<i>Agarwal, 2008</i>	Tren	India	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad	-	500	5 puntos Likert	A papel	1. Análisis Factorial 2. Análisis de Regresión
<i>Aydin, 2017</i>	Tren	Turquía (Estambul)	Establecer una metodología para estudiar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características Socioeconómicas	Anual	17.769	5 puntos Likert	A papel	1. Otras técnicas estadísticas simples 2. TOPSIS difuso
<i>Brons et al., 2009</i>	Tren	Holanda	Evaluar la importancia del aspecto "Accesibilidad" y la disposición a usar el tren	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	Regularmente (No especifican)	1394	10 puntos Likert	-	1. Análisis de Regresión 2. Análisis Factorial
<i>Cavana et al., 2007</i>	Tren	Nueva Zelanda (Wellington)	Establecer una metodología para estudiar la calidad del servicio	Encuesta de Satisfacción	1. SERVQUAL adaptado 2. Hábitos de viaje 3. Características Socioeconómicas	-	340 (800)	9 puntos Likert	A papel, con retorno en mano o por correo.	1. Análisis de Regresión 2. Análisis Factorial
<i>Chou and Kim, 2009</i>	Tren	Taiwán	Analizar el comportamiento de los pasajeros	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Características Socioeconómicas	-	832	10 puntos Likert	-	1. Análisis Factorial 2. SEM
<i>Chowdhury et al., 2018</i>	Tren	Nueva Zelanda (Auckland)	Analizar la calidad del servicio y comparar a los pasajeros frente a los operadores	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	363 pasajeros 7 operadores	9 puntos Likert	A papel con supervisión y entrevista semi-estructurada	1. Otras técnicas estadísticas simples 2. AC

<i>Isikli et al., 2017</i>	Tren	Turquía (Estambul)	Analizar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Hábitos de viaje 3. Características Socioeconómicas	Anual	11.116 (4.966 en 2012 y 6.150 en 2013)	6 puntos Likert	A papel con supervisión	Otras técnicas estadísticas simples
<i>Kuo and Tang (2013)</i>	Tren	Taiwán	Analizar la relación entre calidad, imagen corporativa, satisfacción y lealtad entre la gente mayor	Encuesta de Actitudes	1. SERVQUAL adaptado 2. Actitudes hacia el Servicio 3. Características del Servicio 4. Características Socioeconómicas	-	341 (400)	5 puntos Likert	A papel	1. Análisis Factorial 2. SEM
<i>Nandan (2010)</i>	Tren	India	Analizar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad	-	700	5 puntos Likert	A papel con supervisión	Análisis Factorial
<i>Perucca and Salini, 2014</i>	Tren	Diferentes países europeos	Analizar la satisfacción de los pasajeros	Encuesta de Satisfacción	1. Atributos de Calidad	Anual	> 17.000	11 puntos Likert	-	RB
<i>Shen et al., 2016</i>	Tren	China (Suzhou)	Analizar la satisfacción de los pasajeros	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Hábitos de viaje 4. Características Socioeconómicas	-	813 (840)	10 puntos Likert	A papel sin/con supervisión	1. SEM 2. IPA

<i>Abenoza et al. (2017)</i>	Servicio de Transporte Público	Suecia	a) Identificar y caracterizar los actuales y potenciales pasajeros b) Determinar los atributos más importantes de la satisfacción por segmento c) Análisis, a lo largo del tiempo y en el espacio, de la importancia de los atributos	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Características Socioeconómicas	Anual	453.564 (563.855)	5 puntos Likert	Telefónica	1. AC 2. Otras técnicas estadísticas simples 3. Modelo Logit Ordenado
<i>Lin et al., 2008</i>	Servicio de Transporte Público	Taiwán	Establecer una metodología para estudiar el comportamiento de los pasajeros	Encuesta de Actitudes	1. Atributos de Calidad 2. Actitudes hacia el Servicio 3. Características Socioeconómicas	-	385 (800)	5 puntos Likert	A papel y con devolución	1. Análisis Factorial 2. Path Análisis
Otros										
<i>Alonso et al. (2018)</i>	Taxi	España (Santander)	Establecer una metodología para estudiar la calidad del servicio	Encuesta de Satisfacción	1. Atributos de Calidad 2. Características Socioeconómicas	-	215	5 puntos Likert	A papel	1. Modelo Probit Ordenado

Nota: Se trata de una muestra, no de la totalidad de referencias existentes en la literatura.

La primera conclusión que se puede extraer es que cada estudio es muy diferente al resto, aunque tengan objetivos comunes o similares. Así mismo, se destaca que las singularidades juegan un rol muy importante y son un factor común a tener en cuenta previo a cualquier estudio, tal y como se demuestra a lo largo de todos los trabajos recogidos de la literatura.

Cuestiones sobre si es conveniente considerar las expectativas o simplemente las percepciones, si se quiere analizar de forma global el paradigma Calidad del Servicio-Satisfacción-Lealtad/Intención de Recompra, cuál es el número de atributos a evaluar o incluso qué escala de medida utilizar, son aspectos que se deben de decidir y plasmar en la configuración de las diferentes encuestas empleadas. No obstante, se debe destacar que sí que existen ciertos rasgos comunes en cuanto a atributos seleccionados y escalas de medida. La periodicidad es nula en casi todos los trabajos en la literatura, ya que muchas veces se trata de encuestas realizadas por los propios investigadores para dicha ocasión.

El hecho de que sea una tipología de encuesta u otra, así como el objetivo marcado, se puede apreciar que condiciona la tecnología de recogida de los datos y las posteriores metodologías utilizadas para su análisis.

No obstante, tras la revisión del estado del arte, se observa que las diferentes metodologías recogidas en la literatura, ya tengan su origen en la Minería de Datos o no, no son suficientes por sí mismas para llevar a cabo un análisis completo de la calidad del servicio, especialmente, en el ámbito del transporte público. Hecho que puede contrastarse con la información recogida en la Tabla 2, donde lo más común es utilizar una combinación de varias de esas técnicas y/o metodologías. Aunque bien es cierto, existen ciertas técnicas que presentan más ventajas que otras para solucionar ciertas singularidades o problemas, caso del AC frente a la segmentación tradicional, tal y como se ha comprobado analizando sus diferentes propiedades.

De igual modo, explorando las características de las principales técnicas estadísticas aplicadas, los SEM se erigen como la mejor técnica para validar relaciones y modelos (en definitiva, conocimiento) de una forma robusta, efectiva y eficiente. En consecuencia, se puede afirmar que es idónea la aplicación de dicha técnica en el análisis de la calidad del servicio en el transporte público, ya que, en cada caso de estudio, se presentan diversas características propias y singulares que deben ser validadas a partir de unos datos de estudio.

Sin embargo, presenta la principal limitación de que el conocimiento debe ser hipotetizado previamente en forma de modelo para que el SEM pueda validarlo. Dicho lo cual, ciertas características y relaciones entre las dimensiones del estudio podrían no ser descubiertas y obviadas, perdiendo identidad el modelo y no ajustándose a la realidad analizada.

Por tanto, en cada estudio de la calidad del servicio se hace necesario recurrir a técnicas estadísticas que sean capaces de extraer automáticamente dicho conocimiento sobre cómo se relacionan las dimensiones que explican la calidad, de forma que complemente al SEM y solvante esta situación de omisión de relaciones, que no pueden ser establecidas mediante otras vías, y conforman características fundamentales.

Para tal caso, dentro del amplio abanico de técnicas estadísticas, se presentan las RB capaces de extraer conocimiento directamente de los datos de estudio y sin necesidad de hipótesis previas. Además, la información resultante de dicha técnica es presentada en una forma muy fácil de analizar, de forma que, posteriormente, se pueda extraer la información necesaria para ser utilizada por SEM. De este modo queda evidenciada la complementariedad de ambas técnicas.

Se debe destacar que este problema es un hándicap común no solo en estudios de calidad del servicio en el transporte público, sino también en otros campos, como por ejemplo, cuando se analizan las actitudes de los pasajeros hacia un determinado servicio. Cada caso de estudio presenta una serie de características específicas y singularidades, que hacen difícil hipotetizar de forma acertada todas las relaciones existentes entre los constructos de fenómeno estudiado, recurriendo por un lado a la literatura (en la que muchas veces se encuentran opiniones diferentes y/o contrarias respecto al mismo fenómeno) o por otro lado a conocimiento experto previo (a través del cual se pueden obviar ciertas relaciones fundamentales) para establecer estas hipótesis. Por este motivo se demanda establecer un proceso metodológico más completo, en el que a través de una etapa exploratoria y otra etapa confirmatoria, se puedan establecer las hipótesis de las relaciones existentes de un fenómeno dado de la forma más coherente y ajustada posible a la realidad. Este proceso metodológico de dos etapas se propone, implementa y valida en esta tesis doctoral.



CAPÍTULO 3.

OBJETIVOS

Capítulo 3.

OBJETIVOS

En esta tesis doctoral se propone y valida, a través de dos casos de estudio ligados al transporte público, una metodología de dos etapas, basada en la aplicación combinada de una técnica de carácter exploratorio (RB) y otra técnica de carácter confirmatorio (SEM), con el objetivo de identificar y confirmar las relaciones existentes entre diferentes dimensiones/constructos que conforman la calidad del servicio y otros fenómenos bajo estudio.

Mediante esta metodología se modeliza la calidad del servicio en el Metro de Sevilla desde la perspectiva de los pasajeros, y después se desarrolla un modelo sobre actitudes hacia dicho transporte público.

Diferentes estudios han intentado emplear diversos modelos o técnicas estadísticas para analizar la calidad del servicio en el ámbito del transporte público. Sin embargo, éstos presentan ciertas limitaciones con respecto a la extracción de conocimiento y su modelización automática, de forma que, esta metodología aquí propuesta es novedosa en este campo y genera una nueva visión para solventar diferentes hándicaps ligados a estos temas y a otros propios de la naturaleza de esta tipología de estudios.

3.1. OBJETIVO PRINCIPAL

El objetivo principal de esta tesis doctoral es desarrollar, implementar y validar una metodología de dos etapas, basada en el uso combinado de dos técnicas complementarias, como son las RB y los SEM, para descubrir y confirmar las relaciones existentes entre diferentes dimensiones o constructos que conforman un fenómeno bajo estudio.

Dichas técnicas presentan una serie de propiedades específicas, siendo capaces de resolver cada una las propias limitaciones de la otra. Por lo que, en base a este hecho, se pueden destacar como ventajas de esta metodología, la facilidad y eficacia para la extracción de conocimiento directamente desde los datos en forma de modelo simple e intuitivo, y su capacidad para desarrollar y validar, de manera simultánea, el modelo aprendido y las relaciones establecidas entre las dimensiones que lo componen, posibilitando una generalización posterior de sus resultados.

Se puede afirmar que, dada la dificultad actual para obtener automáticamente conocimiento testeado sobre las relaciones que componen un determinado modelo, esta metodología de dos etapas, presenta una gran contribución para la comunidad científica, en tanto que puede ser utilizada para la exploración de nuevos campos de conocimiento, en los cuales no exista conocimiento previo, como para la confirmación y/o mejora de fenómenos de estudio conocidos, verificando los modelos existentes y/o mejorándolos.

3.2. OBJETIVOS ESPECÍFICOS

Este objetivo principal se concreta en los siguientes objetivos específicos:

- Determinar los diferentes factores latentes en los que se agrupan los atributos de calidad que explican las diferentes características del servicio.
- Identificar y validar un modelo de calidad del servicio del metro de Sevilla con relaciones específicas entre los factores latentes, que no hayan sido previamente verificadas.

- Comprobar si el ranking de importancias que se puede establecer sobre las dimensiones de calidad en base a la estructura de la RB se corresponde con las fuerza de la relaciones extraídas a partir del modelo SEM.
- Identificar cuáles son los factores más influentes en la calidad global del servicio percibida por los pasajeros.
- Determinar qué tipología de efecto tiene cada factor sobre la calidad global del servicio: directo, indirecto o ambos.
- Verificar si los parámetros de ajuste del modelo SEM, calibrado a partir de las RB, mejoran los parámetros de ajuste del modelo tradicional de calidad del servicio en el TP (donde todos los factores influyen de forma directa sobre la calidad del servicio).
- Validar la metodología de dos etapas propuesta aplicándola a otro campo de estudio relacionado con el transporte público, en particular, en el estudio de las actitudes hacia el transporte público.
- Verificar si en el modelo de actitudes extraído y validado con la metodología propuesta se identifica el Paradigma de la Calidad del servicio-Satisfacción-Intención de recompra.

3.3. HIPÓTESIS

Esta tesis doctoral cuenta con las diferentes hipótesis que se plantean contrastar con el trabajo realizado en la misma:

- H1. El uso de RB permite identificar relaciones entre factores latentes que no han sido previamente consideradas en la literatura.
- H2. Ciertos factores latentes en los que se agrupan los diferentes atributos de calidad en el metro de Sevilla no sólo ejercen un efecto directo sobre la calidad del servicio, sino también un efecto indirecto.
- H3. El ranking de importancias que establece la RB sobre las dimensiones de calidad se puede considerar equivalente a los efectos totales calculados con el modelo SEM.
- H4. Los parámetros de ajuste del modelo SEM calibrado a partir de las relaciones extraídas con la RB (para el caso de estudio de la calidad del servicio) son mejores que los parámetros de ajuste del modelo SEM de calidad del servicio tradicional, donde todos los factores tienen una influencia exclusivamente directa hacia la calidad global del servicio.

- H5. La metodología de dos etapas propuesta obtiene resultados satisfactorios cuando se aplica en el estudio de las actitudes hacia el transporte público.
- H6. El modelo de actitudes extraído y validado con la metodología propuesta determina las mismas relaciones que las que se establecen en el Paradigma de la Calidad del servicio-Satisfacción-Intención de recompra, ampliamente aceptado en la comunidad científica.



**CAPÍTULO 4.
METODOLOGÍA Y
MATERIALES**

Capítulo 4.

METODOLOGÍA Y MATERIALES

Este capítulo comienza con una descripción de la metodología aplicada y de las diferentes técnicas empleadas. Del mismo modo, se muestran las diferentes fases del trabajo de investigación que son llevadas a cabo en cada uno de los dos casos de estudio desarrollados en esta tesis doctoral, y los datos utilizados para lograr los objetivos marcados.

4.1. METODOLOGÍA

A continuación se exponen las diferentes técnicas utilizadas en la metodología de dos etapas que se desarrolla en esta tesis doctoral, a saber, el Análisis Factorial (sección 4.1.1.), las Redes Bayesianas (RB) (sección 4.1.2.), y los Modelos de Ecuaciones Estructurales (SEM) (sección 4.1.3.).

4.1.1. ANÁLISIS FACTORIAL

El Análisis Factorial se define como una técnica estadística multivariante que tiene como objetivo analizar las correlaciones existentes entre un grupo de variables observadas (Lawley and Maxwell, 1971; Sharma, 1996). El fin último es encontrar el número mínimo de dimensiones latentes o factores³ capaces de explicar la máxima información contenida en las variables observadas (Barbero-García et al., 2011).

Dentro de esta técnica, se pueden distinguir dos tipos de Análisis: el Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC) (Hair et al., 2010). Ambas técnicas persiguen el mismo objetivo de reducción, pero sus principales diferencias residen en el número y la naturaleza de las hipótesis y restricciones que se realizan a priori en el modelo factorial.

4.1.1.1. ANÁLISIS FACTORIAL EXPLORATORIO (AFE)

El AFE se trata de una técnica estadística exploratoria cuyo objetivo fundamental es evaluar la dimensionalidad de un conjunto de variables observadas mediante la identificación del menor número de factores necesarios para explicar la correlación entre ellos (Brown, 2014). Este tipo de análisis no requiere hipótesis y restricciones previas sobre las variables observadas.

La aplicación del AFE se puede resumir en 5 pasos fundamentales, basados en Williams et al. (2010):

1. Preparación inicial de los Datos.

En este primer paso se deben seleccionar las variables observadas. Se calcula la matriz de correlaciones de las variables, de forma que, si las correlaciones entre ellas son muy bajas, indicará que no tiene sentido hacer el AFE ya que no se encontrarán dimensiones comunes.

2. Método de extracción.

Una vez determinada la necesidad de realizar el AFE, se debe seleccionar el método de estimación del modelo factorial. En esta tesis doctoral se utilizará el método de análisis de componentes principales (ACP) dado que es una de las técnicas más utilizadas por su sencillez para la estimación del modelo factorial (Abad et al., 2011; Barbero-García et al., 2011). Además, éste no realiza ningún tipo de suposición sobre la distribución de las variables observadas y es menos propenso a obtener soluciones incorrectas (Fabrigar et al., 1999).

³ Para explicar dicha técnica, se debe considerar que las dimensiones de calidad son los denominados “factores”, y los atributos son las denominadas “variables observadas”.

3. Selección del número de Factores.

El funcionamiento de este análisis estará basado en la siguiente lógica: Se extrae un primer factor y se elimina de la covarianza total existente entre las variables la parte explicada por el primer factor. Tras lo cual, si todavía queda parte de la covarianza por explicar, se extrae un segundo factor independiente del anterior. Se continúa así sucesivamente hasta que la covarianza residual (no explicada) sea no significativa (Fernando, 1993).

Para la evaluación de la significancia de la covarianza residual se suelen utilizar los valores propios de la matriz de correlaciones, que de acuerdo a Tabachnick and Fidell (2013) se pueden interpretar como la representación de la varianza de las variables observadas que es explicada por los factores.

Se utilizará en esta tesis doctoral el criterio de Kaiser-Guttman, que se basa en la premisa de que si el valor propio de un factor es menor que uno indicará que dicho factor no es capaz de explicar la varianza de las variables observadas que agrupa. Su aplicación consistirá en obtener los valores propios derivados de la matriz de correlaciones y determinar el número de factores con valores propios mayores que uno (Brown, 2014).

4. Rotación.

Determinado el número apropiado de factores, se procede a rotarlos, ya que, generalmente, los métodos de extracción de factores no proporcionan matrices de pesos factoriales de cada variable observada en cada factor adecuadas para la interpretación. Por tanto, se realiza una rotación de forma que se amplifiquen los pesos mayores y se reduzcan los más bajos.

De esta forma, se permite una mejor identificación de aquellas variables sobre las que influye o no el factor, ya que cada factor es definido por un subconjunto de variables observadas con altos pesos en el mismo, y cada variable observada (de forma idealizada) tiene un alto peso en un solo factor, siendo cercanos a cero sus pesos para el resto de factores.

Dos posibles tipos de rotación se pueden encontrar: la ortogonal, que considera que los factores están no correlacionados; y la oblicua, que sí permite dichas correlaciones. En esta tesis doctoral se utilizará la rotación ortogonal llamada "VARIMAX", dada su facilidad para interpretar sus resultados.

Se considera que, tras la rotación, variables con pesos factoriales superiores a 0,4 (Hair et al., 2010) en un determinado factor, son significativos para el mismo. En caso de que alguna variable observada no tenga un peso factorial cercano a estos umbrales en al menos un factor, se recomienda su no consideración.

5. Interpretación y cálculo de los valores.

Tras obtener una apropiada solución factorial, se les intenta dotar de diferentes nombres a cada dimensión en función de los atributos agrupados, así como de significado e interpretación.

En caso de que no se haya conseguido dicha solución apropiada, ya sea debido a la eliminación de alguna variable, o porque no sean interpretables los factores, o simplemente se desee incrementar el número de factores, o variar el método de rotación, se deberá empezar de nuevo el proceso a partir del paso 2.

De forma adicional en este paso, se obtendrán los valores de los factores mediante los pesos o mediante las correlaciones factoriales. Dichos valores serían aquellos que se obtendrían si hubieran sido observados o medidos directamente (Brown, 2014).

A modo de resumen de los pasos del AFE se muestra un esquema en la Figura 10.

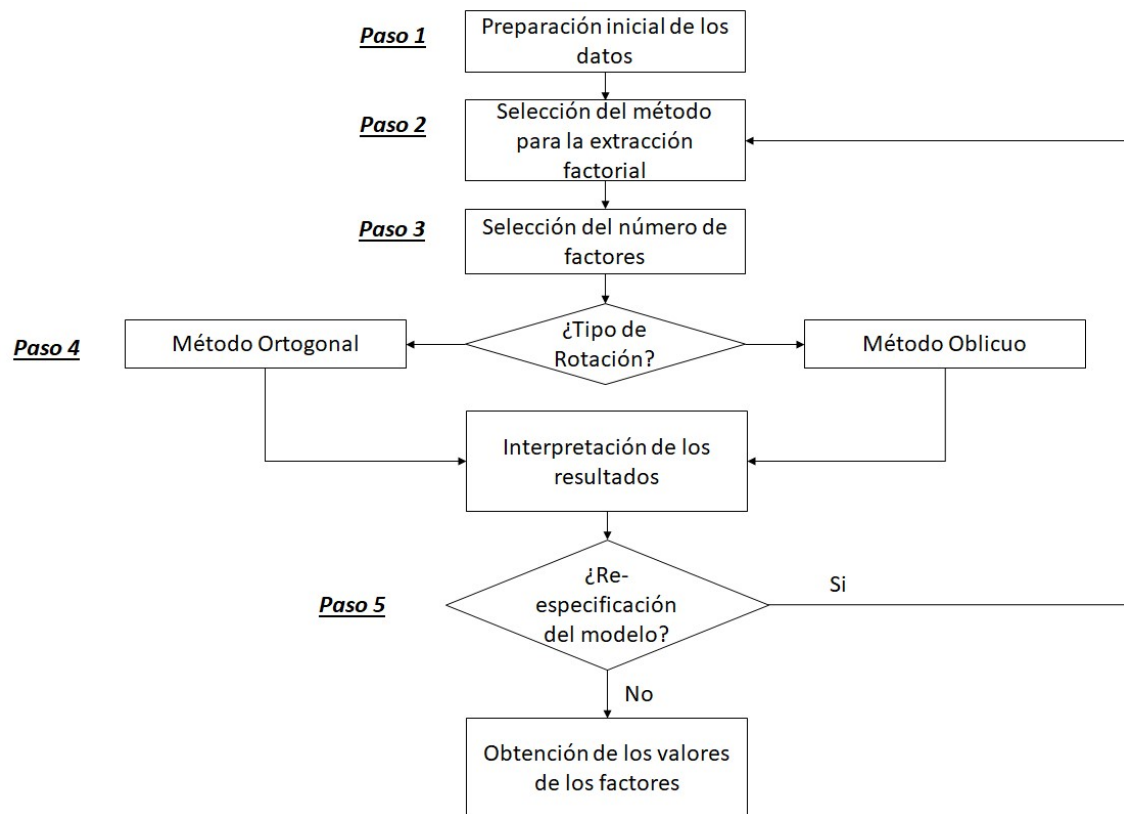


Figura 10. Esquema de los 5 pasos para realizar un AFE (Adaptado Hair et al., 2010).

4.1.1.2. ANÁLISIS FACTORIAL CONFIRMATORIO (AFC)

El AFC se trata de una técnica estadística de naturaleza confirmatoria, cuyo fin es testear el grado en el que una serie de variables observadas miden o representan un número menor de factores (Hair et al., 2010).

A diferencia del AFE, este tipo de análisis requiere el establecimiento previo del modelo, por tanto, requiere hipótesis y restricciones previas, y se debe destacar que en un AFC cada variable observada solo puede tener peso en un único factor.

El AFC se basa en tres tipos de validaciones de los factores: validez de convergencia, validez discriminante y validez nomológica.

- Validez de Convergencia.

Se define como la medición de cómo un conjunto de variables observadas representa un determinado número de factores. En consecuencia, se puede afirmar que hay validez de convergencia si, para cada factor, el conjunto de variables observadas que lo explican converge o muestra una alta proporción de varianza en común.

Diferentes formas hay de calcular esta validez. La primera es a través de los pesos de carga estandarizados, que si son altos mostrarán que las variables observadas convergen en un punto en común (el factor). Como mínimo, estos pesos deben ser significativos, y a modo de recomendación, se consideran adecuados si tienen valores estandarizados superiores a 0,5, e idealmente superiores a 0,7.

La siguiente forma es a través de la Media de la Varianza Extraída (Average Variance Extracted, AVE), calculada como la media de la varianza extraída para los pesos de las diferentes variables observadas en un factor (Fornell and Larcker, 1981). Un AVE superior a 0,5 de un determinado factor indicará que más de la mitad de la varianza del total de variables observadas que lo conforman es explicada por el mismo, y en caso contrario, es más explicado por los errores. Por tanto, cada factor debe tener su propio AVE.

Como última forma, se calcula la fiabilidad de los factores o si todas las variables observadas de un mismo factor lo representan adecuadamente o con consistencia. Diferentes debates hay sobre si es mejor el Alfa de Cronbach (Bacon et al., 1995; Cronbach, 1951) o la Fiabilidad del Constructo (Construct Reliability, CR) (Hair et al., 2010). En esta tesis doctoral se calcularán ambos.

Para el Alfa de Cronbach, calculado como una media ponderada de las correlaciones entre las variables observadas que conforman el factor, se consideran valores superiores a 0,5 como adecuados y de 0,7 como muy buenos para garantizar la fiabilidad del factor.

Para el CR, calculado como el cuadrado de la suma de los pesos de cada factor y la suma de los términos de la varianza del error de cada factor, valores superiores a 0,7 sugieren buena fiabilidad, y valores entre 0,6 y 0,7 deben ir acompañados de otros indicadores, como el Alfa de Cronbach, para ser aceptados.

- Validez Discriminante.

Esta tipología de validez comprueba si un factor es verdaderamente distinto del resto. Para ello se realiza la siguiente comprobación: se trata de comparar el AVE de cada factor con el cuadrado de la correlación estimada entre dicho factor y cada uno de los restantes (Fornell and Larcker, 1981). Dicha correlación debe ser inferior al AVE, mostrando que la varianza explicada por el factor es mayor que la compartida con otro factor, y por tanto, que ambos factores explican aspectos diferentes.

Dado que el cumplimiento de esta prueba suele ser complicado, especialmente en casos como los de esta tesis doctoral, donde se utilizan datos que provienen de encuestas de satisfacción y las dimensiones extraídas suelen solaparse o compartir cierta información, se puede recurrir a otros tipos de test, menos fiables, pero que pueden servir para soportar la validez discriminante, en caso de no cumplir la prueba previamente señalada, como es el caso de los test de diferencia de Chi-Cuadrado (Werner and Schermelleh-Engel, 2010).

En consecuencia, si se considera que un AFE explora los datos y provee información sobre cuántos factores, concebidos como agrupación de variables observadas, son necesarios para conseguir la mejor representación de los datos, AFC trata de confirmar que dichas agrupaciones de las variables observadas y el número de factores son los adecuados.

- Validez Nomológica.

Por último, tal y como se indica en Hair et al. (2010), se comprobará que las agrupaciones de las variables observadas tengan coherencia dentro del marco teórico donde se encuadra dicho modelo factorial. Dicho trabajo es desempeñado por el investigador quien debe chequear las relaciones desde este punto de vista.

Se debe resaltar que si en algún momento cierta variable o factor no cumpliera con algún tipo de validez el modelo deberá ser re-definido.

Un esquema de los pasos a seguir para realizar un AFC, es indicado en la Figura 11.

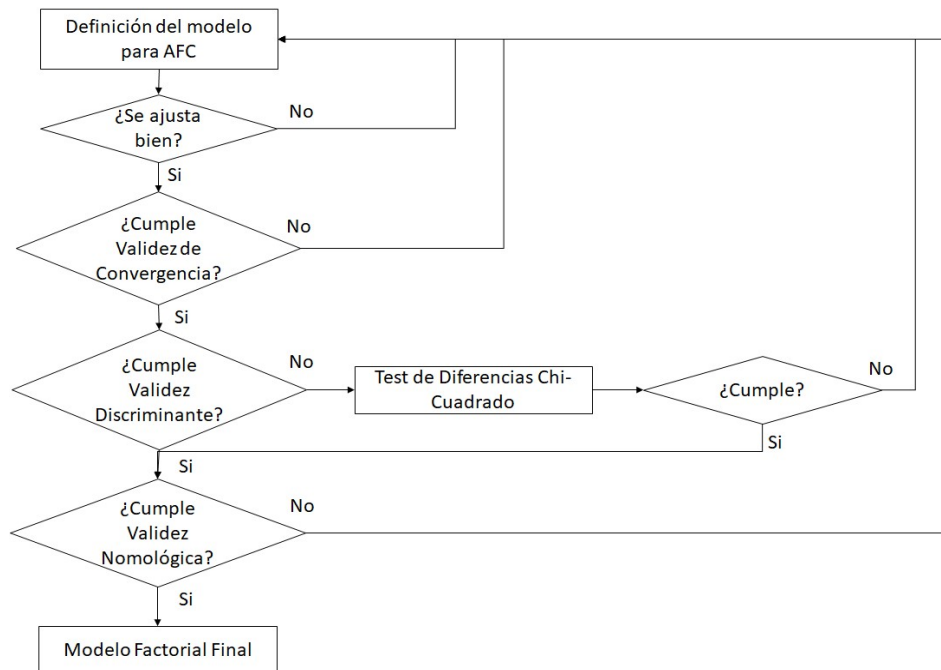


Figura 11. Esquema de pasos para realizar un AFC (Adaptado y completado de Hair et al., 2010).

4.1.2. REDES BAYESIANAS (RB)

Las RB se trata de una técnica estadística avanzada enmarcada en el campo de la Minería de Datos, y que pertenece a la familia de los modelos gráficos probabilísticos usados para representar conocimiento sobre un campo desconocido.

Formalmente, una RB (B) se define como un Grafo Dirigido Acíclico (Directed Acyclic Graph, DAG) que representa una distribución de probabilidad conjunta sobre un conjunto de variables aleatorias V . Dicho lo cual, se pueden diferenciar dos partes fundamentales que componen las RB ($B = (G, \Theta)$). Por un lado, una parte cualitativa (G), que sería el DAG, que está formado por nodos X_1, \dots, X_n , que representan las variables aleatorias consideradas en el estudio, y por arcos o flechas, que representan las dependencias probabilísticas directas entre estas variables. Y, por otro lado, una parte cuantitativa (Θ) que sería aquella en la que se enmarcan todas las distribuciones conjuntas o condicionales de probabilidad, que son establecidas y calculadas a partir de las relaciones existentes y no existentes en el DAG (Pearl, 2000).

Comenzando por la definición de la parte cualitativa, se puede afirmar que la estructura del DAG (G) se define, a su vez, por dos sub-conjuntos: los nodos y los arcos. Los nodos, representan las variables aleatorias consideradas y son dibujados como círculos etiquetados con el nombre de las variables. Por su parte, los arcos representan dependencias directas de probabilidad entre las variables que unen. En particular, se puede decir que un arco que vaya del nodo X_i al X_j , representa una dependencia estadística entre ambas variables, es decir, que el valor tomado por la variable X_j depende del valor tomado por la variable X_i .

La terminología utilizada para referirse a estos nodos es la siguiente: a X_i se le llama “Padre” de X_j , y a este último “Hijo” de X_i . De esta forma, una extensión genealógica puede ser aplicada para definir al conjunto de nodos que pueden alcanzar a este nodo padre siguiendo un camino directo de arcos, denominándolos “Descendientes”. Similarmente, todos los nodos que preceden y pueden alcanzarse por un camino directo de arcos desde un determinado nodo hijo, se les denominan “Ancestros”.

Se destaca que, en el ámbito de la causalidad, aunque los arcos representen conexiones causales directas entre las variables, un proceso de razonamiento puede operar en las RB propagando información en cualquier dirección (Kjaerulff and Madsen, 2013).

La estructura dirigida acíclica garantiza que ningún nodo sea ancestro o descendiente suyo propio, lo cual, es condición fundamental para la aplicación de esta técnica. Teniendo en cuenta esta consideración, se puede afirmar que el grafo G codifica hipótesis de independencia, tales como que cada variable X_i es independiente de sus no descendientes dados sus padres en G , llamados genéricamente como π_i .

Este hecho conforma la afirmación de independencia condicional más efectiva y vital para la realización de los cálculos en las RB, ya que reduce significativamente el número de parámetros requeridos para caracterizar la distribución de probabilidad conjunta de las variables (Joint Probability Distribution, JPD), componente fundamental de la parte cuantitativa, y conforma una vía eficiente de computar las probabilidades a posteriori (Jensen, 2001; Kjaerulff and Madsen, 2013; Lauritzen and Spiegelhalter, 1988; Pearl, 2000).

En cuanto a la parte cuantitativa, se observa que se compone por los parámetros del modelo (Θ), que se obtienen mediante la aplicación de la propiedad de Markov. Esta propiedad expone que la JPD de cada nodo depende exclusivamente de sus padres.

Para el caso de variables aleatorias discretas, como es el caso que ocupa en esta tesis doctoral, esta probabilidad condicional es representada en una tabla, mostrando la probabilidad condicional que en un nodo hijo puede tomar para cada una de sus categorías, de acuerdo a una combinación de los valores con las categorías de sus padres. La distribución conjunta de una colección de variables puede ser determinada únicamente mediante estas tablas de probabilidad condicional local (Kenett and Salini, 2012).

De una manera más formal, este conjunto Θ de la red contienen el parámetro $\Theta_{x_i|\pi_i} = P_B(x_i|\pi_i)$ para cada categoría x_i de X_i condicionada π_i , siendo X_i el conjunto de padres en G . Por consiguiente, B define una distribución de probabilidad conjunta única sobre V , llamada:

$$P_B(X_1, \dots, X_n) = \prod_{i=1}^n P_B(X_i|\pi_i) = \prod_{i=1}^n \Theta_{X_i|\pi_i} \quad (1)$$

Para simplificar la ecuación (1) se ha omitido el subíndice B en la última expresión. Si X_i no tiene padres, su distribución de probabilidad se dice que es “incondicional”, y en caso contrario, “condicional”. Si la variable representada por un nodo es observada, el nodo se dice que es un nodo “evidencia”, en caso contrario, es “escondido o latente”.

Diferentes funciones pueden ser llevadas a cabo con las RB. de Oña et al. (2011) ilustran su función clasificatoria, que es definida como el proceso que consiste en clasificar una variable $y = X_o$, llamada variable clase, dado un conjunto de variables $U = X_1, \dots, X_n$, llamadas variables atributos.

Un clasificador $h: U \rightarrow y$ es una función que mapea una instancia de U a un valor de y . El clasificador es aprendido a partir de una base de datos que consiste en una muestra de (U, y) . La tarea de aprendizaje consiste en encontrar una RB apropiada dada unos datos sobre U .

Otra de las grandes funciones que presentan las RB es la denominada inferencia, definida como su capacidad para simular diferentes escenarios con evidencias hipotéticas u observadas. De manera más formal, dada una RB y sus JPD, si se considera una evidencia $E = (E_1, \dots, E_n)$, que es información recibida a través de fuentes externas y que define un estado específico de una variable o de un conjunto de variables en la red, la probabilidad de ocurrencia de algunos eventos puede ser calculada como:

$$P(U|E) = \frac{P(U, E)}{P(E)} \quad (2)$$

4.1.3.1. ALGORITMOS DE APRENDIZAJE DE LAS REDES BAYESIANAS

En esta sección se procede a describir los algoritmos para calcular las RB. Posteriormente, en la siguiente sección se describe con detalle la metodología de selección de la red más robusta basada en Cugnata et al. (2016).

El primer paso en el aprendizaje de las RB es la determinación de su DAG a partir de unos datos e información previa. Tras lo cual se procede al cálculo de los parámetros que conforman la JPD en la RB. En otras palabras, el primer paso consistirá en el aprendizaje de la parte cualitativa, y posteriormente de la cuantitativa.

Como ya se estableció en la sección 2.2.2.1, diferentes metodologías pueden ser usadas para el aprendizaje del DAG (Tradicional, Aprendizaje automático y Mixto). En esta tesis doctoral se ha seleccionado la metodología Mixta, debido a que se obtienen mejores resultados con esta metodología en lo que a fiabilidad del modelo y ajuste a la realidad se refiere, y a que en la literatura actual, existe una serie de premisas que deben ser consideradas en los modelos de calidad del servicio en el transporte público, previo al aprendizaje del DAG. Dicha metodología combina, por una parte, la imposición de ciertas hipótesis en el DAG en forma de existencia o prohibición de relaciones, y por otra, el aprendizaje automático de la RB con un algoritmo.

Para dicho aprendizaje automático de las RB, existe actualmente en la literatura un amplio rango de algoritmos que presentan diferentes características estadísticas específicas en su proceso de selección y aprendizaje de la RB.

Se definen a continuación los más populares, los cuales serán utilizados en esta tesis doctoral:

- Grow-Shrink (Margaritis, 2003). Está basado en el concepto del Manto de Markov de un nodo o variable que puede definirse como el conjunto de variables que hacen independiente a ésta del resto de la red, y consta de dos fases: Crecimiento (growing) y Reducción (shrinking). Se comienza introduciendo una variable X en un conjunto vacío H . Durante la primera fase (crecimiento) se van añadiendo nuevas variables a H , una a una, siempre que cumplan la condición de que sean dependientes de la variable X , de acuerdo al último contenido de H . En consecuencia, durante dicha fase se puede advertir que algunas de las variables introducidas a lo largo del proceso pueden quedar fuera del Manto de Markov, o lo que es lo mismo, pueden volverse independientes de la variable X al actualizarse el contenido en H . Por lo que se identifican, y durante la segunda fase (reducción) se eliminan.

- Incremental Association Markov Blanket (Tsamardinos et al., 2003). Se basa igualmente en dos fases, una de avance y otra de retroceso. Similar al algoritmo anterior, durante la primera fase, se parte de un conjunto vacío H y se prueban diferentes hipótesis de dependencias, buscando aquellas que maximizan una determinada función heurística. Este hecho significaría que la variable es miembro del Manto de Markov de la variable X . Durante la segunda fase aquellas variables que han proporcionado un falso positivo y se han introducido en un principio se identifican mediante test de independencia y son eliminadas.
- Fast Incremental Association Markov Blanket (Yaramakala and Margaritis, 2005). Se trata de una variante del algoritmo Incremental Association Markov Blanket, que realiza las mismas dos fases que dicho algoritmo, con la diferencia de que utiliza la selección progresiva especulativa para reducir el número de test de independencia condicional de la segunda fase.
- Interleaved Incremental Association Markov Blanket (Yaramakala and Margaritis, 2005). Otra variante del algoritmo Incremental Association Markov Blanket. En este caso, se utiliza la selección paso a paso para evitar falsos positivos en la primera fase, donde se intenta identificar el Manto de Markov de una determinada variable X , intercalando la fase de Reducción con la de Crecimiento para mantener el tamaño del Manto de Markov lo más pequeño posible al eliminar inmediatamente los falsos positivos lo más rápido posible.
- Hill Climbing (Lin and Kernighan, 1973; Selman and Gomes, 2006). Se trata de un algoritmo Greedy, o lo que es lo mismo, un algoritmo que para resolver un problema determinado sigue una heurística consistente en elegir la opción óptima en cada paso local. Dicho algoritmo, partiendo de una red aleatoria, busca máximos locales mediante la eliminación, adición y/o cambio de sentido de los diferentes arcos de la red, usando una función de puntuación previamente seleccionada.
- TABU Search (Glover and Laguna, 1997). Se trata de otro algoritmo Greedy, basado en una modificación del algoritmo Hill Climbing, que evita los óptimos locales mediante la selección de una red que reduce al mínimo la función de puntuación.
- Max-Min Hill Climbing (Tsamardinos et al., 2006). Se trata de un algoritmo que combina el funcionamiento del algoritmo Max-Min Parents and Children (algoritmo basado en la maximización de la medida de asociación mínima observada con cualquier sub-conjunto de nodos seleccionados en previas iteraciones) para restringir el espacio de búsqueda, y el algoritmo Hill Climbing, para encontrar la red óptima en el espacio delimitado.

Los algoritmos de puntuación (Hill Climbing, TABU Search y Max-Min Hill Climbing) necesitan de una función de puntuación para realizar sus búsquedas de máximos y comparar la calidad de las diferentes RB que internamente simulan.

Para ello, en esta tesis doctoral se han utilizado las siguientes funciones de puntuación:

- Función basada en el Criterio de Información de Akaike (AIC) (Akaike, 1974). Calcula su valor en función del número de parámetros en el modelo estadístico y el máximo valor de la función de verosimilitud para el modelo estimado.
- Función basada en el Criterio de Información de Bayesiana (BIC) (Schwarz, 1978). Utiliza los mismos parámetros que el Criterio de Información de Akaike, aunque la expresión para su cálculo es diferente al variar ciertas constantes.
- Función del logaritmo de la verosimilitud multinomial (LogLik) (Scutari, 2010). Se basa en encontrar el valor de cierto parámetro asociado al modelo evaluado, en este caso a la RBs evaluadas, que maximiza la función del logaritmo de la verosimilitud multinomial.
- Función Bayesian Dirichlet Equivalent (BDe) (Heckerman, 1995). Se selecciona la RB que maximiza una determinada función de verosimilitud equivalente basada en la probabilidad conjunta de las variables.
- Función Bayesian Dirichlet Equivalent modificada (mBDe) (Heckerman, 1995). Similar a la anterior, modificando ciertos parámetros y considera distribuciones de probabilidad uniformes.
- Función K2 (Cooper and Herskovits, 1992). Busca maximizar la probabilidad de una RB dado un conjunto de datos. La principal diferencia con las anteriores es la ecuación propuesta para dicha maximización.

Una vez conocido el DAG el objetivo es calcular los valores de los parámetros que conforman la red a través de la maximización del logaritmo de la verosimilitud del conjunto de entrenamiento. Generalmente, cuando se aplica una técnica procedente de la Minería de Datos se realiza una separación de los datos de estudio en dos conjuntos, llamados respectivamente, de entrenamiento y de prueba o test. Cada conjunto tiene su función específica, así como su propia proporción de datos, siendo significativamente mayor en el conjunto de entrenamiento que en el de prueba o test (p.e., 70%/30% o 80%/20%). Si se focaliza en el caso de las RB, el conjunto de entrenamiento sirve para el aprendizaje del DAG y de JPD, mientras que el de prueba o test se utiliza para validar el rendimiento de la RB aprendida.

4.1.3.2. MÉTODO DE OBTENCIÓN DE LA RED BAYESIANA MÁS ROBUSTA

Si se aplican diferentes algoritmos para aprender una RB a partir de una misma base de datos, se obtendrán diferentes RBs que presentarán algunos arcos recurrentes en un alto porcentaje de las redes y otros que aparecen de forma aislada en una única red o en un pequeño número de ellas. Con el objetivo de obtener una RB que explique lo mejor posible el fenómeno estudiado se puede calcular la frecuencia de aparición de los arcos respecto al total de RB aprendidas con diferentes algoritmos, de forma que se puede afirmar que en la RB más robusta reaparecerán los arcos específicos que se repiten en la mayoría de las RB aprendidas. Por consiguiente, estos arcos de la RB más robusta, se puede afirmar que no dependen del algoritmo de aprendizaje utilizado, ni son producidos por outliers o grupos de observaciones con valores atípicos.

En esta tesis doctoral, la metodología para conseguir la RB más robusta va a seguir el proceso detallado en el trabajo de Cugnata et al. (2016) y de una modificación del código en R suministrado en dicho trabajo. El fin será la selección de la RB más robusta de acuerdo a diferentes algoritmos de aprendizaje y a la medición de la fuerza de los arcos (relaciones) usando un método iterativo de re-muestreo. Dicho proceso puede ser aplicado en cualquier caso que se utilice RB, así mismo sus pasos son fácilmente interpretables e intuitivos. De igual modo, se debe destacar que el caso expuesto en Cugnata et al. (2016) prueba su eficiencia y efectividad sobre datos de encuestas de satisfacción, y ha sido probado en otros estudios de mejoras de calidad y estrategias de mercado (p.e., Coscia et al., 2018; Di Pietro et al., 2017). Por lo que su uso en esta tesis doctoral está más que justificado.

Este método de selección de la RB más robusta consta de 3 pasos:

1. Aprendizaje de diferentes RB.

El primer paso, consiste en aprender una amplia gama de RB utilizando diferentes algoritmos. En el caso de esta tesis doctoral, se han considerado los expuestos anteriormente (apartado 4.1.3.1): Grow-Shrink, Incremental Association Markov Blanket, Fast Incremental Association Markov Blanket, Interleaved Incremental Association Markov Blanket, Hill Climbing, TABU Search y Max-Min Hill Climbing.

Para la construcción de dichas RB se utiliza una distribución de 70/30 para establecer el conjunto de entrenamiento y test.

2. Selección de la RB más robusta.

Se selecciona la RB que contiene el mayor número de arcos repetidos en los modelos implementados con los diferentes algoritmos. Para ello se crea una tabla que reporta la ocurrencia de un arco entre dos nodos en cada una de las RB aprendidas. Sus filas son los diferentes posibles arcos que se pueden dar en la red, y sus columnas los diversos algoritmos considerados.

Es decir, cada posible relación (fila) es analizada para cada RB aprendida por un algoritmo (columna), si dicha relación no existe en tal RB, se le asigna un valor de 0 en dicha posición de la tabla, en caso de que exista un arco dirigido, se le asignará el valor de 1, y en el caso de que el algoritmo haya establecido un arco entre esos dos nodos, pero éste arco no esté dirigido se le asignará el valor de 0,5. En la última columna se reporta la puntuación total de cada arco, lo que según Cugnata et al. (2016) puede interpretarse como la fuerza del arco.

Se selecciona la RB que tenga una mayor cantidad de arcos con una fuerza igual o superior a un umbral establecido. Cugnata et al. (2016) recomienda que este umbral se establezca en una ocurrencia del 70% de los modelos aprendidos.

En esta tesis doctoral, se seguirán tales recomendaciones para establecer el umbral, y se utilizarán los valores de 1, 0,5 y 0 para puntuar los arcos.

Finalmente, la RB que presente mayor número de arcos con puntuaciones iguales o superiores al umbral es identificada como la más robusta. Su causa subyace en que se está seleccionando la RB cuya estructura mejor explica el fenómeno estudiado y se encuentra poco influenciada por la presencia de outliers o grupos de observaciones atípicas.

En caso de que haya más de una RB con el mismo número de arcos con una ocurrencia por encima del umbral seleccionado, se recurre a seleccionar aquella que produzca un menor ratio de casos mal clasificados (missclassification rate) utilizando la base de datos de test. Es decir, utilizando las RBs como clasificadores, donde se predice el valor de unos datos de entrada, y la base de test que son datos observados, se puede calcular los casos mal clasificados de cada RB y seleccionar la que produzca un ratio menor.

3. Análisis de la importancia de las dimensiones.

El análisis de la importancia de las dimensiones sobre la calidad global del servicio se realiza mediante un re-muestreo iterativo (bootstrap resampling) sobre la base de datos inicial. Para ello, se seguirá el proceso definido en Cugnata et al. (2016), donde se generan 1.000 sub-bases aleatorias, cada una con proporción de datos de 1/10 sobre el total de la base de datos inicial.

Estos valores son considerados adecuados y suficientes desde el punto de vista de la capacidad computacional necesaria, así como han sido usados por otros autores (Friedman et al., 1999). De esta forma, se vuelve a aprender una RB para cada una de las sub-bases generadas utilizando el mismo algoritmo que ha proporcionado la RB más robusta. A partir de los resultados obtenidos, se muestra el mismo DAG de la RB más robusta, pero mostrando la proporción de ocurrencia de cada arco (fuerza del arco) en dicho proceso de re-muestreo iterativo.

Siguiendo a Albrecht et al. (2014) se define una métrica denominada Influencia de la Distancia-Ponderada (Distance-Weighted Influence, DWI) que determina un ranking de la influencia de las variables sobre un nodo objetivo basándose en la estructura de la red. En este ranking se refleja que cuanto más largo es el camino entre un nodo y el nodo objetivo, menor es la influencia de ese nodo, mientras que la influencia se incrementa conforme se incrementa el número de caminos que conectan ambos nodos.

Por lo tanto el DWI entre X e Y, $DWI(X, Y; w)$ se define de la siguiente forma:

$$DWI(X, Y; w) = \sum_{s \in S(X, Y)} w^{|s|} \quad (3)$$

Donde $S(X, Y)$ es el conjunto de caminos en la RB que unen los nodos X e Y, $|s|$ es la longitud del camino s , y w es el peso del camino. Este peso w se define como el producto de la fuerza de los arcos que componen el camino. Esta métrica DWI se interpreta como la importancia de un nodo respecto a un nodo objetivo (en este caso se referirá a la importancia de una dimensión sobre la calidad global del servicio). El nodo que no sea antecesor de un nodo objetivo (no existe ningún camino que los dirija a este nodo), el índice de importancia se establece en cero.

4.1.4. MODELOS DE ECUACIONES ESTRUCTURALES (SEM)

Los SEM se pueden definir como una técnica multivariada que combina la técnica de regresión, el análisis factorial y el análisis de varianza, y cuya principal función es estimar de forma simultánea relaciones de dependencia inter-relacionadas. Este hecho representa una gran ventaja frente al resto de técnicas estadísticas de modelización.

Dicha tarea se realiza a través de una serie de ecuaciones que describen las relaciones entre los constructos del análisis, y que son similares a las que se utilizan en regresión múltiple.

En consecuencia, si se toma como ejemplo básico de un SEM la Figura 12, se puede observar que está compuesto por:

- Variables observadas (Z_i). Se trata de todas las variables que representan conceptos o fenómenos que pueden ser medidos directamente y se encuentran registradas en una base de datos. En el caso de estudios de calidad del transporte público, estas variables se identifican con los atributos. Suelen ser representadas en los gráficos SEM en rectángulos.
- Variables de error (e_i). Estas variables latentes tienen en cuenta todas las fuentes de variación que no están consideradas en el modelo, como por ejemplo, en la medición de las variables.
- Variables latentes (*Factores*). Son todos aquellos factores latentes que son explicados por las variables observadas. Suelen ser representadas por círculos en los gráficos SEM.
- Relaciones entre las variables. Son las flechas que salen de las variables y que las conectan entre sí, creando finalmente una estructura gráfica. Se trata de hipótesis (H_i) de relaciones entre variables y factores que van a ser testeadas por la metodología SEM.

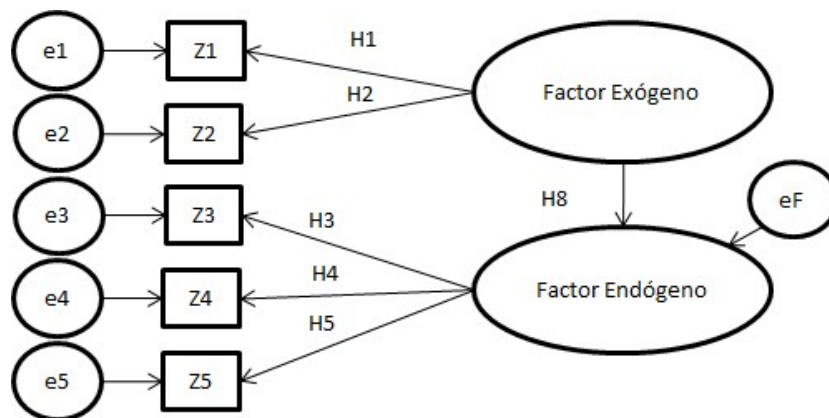


Figura 12. Tipos de variables en un modelo SEM.

Considerando que los constructos son los factores (variables latentes) que, a su vez, están representados por los diferentes atributos (variables observadas), se puede afirmar que esta técnica permite modelizar un fenómeno mediante la consideración de manera simultánea, tanto de variables latentes (factores), como de variables observadas (atributos).

Al igual que ocurre con otras técnicas estadísticas, en SEM también es importante diferenciar entre variables dependientes e independientes (p.e., Análisis de Regresión, MANOVA, Análisis discriminante Múltiple, etc.). No obstante, en estos modelos la terminología utilizada es diferente debido a que unos factores pueden explicar a otros factores (Hair et al., 2010). Dicho lo cual, se pueden distinguir dos clases de factores (Figura 12):

- Factores Exógenos. Son variables latentes equivalentes a las variables independientes. Dichas variables, que pueden ser factores explicados por una gama de atributos o éstos en sí mismo, actúan como variables independientes, es decir, que no son explicadas por otros factores en el modelo. Gráficamente, serían aquellas que no reciben arcos, sino que los arcos parten de ellas.
- Factores Endógenos. Son variables latentes equivalentes a las variables dependientes, y son determinadas teóricamente por otros factores dentro del modelo. Su representación gráfica se detalla con arcos que salen de los factores exógenos hacia ellas.

Estas tipologías de variables conforman las unidades básicas de los modelos SEM, que se pueden descomponer en dos partes. Una denominada “modelo de medición” que calcula los constructos (factores latentes) como una función lineal de las variables observadas (atributos). Y otra denominada “modelo estructural” que muestra la dirección y la fuerza de las relaciones de los diferentes factores latentes.

Atendiendo a la Figura 13, se pueden observar los pasos recomendados para realizar un modelo SEM y que van a ser desarrollados a continuación.

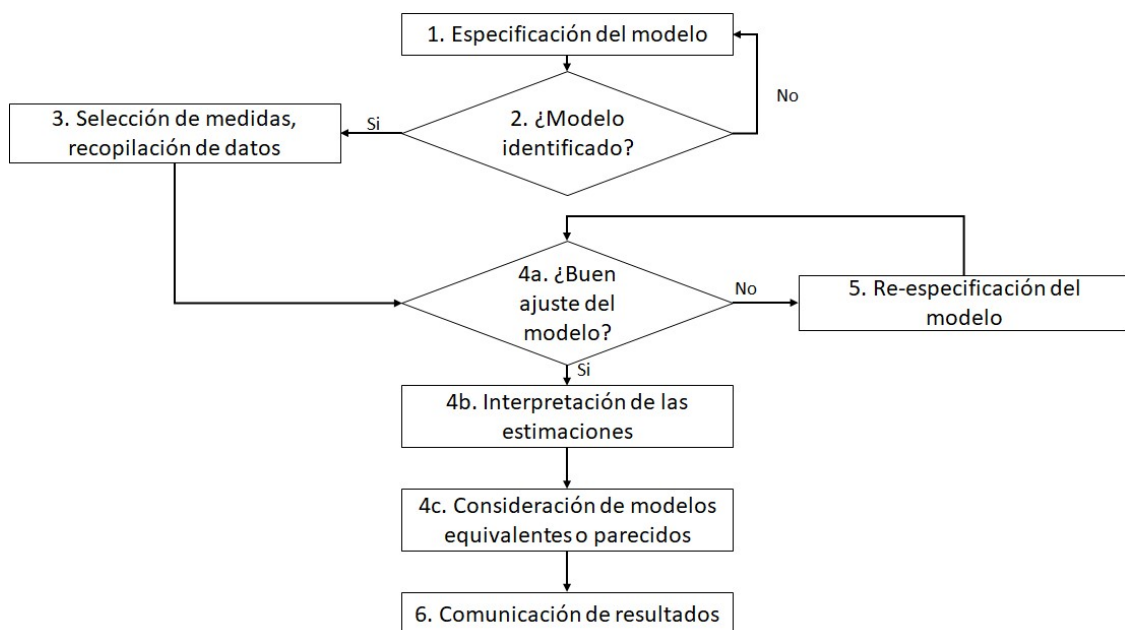


Figura 13. Pasos básicos recomendados del SEM (Kline, 2011).

1. Especificación.

Dado que se trata del desarrollo como tal del modelo, se debe destacar que cada parte del mismo (estructural y de medición) tiene sus propias ecuaciones que la gobiernan.

En el caso del modelo estructural se establece que habrá tantas ecuaciones estructurales como factores sean explicados por otras variables exógenas. Bollen (1989) establece la ecuación básica de este modelo:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (4)$$

Donde η es un vector $m * 1$ de factores endógenos, ξ es un vector $n * 1$ de factores exógenos, B es una matriz $m * n$ de coeficientes asociados con los factores endógenos, Γ es una matriz $m * n$ de coeficientes asociados con los factores exógenos, y ζ es un vector $m * 1$ de términos de error asociados con los factores endógenos.

Por su parte, en el modelo de medición habrá tantas ecuaciones como relaciones entre atributos y factores. Sus ecuaciones básicas pueden ser expresadas como sigue:

$$x = \Lambda_x\xi + \delta \quad (5)$$

$$y = \Lambda_y\eta + \varepsilon \quad (6)$$

Donde x y δ son q-vectores columna relacionados con los atributos exógenos y los errores, respectivamente, Λ_x es una matriz $q * n$ de coeficientes estructurales de los efectos de los factores exógenos sobre los atributos, y y ε son p-vectores columna relacionados con las atributos endógenos y los errores, respectivamente, y por último, Λ_y es una matriz $p * n$ de coeficientes estructurales de los efectos de los factores endógenos sobre las atributos.

2. Identificación.

Una vez especificado el modelo, se pasa a identificarlo. Es decir, a comprobar si los parámetros del modelo pueden estimarse a partir de los elementos de la matriz de covarianzas de los atributos. Para ello se utiliza el concepto de grado de libertad, definido como:

$$g = \frac{s * (s + 1)}{2} - t \quad (7)$$

Siendo s el número total de variables, t el número total de parámetros que han de ser estimados y g los grados de libertad.

Obtenido este valor, se pueden clasificar los modelos en identificables o no:

- $g < 0$. No identificable, los parámetros toman infinitos posibles valores y, por tanto, están indeterminados.
- $g = 0$. Posiblemente identificados, es decir, puede existir una única solución para los parámetros que igualen a la matriz de covarianzas observada e implicada.
- $g > 0$. Posiblemente sobre-identificados. No existe ninguna solución para los parámetros que igualen la matriz de covarianzas observada, pero puede existir una única solución que minimice los errores entre ambas matrices.

3. Selección de medidas y recopilación de datos.

En esta tesis doctoral se van a seguir algunos de los índices y umbrales propuestos por Hooper et al. (2008), los cuales, en su mayoría, son los utilizados con general reconocimiento en los trabajos de investigación sobre calidad del servicio con modelos SEM (p.e., de Oña et al., 2013a; de Oña et al., 2016b; Rahman et al., 2016; Su et al., 2016; etc.).

Tal y como Crowley and Fan (1997) citan, no hay una regla de oro para elegir los índices de ajuste, pero sí es necesario utilizar una variedad de los mismos, ya que diferentes tipos de índices reflejan diferentes aspectos de ajuste del modelo.

Se pueden distinguir tres tipos de índices que miden el ajuste del modelo:

1. Índices de Ajuste Absoluto.

Esta tipología determina cómo de bien un modelo a priori se ajusta a los datos (McDonald and Ho, 2002). Las medidas de esta tipología representan un indicador fundamental sobre cómo se ajusta la teoría propuesta a los datos y, además, las proporciona sin necesidad de realizar ningún tipo de comparación (Jöreskog and Sörbom, 1993).

Entre los índices se encuentran:

- Modelo Chi-Cuadrado (X^2).

El valor X^2 proporciona la magnitud de discrepancia entre la muestra y la matriz de covarianza ajustada (Hu and Bentler, 1999). De esta forma, se considerará un modelo adecuado si este ofrece un resultado insignificante a un nivel umbral de 0,05 (Barrett, 2007).

A pesar de su popularidad como estadístico de ajuste presenta ciertas desventajas, como que asume normalidad multivariable (McIntosh, 2006), lo que hace que si se producen desviaciones severas puede llevar a rechazar modelos que estén adecuadamente especificados. De igual modo, es sensible al tamaño de muestra, de forma que en muestras grandes casi siempre rechaza el modelo (Joreskog and Sorbom, 1993) y en pequeñas no tiene suficiente entidad como para diferenciar entre buen o mal ajuste (Kenny and McCoach, 2003).

- Raíz Cuadrada del Promedio de los Residuos de Aproximación (Root Mean Square Error of Aproximation, RMSEA).

Propuesto por Steiger and Lind (1980), indica cómo de bien un modelo desconocido se ajustaría a la matriz de covarianza de los datos, pero con la estimación de los parámetros elegidos óptimamente (Byrne, 1998). Una de las propiedades de este indicador es que favorece la parsimonia ya que elige el modelo que menor número de parámetros presente. Tradicionalmente los valores entre 0,08 y 0,10 se consideraban valores mediocres de ajustes, y si estos estaban por debajo de 0,08 se consideraban buenos. Recientemente, se establece que deben ser cercanos a 0,06 (Hu and Bentler, 1999) y con un límite superior estricto a 0,07 (Steiger, 2007) para considerarse apropiados.

Una de sus grandes ventajas es su capacidad para calcular un intervalo de confianza en torno al valor del estadístico (MasCallum et al., 1996).

- Índice de Bondad de Ajuste (Goodness of Fit Index, GFI) e Índice de Bondad de Ajuste Corregido (Adjusted Goodness of Fit Index, AGFI).

El GFI calcula la proporción de varianza que es representada por la covarianza de la población estimada (Tabachnick and Fidell, 2013), y presenta valores entre 0 y 1, estableciéndose tradicionalmente su umbral en 0,90.

Se debe destacar que el GFI se encuentra influenciado por el tamaño de la muestra y por el número de parámetros (MacCallum et al., 1996; Sharma et al., 2005), de forma que con un número elevado de parámetros o en muestras grandes este tiende a presentar un mayor sesgo (Bollen, 1990; MasCallum and Hong, 1996; Miles and Shevlin, 1998). Por el contrario, cuando el número de grados de libertad es grande en comparación con el tamaño de la muestra, tiende a mostrar un menor sesgo (Sharma et al., 2005).

Por su parte, el AGFI ajusta el GFI en función de los grados de libertad, con modelos más saturados que reducen el ajuste (Tabachnick and Fidell, 2013). De esta forma los modelos más parsimoniosos son premiados frente a los complejos. Este indicador se encuentra influenciado también por el tamaño de la muestra, tendiendo a incrementar su valor con el tamaño de la muestra. Al igual que el GFI, puede tomar valores entre 0 y 1, siendo los valores superiores a 0,90 considerados como buena señal de ajuste del modelo.

En conclusión, debido a su histórica importancia suelen ser utilizados en estudios de análisis de estructuras de covarianza. Sin embargo, ya que ambos índices son influenciados de forma relevante por el tamaño de la muestra no se recomienda su utilización de forma aislada, sino en conjunto con otros índices (Hooper et al., 2008).

- Raíz Cuadrada del Promedio de los Residuos (Root Mean Square Residual, RMR).

Este indicador calcula la raíz cuadrada de la diferencia entre los valores residuales de la matriz de covarianza de la muestra y la covarianza hipotética del modelo. Este índice se calcula en base a las escalas de cada variable observada. Por tanto, si hay diferentes escalas puede ser muy difícil interpretarlo (Kline, 2011). No obstante, se puede afirmar que generalmente valores cercanos a 0 son considerados como sinónimos de buen ajuste del modelo (Schermelleh-Engel et al., 2003).

2. Índices de Ajuste Incrementales.

Esta tipología de índices compara el X^2 del modelo calculado con el X^2 del modelo nulo, siendo éste aquel en el que todas las variables observadas están no correlacionadas (McDonald and Ho, 2002).

- Índice de Ajuste Normalizado (Normed Fit Index, NFI).

Aunque sus valores pueden oscilar entre un rango 0 y 1, recientemente, Hu and Bentler (1999) han señalado que valores superiores o iguales a 0,95 son sinónimo de buen ajuste del modelo, mientras que Bentler and Bonnet (1980) establecieron este umbral en 0,90. Una gran desventaja de este índice es su sensibilidad al tamaño de la muestra, subestimando el ajuste para muestra menores de 200 (Bentler, 1990; Mulaik et al., 1989). Por lo que su aplicación se recomienda que sea acompañando a otros índices (Kline, 2005).

- Índice de bondad de ajuste comparativo (Comparative Fit Index, CFI).

Se trata de una revisión del NFI en el que se considera la influencia del tamaño de la muestra (Byrne, 1998) de forma que produzca resultados satisfactorios para muestras de pequeño tamaño (Tabachnick and Fidell, 2013). Su funcionamiento es similar al NFI y su rango es también de 0 a 1, siendo 1 un ajuste perfecto. Hu and Bentler (1999) establecen que es necesario un valor superior o igual a 0,90 para asegurarse que modelos no específicos no sean aceptados y que, a partir de 0,95, se reconoce como buen ajuste.

Se debe destacar que, actualmente, se trata de uno de los índices más populares e incluido en todos los estudios de SEM debido a su mínima influencia por parte del tamaño de la muestra (Fan et al., 1999).

3. Índices de Ajuste Parsimoniosos.

Estos índices fueron desarrollados para evitar que en modelos complejos casi saturados, donde el proceso de estimación depende de los datos de la muestra, den como resultado un modelo teórico menos riguroso pero con mejores índices de ajuste (Mulaik et al., 1989; Crowley and Fan, 1997).

Los índices desarrollados para tal fin y que se pueden enmarcar dentro de esta categoría, fueron:

- Índice de bondad de ajuste parsimonioso (Parsimonious Goodness of Fit Index, PGFI) e Índice de ajuste normalizado parsimonioso (Parsimonious Normed Fit Index, PNFI).

Ambos están basados en el GFI y NFI respectivamente, ajustados por la pérdida de grados de libertad. Se debe destacar que no se han establecido umbrales para estos índices, ya que Mulaik et al. (1989) mostró que es posible obtener valores de estos índices parsimoniosos en torno a 0,5 y valores para otros índices de ajuste sobre 0,90.

En consecuencia, Hooper et al. (2008) recomienda el uso de estos índices de ajuste parsimoniosos junto con otras medidas debido a que ningún tipo de umbral ha sido establecido rigurosamente para los mismos, lo que implica una mayor dificultad en su interpretación.

Es cierto que, tal y como menciona Hooper et al. (2008), no se recomienda utilizar todos los índices que muestra el programa como resultados. No obstante, dado que no existe una regla universal sobre este tema, es necesario reportar una variedad de índices debido a que diferentes índices de ajuste reflejan diferentes aspectos del modelo (Crowley and Fan, 1997).

En conclusión, siguiendo las recomendaciones de diversos autores que han investigado sobre cuáles deberían ser los índices de ajuste que deben estar siempre presentes en cualquier estudio de SEM (Boomsma, 2000; Hayduk et al., 2007; Hu and Bentler, 1999; Kline, 2005; McDonal and Ho, 2002) y de los índices que aparecen en otros trabajos de calidad del servicio en el transporte público (De Oña et al., 2015b; Eboli and Mazzulla, 2007; Karlaftis et al., 2001; etc.), la Tabla 3 muestra los diferentes índices y los umbrales recomendados por otros autores, que han sido establecidos en esta tesis doctoral.

Tabla 3. Resumen de los Índices de Ajuste seleccionados para modelos SEM y sus umbrales recomendados.

Índice de Ajuste	Umbrales recomendados
<i>Índice de Ajuste Absolutos</i>	
χ^2	Valores significativos con p-valor menor de 0,05
RMSEA	Valores menores de 0,07
GFI	Valores mayores de 0,90
AGFI	Valores mayores de 0,90
RMR	Valores bajos
<i>Índices de Ajuste Incrementales</i>	
NFI	Valores mayores de 0,90
CFI	Valores mayores de 0,90
<i>Índices de Ajuste parsimoniosos</i>	
PGFI	Valores mayores de 0,90
PNFI	Valores mayores de 0,90

4. Estimación.

Una vez comprobada la existencia o no de solución, los parámetros del modelo SEM pueden ser estimados mediante el uso de diferentes métodos: Método de Máxima Verosimilitud (ML), Estimación por mínimos cuadrados ponderados y Estimación por mínimos cuadrados generalizados (de Oña et al., 2013a).

Todos ellos se basan en métodos de análisis de la covarianza cuyo objetivo principal es que se obtengan los parámetros del modelo que minimizan las diferencias entre la matriz de varianza-covarianza hipotética en el modelo y la de las observadas, mientras se respetan las diferentes restricciones o condicionantes del modelo.

El método ML es el más popular y el que ha sido utilizado en esta tesis doctoral, ya que proporciona estimaciones consistentes, eficientes y no sesgadas, con tamaños de muestra suficientemente grandes.

Este método debe ser elegido considerando las diferentes hipótesis sobre la probabilidad de distribución de cada caso, las propiedades de la escala de las variables, la complejidad en sí misma del SEM y el tamaño de la muestra (Golob, 2003). Sin embargo, se destaca que la violación de alguna de sus condiciones, no afecta a la capacidad del método para estimar de forma no sesgada los parámetros (Lara, 2014).

Definiendo la función de log-verosimilitud:

$$\text{Log } L = -1/2(N - 1)(\text{Log}(\Sigma(\theta)) + \text{tr}(S\Sigma(\theta)^{-1})) + c \quad (8)$$

Su maximización, es equivalente a minimizar la siguiente función:

$$F_{ML} = \text{Log}(\Sigma(\theta)) - \text{Log}(S) + \text{tr}(S\Sigma(\theta)^{-1}) - p \quad (9)$$

Donde L es la función de verosimilitud, N el tamaño de la muestra, S la matriz de covarianza de la muestra, $\Sigma(\theta)$ es la matriz de covarianzas del modelo y (θ) es el vector de los parámetros.

Con posterioridad a dicha estimación de los parámetros, se procede a evaluar el modelo mediante la comparación de los resultados obtenidos para los índices de ajuste seleccionados con sus umbrales recomendados. Esta una de las partes más importantes de los modelos SEM como técnica de naturaleza confirmativa (Hooper et al., 2008; Yuan, 2005).

5. Re-especificación.

En caso de que no se consigan buenos resultados de ajuste se deberá volver a re-especificar el modelo y empezar de nuevo estos pasos.

6. Comunicación de resultados.

Tal y como su nombre indica, una precisa descripción de los resultados de tal análisis debe ser llevada a cabo.

4.2. FASES DE LA INVESTIGACIÓN

Dos casos de estudio son considerados en esta tesis doctoral. Las fases de la investigación que se tienen que llevar a cabo para aplicar la metodología de dos etapas propuesta es esta tesis doctoral son las mismas. No obstante, aunque ambos casos de estudio se encuentran dentro del campo del transporte público, las bases de datos así como las variables consideradas son distintas.

En el primer caso, se aplica la metodología propuesta en una base de datos de calidad del transporte, obteniéndose un modelo general de calidad desde el punto de vista de los pasajeros.

En el segundo caso, se emplea esta metodología en una base de datos de actitudes hacia el transporte público para conseguir un modelo que explica las intenciones de comportamiento de los pasajeros.

4.2.1. CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA)

Este primer caso de estudio sirve para exponer y describir la metodología de dos etapas propuesta mediante su aplicación en una base de datos de calidad que proviene de una encuesta de satisfacción sobre el Metro de Sevilla (España), y tratará sobre cómo aprende, desarrolla y valida un modelo de calidad del servicio desarrollado directamente a partir de los datos.

Esta metodología está fundamentada en dos etapas principales. Sin embargo, debido a la naturaleza y forma de los datos disponibles, se hace necesario un pre-tratamiento de los mismos. Por tanto, en este caso de estudio se pueden identificar tres bloques claros y diferenciados con sus correspondientes fases. Un resumen de los mismos puede ser encontrado en la Tabla 4.

Bloque 1. Pre-tratamiento de los datos.

En esta fase se pre-procesan los datos para que puedan ser utilizados de manera eficiente por las diferentes técnicas de la metodología de dos etapas (RB y SEM), de tal forma que no se vea afectada la precisión del modelo conseguido, su complejidad o la probabilidad de cada estado, influenciadas por el número de categorías, atributos y/o dimensiones (Kashani and Mohaymany, 2011).

Por tanto, las fases que se engloban en este bloque son:

1. Re-categorización de las variables. Dado el amplio número y tipos de variables (p.e., de calidad, hábitos de viaje, socioeconómicas o de actitudes), así como de sus diferentes escalas de valoración (escala Likert 0-10 o 0-5, nominales, etc.) que provienen de la encuesta de satisfacción realizada en el Metro de Sevilla (España) en 2014, se hace necesario seleccionar las variables de calidad y homogeneizarlas re-categorizándolas. Los resultados son mostrados en la sección 5.1.1.1.
2. Realización del AFE. Se reduce el número de atributos considerados a priori, agrupándolos en dimensiones de calidad y comprobando la significancia de sus pesos factoriales. De esta forma, aquellos con pesos factoriales por debajo de 0,4, lo que significa que son poco significativos (Hair et al., 2010), son eliminados uno a uno, analizando posibles efectos en el resto de pesos factoriales del resto de atributos.

Tal análisis es explicado en la sección 4.1.1.1. y los resultados son mostrados en la sección 5.1.1.2.

3. Realización del AFC. Se realiza un análisis de validez de convergencia, discriminante y nomológica de las dimensiones extraídas en la fase anterior. En caso de que no cumpla algún atributo o dimensión, es eliminado uno a uno por posibles efectos en el resto de valores (AVE, Alfa de Cronbach, Fiabilidad del Constructo, etc.) de los otros atributos o dimensiones. Tal análisis es explicado en la sección 4.1.1.2. y los resultados son mostrados en la sección 5.1.1.3.

Bloque 2. Metodología de dos etapas. Etapa 1: Extracción de conocimiento de los datos.

En este bloque se desarrollan las fases ligadas a la aplicación de las RB con el fin de extraer conocimiento directamente de los datos (Etapa 1). Su aplicación, basada en el trabajo realizado en Cugnata et al. (2016) y una modificación del código en R suministrado en este trabajo, está realizada de forma que se consiga la RB más robusta.

Se debe destacar que en este caso se trabajará con los factores obtenidos en el análisis factorial (AFE y AFC) que representan a las dimensiones de calidad en la técnica de RB. El valor de estos factores puede ser calculado por infinidad de métodos, más o menos avanzados. No obstante, debido a su simplicidad y buenos resultados obtenidos en otros trabajos que han utilizado la base de datos de este estudio (p.e., de Oña et al., 2015b; 2016a; Machado-Leon, 2016), se ha utilizado la media ponderada para dicho cálculo utilizando la matriz de coeficientes de puntuación proporcionada durante el análisis factorial.

Las fases que se enmarcan dentro de este bloque son:

4. Aprendizaje de diferentes RB. Se aplica una amplia gama de algoritmos, cubriendo las diferentes clases existentes (Constrained-Based, Score-Based e Híbridos) con el fin de aprender diferentes RBs directamente de los datos. Como variables se utilizan las dimensiones extraídas del AFC. En base a la propia definición de calidad del servicio y a diversos autores y trabajos en la literatura (p.e., de Oña et al., 2013a; Eboli and Mazzulla, 2008; Redman et al., 2013; etc.), se impone la hipótesis de que todas las dimensiones tengan relación (influencia) directa con la dimensión de Calidad del Servicio. Dicha hipótesis será testeada posteriormente con la técnica SEM.

5. Selección de la RB más robusta. Se analiza cada red aprendida y se crea una tabla resumen de las relaciones extraídas (filas) en cada algoritmo (columnas), asignándoles una determinada puntuación: 0, 0,5 y 1, en el caso de no existencia de arco, existencia de arco no dirigido y existencia de arco dirigido, respectivamente. Una columna final mostrará la suma de dichas puntuaciones.

A continuación, se establece un umbral que mide la fuerza de ocurrencia de los arcos. Éste se establece en un valor del 70%, siguiendo las recomendaciones de Cugnata et al. (2016), lo cual indica que los arcos tienen presencia en un 70% o más de las RB aprendidas.

La RB que presente más relaciones con puntuaciones iguales o mayores que el umbral dado, es identificada como la más robusta. En caso de existir varias RBs con el mismo número de relaciones se elige aquella que tenga el menor ratio de casos mal clasificados utilizando la base de datos de test.

6. Análisis de la importancia de las dimensiones sobre la calidad global del servicio. Se generan 1.000 sub-base aleatorias, cada una con una proporción de datos de 1/10 sobre el total de la base de datos inicial, y se vuelve a aprender una RB para cada una de estas sub-bases, utilizando el mismo algoritmo que ha proporcionado la RB más robusta. A partir de estos resultados se calcula el porcentaje de ocurrencia de cada relación. Este porcentaje nos indicará la fuerza de los arcos. Finalmente, se calcula el DWI de cada dimensión sobre la calidad global del servicio, teniendo en cuenta el número de caminos que conectan cada dimensión con el nodo objetivo, la longitud de estos caminos y la fuerza de estos arcos.
7. Extracción de las diferentes relaciones entre las dimensiones de calidad. Obtenida la RB más robusta, se identifican en su DAG todas las relaciones existentes entre las diferentes dimensiones de calidad, las cuales, serán utilizadas durante el procedimiento de construcción del modelo estructural en SEM.

El proceso de las fases 4, 5, 6 y 7 es explicado en la sección 4.1.3. y sus resultados en la sección 5.1.2.

Bloque 3. Metodología de dos etapas. Etapa 2: Validación de las relaciones.

En este bloque las relaciones entre las diferentes dimensiones de calidad identificadas en las RB son modelizadas y validadas mediante la aplicación de SEM.

Por lo que se comprenden las siguientes fases:

8. Construcción del modelo en SEM. Se introducen como variables observadas los diferentes atributos considerados en la base de datos pre-tratada, que conforman el modelo de medición. Como variables latentes se utilizan las dimensiones de calidad extraídas del análisis factorial. Por último, las relaciones para el modelo estructural son extraídas directamente del DAG de la Red Bayesiana.
9. Depuración del modelo en SEM. Se van eliminando, una a una, las relaciones no significativas al nivel 0.05, partiendo desde la que mayor p-valor muestre, por posible efecto en el resto de valores de significancia. Por lo que se vuelve a simular y a analizar el modelo obtenido en cada paso.
10. Análisis de los resultados obtenidos. Se comprueban los diferentes parámetros de ajuste del modelo SEM, así como, en caso de que cumplan todos ellos, se comparan con los obtenidos en un modelo tradicional, donde solamente hay relación de las dimensiones con la calidad, no entre ellas, y con otros trabajos en la literatura. Finalmente, se analiza el modelo.

La técnica estadística SEM es explicada en la sección 4.1.4. y los resultados de las fases 8, 9 y 10 son mostrados en la sección 5.1.3.

Tabla 4. Esquema de las fases de investigación.

1º Bloque	<i>Pre-Tratamiento de los datos</i>		1º Re-categorización de las variables
			2º Realización del AFE
			3º Realización del AFC
2º Bloque	<i>Metodología de dos etapas</i>	<i>Extracción de Conocimiento (Etapa 1)</i>	4º Aprendizaje de diferentes RBs
			5º Selección de la RB más robusta
			6º Análisis de la importancia de las dimensiones
			7º Extracción de las relaciones entre dimensiones a partir del DAG
3º Bloque		<i>Validación de las Relaciones (Etapa 2)</i>	8º Construcción del modelo en SEM
			9º Depuración del modelo
			10º Análisis de los resultados obtenidos

4.2.2. CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO

El segundo caso de estudio sirve para mostrar que la metodología de dos etapas puede aplicarse en otros ámbitos de estudio. Para ello, se aplica la metodología en una base de datos de actitudes hacia el servicio de Metro de Sevilla (España), diferente a la del caso de estudio 1, y se obtiene un modelo sobre las intenciones de comportamiento de los pasajeros.

El proceso seguido en esta ocasión es análogo al del caso de estudio 1, siguiendo las mismas fases de investigación (Tabla 4). Los únicos aspectos diferenciadores radican en las variables observadas consideradas para el estudio y los constructos latentes obtenidos con el AFE y el AFC, que son de actitudes hacia el servicio.

4.3. DATOS

Para los dos casos de estudio anteriormente expuestos, se dispone de una base de datos sobre el servicio de Metro Ligero en Sevilla (España), que proviene del proyecto de investigación *“Mejora de la calidad del transporte público para fomentar la movilidad sostenible: Metro de Sevilla (G-GI3002/IDIK)”*, financiado por la Agencia de Obra Pública de la Junta de Andalucía.

Dichos datos fueron obtenidos a través de una encuesta cuyo principal propósito era el de obtener la percepción de los pasajeros sobre el servicio, sus actitudes hacia el transporte público, hábitos de viaje y otras características socioeconómicas.

En las siguientes secciones (4.3.1 y 4.3.2.) se detallan las principales características del servicio, la configuración y desarrollo de la encuesta, su distribución y la tecnología aplicada para la toma de datos. Así mismo, se realiza una breve descripción de los resultados de las diferentes partes en las que se divide la encuesta.

4.3.1. EL METRO LIGERO DE SEVILLA

El servicio de transporte público analizado en esta tesis doctoral corresponde al servicio de Metro Ligero en la ciudad de Sevilla (España). Se trata de una ciudad de gran tamaño en el sur de España que cuenta con un área metropolitana de 858.000 habitantes.

Metro de Sevilla Sociedad Concesionaria de la Junta de Andalucía S.A. se creó en 2003 para la redacción del proyecto, ejecución de las obras y explotación comercial de la Línea 1 interurbana del Metro de Sevilla. Dicho servicio quedó inaugurado de forma parcial el 2 de abril de 2009 y, desde el 23 de noviembre del mismo año, de forma completa, transportando actualmente alrededor de 14 millones de viajeros al año.

La línea 1 cuenta con un total de 18,1 kilómetros de longitud y un total de 22 estaciones. Atraviesa la ciudad transversalmente de oeste a este, conectando la capital con los municipios de San Juan de Aznalfarache, Mairena de Aljarafe, Dos Hermanas y Alcalá de Guadaíra. El servicio está compuesto por 21 coches de 31 metros de largo, en los que caben hasta 202 pasajeros (145 de pie y 57 sentados) y permiten acoplar módulos adicionales para aumentar su capacidad. La velocidad máxima de circulación es de 70 Km/h y consigue una frecuencia en hora punta de 4 minutos.

La configuración y estructura de la red del Metro fue diseñada de forma que se consiguiera conectar a los principales núcleos del área metropolitana de Sevilla, que representan a más del 60% de la población total de dicha área, con Sevilla capital, donde se generan y se localizan la mayoría de los viajes (p.e., centros administrativos, centros de salud, educativos y comerciales). Además, dicho diseño fue realizado siguiendo diferentes pautas para favorecer la intermodalidad con el resto de modos de transporte público de la ciudad. De esta forma, la Estación de San Bernardo se convierte en un gran intercambiador, donde confluyen múltiples modos de transporte de corta y media distancia (p.e., autobús urbano y metropolitano, servicio de bici, cercanías, taxi, etc.). Además, se debe destacar que cuenta con aparcamientos para vehículos privados.

Metro de Sevilla, a través del Instituto de Investigación y Marketing y Comunicación (IMC), conduce anualmente desde el año 2009 una encuesta de satisfacción al cliente, que les permite analizar la evolución en el tiempo de la prestación del servicio, además de constituir una herramienta de notable valor para la Agencia de Obra Pública de la Junta de Andalucía en su misión de supervisar y mejorar el servicio, de acuerdo a los parámetros de calidad establecidos.

Sin embargo, los datos usados para este trabajo de investigación provienen de una Encuesta ad-hoc elaborada, validada y llevada a cabo por el equipo de investigación TRYSE de la Universidad de Granada (del que formo parte), durante el año 2014.

4.3.2. ENCUESTA DEL PROYECTO “MEJORA DE LA CALIDAD DEL TRANSPORTE PÚBLICO PARA FOMENTAR LA MOVILIDAD SOSTENIBLE: METRO DE SEVILLA (G-GI3002/IDIK)”

La encuesta conducida durante 2014 se puede encontrar en el Anexo 1 de esta tesis doctoral.

Para la configuración de dicha encuesta, primero se realizó una revisión de la literatura y normativas existentes, así como de otras encuestas realizadas en servicios similares de otras ciudades (p.e., servicio de metro de Madrid, Metro de Sevilla, servicio de autobuses de Granada, etc.).

Este primer paso permitió la extracción de los aspectos más importantes y de las preguntas que se debían realizar, configurándose de este modo, un borrador de la encuesta.

Como segundo paso, atendiendo a su experiencia y conocimiento en la materia, se seleccionó un listado de 20 expertos a los que enviar el borrador de la encuesta, solicitando sus valoraciones y críticas.

Dichos expertos podemos agruparlos en dos grupos distintos:

- Profesionales que trabajan en los Consorcios de Transporte de Andalucía y en el Consorcio Regional de Transportes de Madrid (un total de 12 expertos).
- Profesionales que se dedican a la investigación y/o enseñanza dentro del área de Ingeniería del Transporte en diferentes centros de investigación y universidades (8 expertos).

En base a las opiniones y valoraciones de los expertos, se llevó a cabo diferentes modificaciones en el borrador de la encuesta, implementando la versión final de la encuesta piloto.

Esta encuesta piloto se realizó en el metro de Sevilla durante los días 29 y 30 de abril de 2014 para comprobar su validez. Estas encuestas fueron realizadas por un total de seis encuestadores, mediante entrevistas cara a cara, y se llevaron a cabo a lo largo de todo el día, por lo que abarcaron una gran diversidad horaria.

Como resultado se obtuvieron un total de 216 encuestas, a partir de cuyos datos, se depuraron y corrigieron diferentes aspectos, así como preguntas no significativas para el análisis, quedándose finalmente configurada la encuesta definitiva.

Para la divulgación de la encuesta, se decidió llevar a cabo una encuesta on-line, fomentando su participación mediante premios. La plataforma utilizada para la implementación de la encuesta fue el servidor Survey Monkey.

El conocimiento de dicha encuesta fue realizado a través de tarjetas (Figura 14), repartidas a los pasajeros en las estaciones del metro, y durante los periodos del 23 al 29 de Mayo de 2014 y del 3 al 6 de Junio del 2014. De esta forma se abarcó toda la franja horaria (horas valle y punta) y días de la semana en que este servicio se encuentra en operación.

Dichas tarjetas contenían diferente tipo de información: un enlace web, que era el proporcionado por Survey Monkey para realizar la encuesta on-line; un código, que permitía la identificación personal y singular del individuo, así como optar al sorteo de los premios; el último día hábil para la realización de la encuesta on-line (el 15 de Junio de 2014, disponiendo un total de 25 días hábiles desde el primer día de divulgación, y 7 días desde el último de divulgación); día del sorteo de los premios; información de la encuesta y contacto para ayuda; etc.

Figura 14. Tarjetas divulgativas de la Encuesta de Actitudes del Metro de Sevilla.

Se repartieron unas 19.683 tarjetas en el Metro de Sevilla (10.000 durante los días 23 al 29 de Mayo de 2014 y 9.683 del 3 al 6 de Junio de 2014). Esto ocasionó que se registraran unas 3.365 respuestas en el servidor Survey Monkey, lo que supone una ratio de respuesta del 17,09%. De dichas encuestas, 3.198 fueron consideradas como válidas para análisis posteriores.

La estructura de la encuesta consta de 4 secciones:

- Parte A: Actitudes hacia el servicio de Metro. Este apartado valora las distintas actitudes de los pasajeros del servicio de metro hacia el servicio o el transporte público (27 preguntas).
- Parte B: Su opinión sobre la calidad percibida en el uso del metro de Sevilla. Este apartado valora directamente los diferentes aspectos del servicio que se ponen a disposición del cliente en el metro de Sevilla, así como la valoración del servicio global (38 preguntas).

- Parte C: Su viaje. Este apartado valora los distintos hábitos de viaje adquiridos por el cliente (19 preguntas).
- Parte D: Sobre usted. Este apartado valora las características socioeconómicas de los distintos clientes que usan el metro de Sevilla (7 preguntas).

A continuación, se describen cada una de las partes:

Parte A. Actitudes hacia el Servicio de Metro.

Esta sección de la encuesta se centra en las actitudes de los pasajeros hacia el servicio de Metro. Las preguntas que componen esta parte valoran el grado de acuerdo o desacuerdo que tiene el pasajero con diferentes afirmaciones relacionadas con la satisfacción, los costes percibidos, los beneficios percibidos, la existencia de alternativas de transporte público atractivas, la actitud hacia el transporte público y las intenciones de comportamiento. Para ello se utilizó una escala Likert de 11 puntos siendo 0-Totalmente en desacuerdo, y 10-Totalmente de acuerdo, más la opción “No sabe/No contesta” (NS/NC). Esta parte estuvo formada por 27 preguntas. Se debe destacar que la última pregunta de dicha parte fue una valoración sobre la satisfacción global de los pasajeros con el funcionamiento del metro, utilizando para ello una escala Likert de 5 puntos (de 1- mínima satisfacción, a 5-máxima satisfacción).

Si a continuación se analizan los resultados de las preguntas de esta parte (Tabla 5), se puede observar que la desviación típica es mayor cuanto más baja es la valoración media (es decir, que los pasajeros no están de acuerdo con la afirmación), y viceversa.

En referencia a las medias, se observa que existe un amplio rango, siendo las afirmaciones más respaldadas la A20, A13 y A2 (9,00, 8,35 y 8,12, respectivamente), y con las que menos de acuerdo se muestran los pasajeros fueron la A7, A24 y A17 (3,53, 3,68 y 4,23, respectivamente). No obstante, se debe destacar que una baja valoración en A7 (3,53) es positivo, pues indica que los pasajeros creen que el tiempo de espera en los andenes es bajo.

En cuanto a las desviaciones típicas, se debe destacar que la A6, A16 y A8 son las que mayor diversidad de opiniones comprenden y menos acuerdo hay (sus desviaciones típicas son 3,47, 2,91 y 2,89, respectivamente). Mientras que en las que muestran mayor consenso son A20, A9 y A2 (1,57, 1,69 y 1,73, respectivamente).

En cuanto a las afirmaciones más respaldadas, éstas indican que los pasajeros se encuentran cómodos con el servicio de metro (A2), consideran que es un modo de transporte rápido (A13), y afirman que volverán a utilizar el servicio de metro (A20). Por otro lado, las preguntas A17 y A15 indican una falta de alternativas adecuadas al metro, lo que puede ser un problema ya que muchos pasajeros se sienten cautivos (no disponen de otras posibilidades para realizar el viaje).

Además, la valoración de la A24, muestra que los pasajeros no consideran que usar el servicio de metro varíe la forma de pensar de la gente sobre ellos. En referencia a su grado de satisfacción global (SQ2), se puede observar una nota de casi 4 (3,96) sobre 5, lo que demuestra un alto grado de satisfacción en general del pasajero con el servicio.

Tabla 5. Resumen de los resultados de la Encuesta de Actitudes: Parte A.

	Variable	Media	Desv. Típ.
A1	Me atrae viaje en metro	7,56	2,08
A2	Me siento cómodo viajando en metro	8,12	1,73
A3	El servicio de metro suministrado cumple con mis expectativas	7,40	2,19
A4	El precio me parece alto	7,26	2,65
A5	Creo que el billete supera ampliamente los costes del metro	6,25	2,65
A6	Las estaciones se encuentran alejadas de mi origen y/o destino	4,80	3,47
A7	Considero que el tiempo de espera en los andenes es excesivo	3,53	2,84
A8	Considero altos los costes de viajar en metro	5,38	2,89
A9	El servicio es bueno	7,92	1,69
A10	Me parece adecuada la relación calidad-precio	5,54	2,52
A11	El horario satisface mis necesidades	6,77	2,88
A12	La atención al cliente es buena	7,67	2,03
A13	Me gusta el metro porque me lleva rápido	8,35	1,76
A14	Prefiero el metro antes que otros Transportes Públicos	7,72	2,30
A15	Considero que existen buenas alternativas de Transporte Público al metro	4,73	2,67
A16	Da igual el modo de transporte empleado, mientras cumpla mis necesidades	5,87	2,91
A17	Pienso que otros modos de transporte ofrecen más ventajas que el metro	4,23	2,55
A18	Volveré a utilizar el metro con las mismas condiciones	7,99	2,02
A19	Suelo recomendar a otros que usen el metro	7,56	2,28
A20	Seguro que volveré a utilizar el servicio de metro	9,00	1,57
A21	Siento que usar el Transporte Público es acorde con mi estilo de vida	6,73	2,53
A22	No importa cuál sea el motivo de viaje, que siempre prefiero coger el Transporte Público	5,15	2,97
A23	Siento que viajando en Transporte Público colaboro a proteger el medio ambiente	7,79	2,26
A24	Creo que usar el Transporte Público influye en la opinión de la gente sobre mí	3,68	2,88
A25	Me gusta que la gente sepa que utilizo el Transporte Público	5,67	2,65
A26	Me gusta la gente que utiliza el Transporte Público	6,72	2,34
SQ2	Grado de Satisfacción Global	3,96	0,79

Nota: Desde la A1–A26 están medidas según una escala Likert de 0-10, mientras que SQ2 se mide en una escala Likert 1-5.

Parte B. Calidad Percibida en el Uso del Metro.

Esta parte de la encuesta se centró en las percepciones de los pasajeros sobre diferentes aspectos que conforman la calidad del servicio. Debido a esta razón, dicha parte está formada por la valoración de los pasajeros sobre un conjunto de atributos del servicio, que lo caracterizan y están relacionados con dimensiones de calidad. Así mismo, se invitó a realizar, como última pregunta de dicha parte, una valoración global del servicio. Las valoraciones de los atributos y la global se realizaron utilizando una escala Likert de 11 puntos (0- muy baja calidad y 10- muy alta calidad), y, adicionalmente, los encuestados tuvieron disponible la opción NS/NC. Dicha parte constó de 38 preguntas. La última pregunta hizo referencia a la valoración global del servicio, mientras que las otras 37 preguntas hicieron referencia a los atributos o características de la calidad del servicio.

Si se analizan los resultados obtenidos (Tabla 6), se observa que los atributos con mejor puntuación son el B22, B15 y B24 (8,55, 8,49 y 8,36, respectivamente), que además son los que tienen menor desviación típica (1,51, 1,54 y 1,52, respectivamente): Los atributos con menor puntuación son B30, B1 y B26 (2,69, 5,82 y 6,16, respectivamente), y presentan mayor desviación típica (2,94, 2,86 y 2,63, respectivamente). Este hecho muestra que los pasajeros muestran un mayor acuerdo en cuanto a decidir qué atributos son mejores, sin embargo, para clasificar los peores, se atisban ciertas discrepancias.

Si se analizan más en detalle, aspectos relacionados con la limpieza (B22 y B23) e iluminación (B24 y B25) (Equipamiento Tangible) son muy bien valorados (8,55, 8,11, 8,36 y 8,33, respectivamente), al igual que la puntualidad del servicio (B15 con un valor de 8,49). En cuanto a los peor valorados, el que obtiene una menor puntuación es el aspecto sobre la conexión a internet (B30 con un valor de 2,69), ya que, durante la realización de dicha encuesta, las partes soterradas del servicio no disponían de cobertura. Similarmente, las horas de funcionamiento (B1) y la disponibilidad de asientos en estaciones (B26), parecen ser dos aspectos problemáticos para el pasajero y, de ahí, que no tengan tan buena valoración como el resto (5,81 y 6,16, respectivamente).

Por último, la valoración global media del servicio es 7,61 sobre 10 con una desviación típica de 1,51, lo que significa una notable percepción del servicio en general, coherente con el hecho de que solo el atributo B30 es menor de 5 y el B1 menor de 6, poseyendo el resto una nota superior a este umbral. En consecuencia, se puede afirmar en líneas generales que el servicio es percibido con una buena calidad.

Tabla 6. Resumen de los resultados de la Encuesta de Actitudes: Parte B.

	Variables	Media	Desv. Típ.
B1	Horas de funcionamiento del servicio	5,82	2,86
B2	Número de servicios diarios	7,29	2,10
B3	Proximidad de las paradas a su origen y/o destino	6,30	2,87
B4	Regularidad del servicio	7,72	2,04
B5	Facilidad de conexión con otros medios de transporte	7,46	2,03
B6	Facilidad de acceso desde la calle a las estaciones y andenes	8,15	1,69
B7	Funcionamiento de ascensores, escaleras mecánicas, etc.	8,24	1,76
B8	Facilidad de acceso para las personas con movilidad reducida	7,96	1,95
B9	Funcionamiento de los lectores o validadoras de tickets para entrar y salir de la zona de andenes	7,86	2,02
B10	Facilidad de uso de las máquinas de venta de billetes	7,53	2,13
B11	Información actualizada, precisa y fiable a bordo	7,84	1,92
B12	Información actualizada, precisa y fiable en las estaciones	7,84	1,90
B13	Información disponible a través de otros sistemas de comunicación	6,37	2,42
B14	Paneles informativos y direccionales simples y claros en las estaciones	8,11	1,80
B15	Puntualidad	8,49	1,54
B16	Rapidez en la realización del viaje	8,14	1,77
B17	Tiempo de espera en el andén	7,31	2,09
B18	Apariencia del personal	7,99	1,76
B19	Amabilidad del personal	7,78	1,99
B20	Eficacia y rapidez del personal para informar, atender y resolver problemas cotidianos de los pasajeros	7,55	2,13
B21	Funcionamiento del "Servicio de atención al cliente"	7,05	2,14
B22	Limpieza de las estaciones	8,55	1,51
B23	Limpieza del vehículo	8,11	1,74
B24	Iluminación de las estaciones	8,36	1,52
B25	Iluminación del vehículo	8,33	1,53
B26	Disponibilidad de asientos en estaciones y andenes	6,16	2,63
B27	Nivel de comodidad a bordo de los vehículos	6,41	2,41
B28	Temperatura y ventilación de las estaciones y a bordo	7,29	2,22
B29	Conducción adecuada y no brusca	7,20	2,17
B30	Disponibilidad de cobertura telefónica y 3G en las estaciones y a bordo	2,69	2,94
B31	Sensación de seguridad frente a accidentes cuando viaja en el metro	7,28	2,20
B32	Sensación de seguridad frente a robos y violencia en las estaciones y a bordo	7,27	2,19
B33	Sensación de seguridad frente a resbalones, caídas, atrapamientos por puertas y escaleras mecánicas	7,07	2,29
B34	Señalización de las "Salidas de emergencia" y extintores	7,61	1,97
B35	Niveles de ruido en estaciones	6,49	2,41
B36	Niveles de ruido a bordo	6,42	2,38
B37	Niveles de vibraciones a bordo	6,30	2,34
SQ1	Valoración global del servicio	7,61	1,51

Parte C. Su Viaje.

Esta parte de la encuesta se centró en los hábitos de viaje (p.e., disponibilidad de vehículo privado para realizar el viaje, el motivo del desplazamiento, el motivo por el que utiliza el servicio de metro, la frecuencia de uso, el modo utilizado para desplazarse del origen a la parada de metro y desde la parada de metro a su destino, etc.). Dicha parte consta de 12 preguntas: seis son de respuesta simple, tres de respuesta múltiple, donde se pregunta por el motivo del uso del metro, y tres requieren que el encuestado indique un número positivo entero para obtener respectivamente el tiempo de viaje total, el desglosado en partes, y el número de viajes motorizados que realiza a la semana.

Por conveniencia, para una mejor caracterización de los pasajeros y lectura de los resultados de esta parte (Tabla 7), se analizan conjuntamente con los de la parte D.

Tabla 7. Resumen de los resultados de la Encuesta de Actitudes: Parte C.

<i>Características</i>	<i>Porcentajes</i>
C1 Razones de uso del metro	Rapidez (66,64%), Comodidad (50,00%), Falta de aparcamiento (32,18%), Frecuencia (28,89%), Congestión del tráfico (24,80%), No dispongo de vehículo (23,11%), Razones ecológicas (16,57%), No dispongo de carnet de conducir (14,54%), Es mi única alternativa (13,57%), Precio (10,16%), Otra razón (6,69%), Mi vehículo no lo puedo usar por algún motivo (6,04%).
C2 Razones de viaje	Estudios (38,90%), Trabajo (35,49%), Ocio (15,29%), Otros (10,32%).
C4 Modos de transporte complementarios desde el origen a la parada de autobús	A pie (62,50%), Coche (22,30%), Autobús (9,63%), Otros (5,57%).
C5 Duración del viaje desde el origen a la estación de metro	Menos de 10 min (60,48%), Más de 15 min (20,48%), Entre 10 y 15 min (19,04%).
C6 Modos de transporte complementarios desde la parada de autobús al destino	A pie (86,34%), Autobús (6,00%), Coche (4,16%), Otros (3,50%).
C7 Duración del viaje desde el origen a la estación de metro	Menos de 10 min (43,28%), Entre 10 y 15 min (30,08%), Más de 15 min (26,64%).
C8 Duración total del viaje desde el origen hasta el destino	Entre 25 y 40 min (39,02%), Menos de 25 min (34,87%), Más de 40 min (26,11%).
C9 Tipo de billete utilizado	Tarjeta del Consorcio (58,76%), Billete sencillo (9,69%), Bonometro (31,55%).
C10 Frecuencia de uso	Diariamente (52,03%), 3 o 4 veces por semana (17,95%), Esporádicamente (16,35%), 1 o 2 veces por semana (13,66%).
C12 Alternativas al Metro	Coche (33,18%), Autobús urbano (28,52%), Autobús interurbano (14,63%), Combinación de vehículos (8,01%), Bicicleta (7,72%), A pie (3,66%), Motocicleta (1,84%), Tranvía (1,13%), Otros (1,31%).

Parte D. Sobre Usted.

Finalmente, en esta parte se recogieron datos sobre las características socioeconómicas de los pasajeros (p.e., género, edad, nivel de estudios, etc.). Es decir, dichos datos proporcionarán información de los pasajeros sobre cómo se distribuyen en las franjas horarias, niveles de ingresos, código postal, etc. Consta de 9 preguntas, una de ellas con respuesta múltiple, dos requerían introducir un número entero positivo, y el resto son preguntas de respuesta simple.

Si atendemos a estas dos últimas partes (parte C y D) podemos caracterizar la muestra de la siguiente forma (Tabla 7 y 8): la mayoría de los que respondieron a dicha encuesta fueron mujeres (53,35%). En cuanto a la distribución de edad, la franja más representada fue la de 18-25 años (41,59%), seguida de las de 26-40 (28,89%) y 41-65 (25,58%). Las franjas de más jóvenes de 18 y mayores de 65 se encuentran infrarrepresentados en esta muestra (2,81% y 1,03%, respectivamente). Las principales razones para viajar fueron por estudios (38,90%) y trabajo (35,49%), seguido de los motivos ocio (15,29%) y otros (10,32%) con un menor porcentaje, y los tres principales motivos para utilizar el metro fueron su rapidez (66,64%), comodidad (50,00%) y a la falta de aparcamiento (32,18%). Igualmente serían resaltables los motivos de frecuencia, congestión del tráfico y la no disposición de vehículo (28,89%, 24,80% y 23,11%, respectivamente) por ser elegidas por alrededor el 25% del total de encuestados. La frecuencia de uso de esta muestra es diaria (52,10%), por lo que más de la mitad de los encuestados disponen de la Tarjeta del Consorcio (58,76%), y la mayoría tiene estudios Universitarios (48,47%) o de Bachiller/Ciclo de grado medio (41,96%). No obstante, se hace destacar que hay un pequeño grupo de pasajeros que se encontraban sin estudios o simplemente la educación obligatoria (9,01%). Similarmente, una parte notable de la muestra tiene unas rentas mensuales medias-bajas (<1.801€) y no hay mucha diferencia entre los que poseen vehículo privado para realizar el viaje y los que no (54,69% y 45,31% respectivamente), aunque el vehículo privado sería la alternativa favorita al metro (33,18%). Y como principal modo de transporte para llegar a la parada de bus y desde la parada final hasta el destino es a pie. El resto de modos tienen representación, pero no destacable.

Tabla 8. Resumen de los resultados de la Encuesta de Actitudes: Parte D.

<i>Características</i>		<i>Porcentajes</i>
D1	Género	Mujer (53,35%), Hombre (46,65%).
D2	Minusvalía	Ninguna (98,56%), Permanente (0,97%), Temporal (0,47%).
D3	Disponibilidad de...	Carnet de conducir (74,95%), Vehículo privado (54,69%), Bicicleta (43,18%), Ninguna de las anteriores (11,98%), Motocicleta (6,72%).
D4	Edad	18-25 (41,59%), 26-40 (28,89%), 41-65 (25,58%), <18 (2,81%), >65 (1,03%), NS/NC (0,1%).
D5	Estudios	Universitario (48,47%), Bachiller o grado medio (41,96%), Sin estudios o Educación Obligatoria (9,01%), NS/NC (0,56%).
D6	Situación Laboral	Empleado (43,68%), Estudiante (41,53%), Otro (12,23%), Jubilado (2,56%).
D7	Número de personas que habitan su hogar	3 o 4 personas (60,44%), Hasta 2 personas (24,04%), Más de 4 personas (15,52%).
D8	Renta	<1,201€ (28,80%), 1,201 -1,800€ (21,04%), NS/NC (17,70%), 1,801-2,400€ (16,48%), >2,401€ (15,98%).



**CAPÍTULO 5.
RESULTADOS Y
DISCUSIÓN**

Capítulo 5.

RESULTADOS Y DISCUSIÓN

En este capítulo se estudian, analizan y discuten los resultados de los dos casos de estudio propuestos. Cada uno de ellos presenta una aplicación de la metodología de dos etapas en diferentes tipologías de datos (datos de calidad del servicio, y de actitudes hacia el servicio). El fin último de utilizar dicha metodología de dos etapas, en ambos casos de estudio, es el aprendizaje automático de modelos directamente de los datos y su validación mediante la aplicación combinada de las RB (1º etapa) y el SEM (2º etapa).

La obtención de los diferentes modelos, su análisis y la discusión de sus resultados, justifican la utilidad real y potencial de esta metodología en el campo del transporte público, así como su posible aplicación a otros ámbitos, quedando de esta forma reconocida como una vía de extracción y validación automática de conocimiento.

5.1. CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA)

En esta sección se realiza un estudio sobre la calidad del servicio en el Metro de Sevilla (España) mediante la aplicación de la metodología de dos etapas propuesta. Como resultado de la primera etapa, es decir, de la aplicación de las RB, se extraen directamente de los datos las relaciones entre las diferentes dimensiones, que conforman un determinado modelo sobre la calidad. Posteriormente, en la segunda etapa, se desarrolla dicho modelo en SEM y se evalúa de forma que los resultados obtenidos permiten depurar y confirmar dicho modelo.

Similarmente, los resultados de ajuste de dicho modelo sobre calidad del servicio, conseguidos mediante la metodología de dos etapas, son comparados con los resultados del modelo básico sobre calidad, en que solamente se consideran relaciones directas entre las dimensiones que conforman la calidad (p.e., Equipamiento Tangible, Seguridad, Información, etc.) y la dimensión de calidad global.

Debido a la gran cantidad de atributos utilizados en este caso de estudio (39 atributos) y a sus respectivas categorías de las escalas de medida empleadas (11 en todos los casos, salvo en el atributo SQ2, que son 5), previamente a la aplicación de la metodología de dos etapas, es necesario llevar a cabo un pre-procesamiento de los datos (re-categorización, AFE y AFC), de forma que la precisión de las técnicas y la coherencia de los resultados no se vean influenciados por este hecho.

5.1.1. PRE-PROCESAMIENTO DE LOS DATOS

De entre el amplio abanico de datos disponibles de la encuesta, como ya se ha comentado, se han utilizado 39 variables que describen atributos relacionados con diferentes aspectos de la calidad del servicio de metro (p.e., Disponibilidad, Accesibilidad, Seguridad, etc.) o con la calidad global del mismo (p.e., valoración global del servicio o satisfacción).

- 37 atributos de la calidad del servicio (Parte B: B1-B37).
- Atributo “Valoración global del servicio” (Parte B: SQ1).
- Atributo “Grado de Satisfacción Global” (Parte A: SQ2).

Se debe señalar que los atributos que proceden de la Parte B de la encuesta (B1-B37 y SQ1) siguen una escala de 11 puntos Likert (0- valoración muy baja de calidad y 10- valoración muy alta de calidad). La cuestión procedente de la Parte A (SQ2) fue preguntada siguiendo una escala de 5 puntos Likert (1- mínima satisfacción y 5- máxima satisfacción).

El pre-procesamiento seguido, se puede dividir en tres fases que ocupan cada una las siguientes secciones.

5.1.1.1. RE-CATEGORIZACIÓN DE LOS ATRIBUTOS

En esta primera fase, los diferentes atributos que conformaron la base de datos para este caso de estudio, son re-categorizados. Dicha fase es justificada debido a que la precisión de los modelos de RB, su complejidad y la probabilidad de cada clase, se encuentran influenciados por el número de categorías de las diferentes variables (Kashani and Mohaymany, 2011).

La re-categorización consistió en una transformación de una escala de 11 puntos Likert a una de 5 puntos Likert. Se realizó dicha re-categorización de los valores de las observaciones de la siguiente forma:

- De [10-9] se les re-categorizó con el valor 5.
- De (9-7] se les re-categorizó con el valor 4.
- De (7-5] se les re-categorizó con el valor 3.
- De (5-3] se les re-categorizó con el valor 2.
- De (3-0] se les re-categorizó con el valor 1.

Como resultado principal de esta fase, todos los atributos seleccionados para este caso de estudio contaban con las mismas 5 categorías.

5.1.1.2. ANÁLISIS FACTORIAL EXPLORATORIO (AFE)

En la siguiente fase, con el objetivo de reducir el número de atributos, se extrajeron dimensiones latentes subyacentes en la base de datos, que conforman la calidad del servicio de este estudio. Para ello, se realizó un AFE sobre los 37 atributos de calidad utilizando el software *SPSS v.20*. Del mismo modo, realizando las mismas consideraciones y dada la necesidad de establecer una dimensión latente que representara la Calidad del Servicio (CS), otro AFE fue realizado únicamente sobre las variables “Grado de Satisfacción” (SQ2) y “Valoración Global del Servicio” (SQ1).

Dado que ambos análisis fueron realizados de forma independiente, se procederá igual en cuanto a exposición y discusión de sus resultados.

Se debe destacar que, con el objetivo de depurar y mejorar las dimensiones obtenidas, la decisión de eliminar (o no) atributos no significativos en cada una de ellas se realiza en base a los resultados obtenidos conjuntamente con el AFE, AFC y el Alfa de Cronbach. En base a lo cual, debido a que se quiere conseguir una mejor explicación de por qué se eliminan ciertos atributos, se explicarán dichas decisiones en detalle en la siguiente sección (5.1.1.3.), destacándose solamente aquí aquellos atributos cuyos resultados recomiendan su eliminación.

En relación al primer AFE sobre los 37 atributos de calidad, se comenzó evaluando la adecuación de los datos a la realización de dicho análisis. Considerando que el número de observaciones válidas que tenía la base de datos era 3.198, se calculó la ratio de observaciones por atributo obteniéndose un valor de 86. Dado que debe ser mayor de 5 observaciones por atributo, se comprobó que cumplía. Además, el Test de Bartlett (Bartlett, 1950) obtuvo un valor de 58.171, significativo al nivel 0,001, mientras que la Medida de Adecuación de la Muestra (MSA), mostró un valor aceptable de 0,959 (superior al ideal de 0,8). Por lo tanto, se pudo establecer que existían correlaciones significativas entre los atributos y que los datos eran adecuados para realizar el AFE.

Fue posible extraer 8 dimensiones latentes cuyos valores propios eran mayores que 1, de acuerdo a la regla Kaiser-Guttman, y conjuntamente explicaban una proporción bastante adecuada de varianza (62,84%) (Hair et al., 2010).

La Tabla 9 muestra las dimensiones resultantes, así como los pesos factoriales y los coeficientes de las puntuaciones factoriales de cada atributo. Utilizando la media ponderada de la matriz de pesos factoriales y de coeficiente de puntuación se consiguen los valores de los factores que obtendrían si hubieran sido observados directamente. Una descripción en detalle del AFE viene recogida en la sección 4.1.1.1.

Se puede destacar que el único atributo propuesto para su eliminación es B3 “Proximidad de las paradas a su origen y/o destino”, cuyo peso factorial es inferior a 0,4 (0,381). No obstante, se consideró en el AFC antes de tomar una decisión sobre su eliminación.

Para el segundo AFE sobre SQ1 y SQ2, el ratio de observaciones por atributo fue de 1.599 (>5), su test de Bartlett fue significativo al nivel 0,001, con un valor de 2.111, y su MSA fue de 0,505, que aunque bajo, cumple con la condición de que sea mayor de 0,5. Por tanto, se pudo concluir que la base de datos era adecuada para realizar el AFE, que extrajo una única dimensión que explicaba una varianza del 84,77%.

Se debe destacar que tanto SQ1 como SQ2 mostraron pesos factoriales aceptables (Tabla 9).

Agrupados los atributos y extraídos sus pesos factoriales, se interpretaron las diferentes dimensiones, desde el marco teórico de este caso de estudio, como sigue: “Calidad del Servicio” (CS), “Equipamiento Tangible” (Eq_Tan), “Accesibilidad” (Acces), “Disponibilidad” (Disp), “Atención al Cliente”(At_Cl), “Seguridad” (Seg), “Información” (Info), “Medioambiente” (Med_Am) y “Espacio Individual” (Esp_Ind).

Se puede destacar que, inicialmente los atributos estaban agrupados de esta misma forma en la encuesta y bajo las mismas dimensiones obtenidas por el AFE, salvo el caso de la dimensión Espacio Individual, que no había sido contemplada, y la dimensión Disponibilidad, que es la unión de la original dimensión Disponibilidad en la encuesta y la desaparecida dimensión Tiempo.

Tabla 9. Resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 1.

Dimensiones		Peso Factorial	Coef. P.F.
Equipamiento Tangible (Eq_Tan)			
B22	Limpieza de las estaciones	0,714	0,315
B24	Iluminación de las estaciones	0,692	0,291
B25	Iluminación del vehículo	0,691	0,290
B23	Limpieza del vehículo	0,662	0,282
B28	Temperatura y ventilación de las estaciones y a bordo	0,542	0,254
B29	Conducción adecuada y no brusca	0,440	0,153
Accesibilidad (Acces)			
B8	Facilidad de acceso para las personas con movilidad reducida	0,695	0,343
B6	Facilidad de acceso desde la calle a las estaciones y andenes	0,691	0,323
B7	Funcionamiento de ascensores, escaleras mecánicas, etc.	0,667	0,320
B9	Funcionamiento de los lectores de Tickets	0,629	0,303
B10	Facilidad de uso de las máquinas de venta de billetes	0,615	0,300
B5	Facilidad de conexión con otros medios de transporte	0,556	0,248
Disponibilidad (Disp)			
B2	Frecuencia del servicio	0,755	0,385
B17	Tiempo de espera en el andén	0,711	0,346
B1	Horas de funcionamiento del servicio	0,581	0,285
B16	Rapidez en la realización del viaje	0,572	0,246
B4	Regularidad del servicio	0,542	0,255
B15	Puntualidad	0,490	0,184
B3	Proximidad de las paradas a su origen y/o destino	0,381	0,177
Atención al Cliente (At_Cl)			
B20	Eficacia y rapidez del personal para informar, atender, resolver problemas cotidianos	0,799	0,406
B19	Amabilidad del personal	0,791	0,399
B21	Funcionamiento del "Servicio de atención al cliente"	0,724	0,351
B18	Apariencia del personal	0,696	0,327

Seguridad (Seg)			
B32	Sensación de seguridad frente a robos y violencia en las estaciones y a bordo	0,741	0,459
B31	Sensación de seguridad frente a accidentes cuando viaja en el metro	0,734	0,431
B33	Sensación de seguridad frente a resbalones, caídas, atrapamientos por puertas y escaleras	0,707	0,412
B34	Señalización de las "Salidas de Emergencia" y extintores	0,588	0,312
Información (Info)			
B11	Información actualizada, precisa y fiable a bordo	0,727	0,444
B12	Información actualizada, precisa y fiable en las estaciones	0,723	0,431
B14	Paneles informativos y direccionales simples y claros en las estaciones	0,643	0,360
B13	Información disponible a través de otros sistemas de comunicación	0,596	0,361
Medioambiente (Med_Am)			
B36	Niveles de ruido a bordo	0,866	0,423
B37	Niveles de vibraciones a bordo	0,843	0,404
B35	Niveles de ruido en estaciones	0,814	0,396
Espacio Individual (Esp_Ind)			
B30	Disponibilidad de cobertura telefónica y 3G en las estaciones y a bordo	0,723	0,441
B26	Disponibilidad de asientos en estaciones y andenes	0,634	0,359
B27	Nivel de comodidad a bordo de los vehículos	0,587	0,315
Calidad del Servicio (CS)			
SQ1	Valoración global del Servicio	0,921	0,543
SQ2	Grado de satisfacción con el Servicio	0,921	0,543

5.1.1.3. ANÁLISIS FACTORIAL CONFIRMATORIO (AFC)

Como última fase del pre-procesamiento, se realizó un AFC que permitió testear las diferentes agrupaciones de los atributos en las 8 dimensiones que conforman la calidad extraídas del AFE. Tres atributos (B1, B3 y B30) fueron eliminados durante el AFC siguiendo un proceso gradual basado en los valores obtenidos, que mostraban bajo nivel de significancia en relación a la dimensión latente que estaban midiendo.

Para realizar dicho AFC se definió un modelo de medición en SEM que estaba basado en los resultados del AFE realizado, y que contaba con 8 variables exógenas: Equipamiento Tangible, Accesibilidad, Disponibilidad, Atención al Cliente, Seguridad, Información, Medio Ambiente y Espacio Individual. De esta forma se pudieron verificar los atributos y las dimensiones que resultaron del AFE en base a los resultados del AFC.

Se observó que el atributo B30 “Disponibilidad de cobertura telefónica y 3G en las estaciones y a bordo” debía ser excluido del análisis, ya que se testeó la fiabilidad de la dimensión que lo englobaba mediante el Alfa de Cronbach. Esta dimensión (Espacio Individual) obtenía un valor del Alfa de Cronbach igual a 0,631 cuando se consideraba este atributo, y alcanzaba un valor de 0,699 cuando se eliminaba. Además, en el AFC su peso de carga estandarizado fue inferior a 0,5, justificando de esta forma su eliminación. El atributo B3 “Proximidad de las paradas a su origen y/o destino” fue eliminado, ya que su peso factorial es inferior a 0,4, su peso de carga estandarizado fue inferior a 0,5 en el AFC, y el Alfa de Cronbach de su dimensión mejoraba cuando no se consideraba (de 0,773 a 0,791).

Casos especiales fueron B1 “Horas de funcionamiento del servicio” y B29 “Conducción adecuada y no brusca”. Por un lado, B1 obtuvo un peso factorial por encima de 0,4 en el AFE. Sin embargo, durante el AFC su peso de carga estandarizado bajaba considerablemente de 0,5, y si no era considerado el Alfa de Cronbach de la dimensión donde se agrupaba pasaba de 0,791 a 0,803, por lo que se determinó su eliminación. Se hace notar que, por conveniencia de la redacción y discusión del proceso, la Tabla 10 ya proporciona los atributos finalmente considerados.

Tras depurar los atributos anteriores, el atributo B29 mostró un peso factorial por debajo de 0,4 y, por tanto, se recomendaba su eliminación. No obstante, fue aceptada su consideración, ya que durante el AFC mostró un rol importante para su dimensión (Equipamiento Tangible), su eliminación producía un descenso brusco del Alfa de Cronbach de la misma (0,834 a 0,808), y su peso factorial en el AFE no es muy lejano al del umbral establecido (0,340).

La Tabla 10 muestra los pesos factoriales y los coeficientes de las puntuaciones factoriales de todos los atributos considerados, así como las dimensiones que constituyen.

Tabla 10. Resultados del Análisis Factorial Exploratorio tras la depuración y mejora de las dimensiones, caso de estudio 1.

Dimensiones	Peso Factorial	Coef. P.F
Equipamiento Tangible (Eq_Tan)		
B22 Limpieza de las estaciones	0,732	0,370
B24 Iluminación de las estaciones	0,707	0,341
B25 Iluminación del vehículo	0,689	0,325
B23 Limpieza del vehículo	0,643	0,298
B28 Temperatura y ventilación de las estaciones y a bordo	0,418	0,160
B29 Conducción adecuada y no brusca	0,340	0,085

Accesibilidad (Acces)			
B8	Facilidad de acceso para las personas con movilidad reducida	0,696	0,348
B6	Facilidad de acceso desde la calle a las estaciones y andenes	0,691	0,327
B7	Funcionamiento de ascensores, escaleras mecánicas, etc.	0,671	0,326
B9	Funcionamiento de los lectores de Tickets	0,635	0,311
B10	Facilidad de uso de las máquinas de venta de billetes	0,613	0,301
B5	Facilidad de conexión con otros medios de transporte	0,548	0,242
Disponibilidad (Disp)			
B17	Tiempo de espera en el andén	0,769	0,434
B2	Frecuencia del servicio	0,728	0,414
B16	Rapidez en la realización del viaje	0,632	0,315
B4	Regularidad del servicio	0,565	0,296
B15	Puntualidad	0,554	0,243
Atención al Cliente (At_Cl)			
B20	Eficacia y rapidez del personal para informar, atender, resolver problemas cotidianos	0,806	0,412
B19	Amabilidad del personal	0,794	0,400
B21	Funcionamiento del "Servicio de atención al cliente"	0,733	0,359
B18	Apariencia del personal	0,693	0,323
Seguridad (Seg)			
B32	Sensación de seguridad frente a robos y violencia en las estaciones y a bordo	0,740	0,458
B31	Sensación de seguridad frente a accidentes cuando viaja en el metro	0,735	0,431
B33	Sensación de seguridad frente a resbalones, caídas, atrapamientos por puertas y escaleras	0,715	0,416
B34	Señalización de las "Salidas de Emergencia" y extintores	0,592	0,315
Información (Info)			
B11	Información actualizada, precisa y fiable a bordo	0,733	0,451
B12	Información actualizada, precisa y fiable en las estaciones	0,733	0,442
B14	Paneles informativos y direccionales simples y claros en las estaciones	0,648	0,368
B13	Información disponible a través de otros sistemas de comunicación	0,596	0,356
Medioambiente (Med_Am)			
B36	Niveles de ruido a bordo	0,868	0,425
B37	Niveles de vibraciones a bordo	0,847	0,407
B35	Niveles de ruido en estaciones	0,819	0,401
Espacio Individual (Esp_Ind)			
B26	Disponibilidad de asientos en estaciones y andenes	0,739	0,514
B27	Nivel de comodidad a bordo de los vehículos	0,728	0,486
Calidad del Servicio (CS)			
SQ1	Valoración global del Servicio	0,921	0,543
SQ2	Grado de satisfacción con el Servicio	0,921	0,543

Por su parte, los resultados obtenidos sobre la validez de convergencia y la validez discriminante de estas dimensiones y sus atributos en el AFC, son mostrados en la Tabla 11.

Tabla 11. Resultados del Análisis Factorial Confirmatorio, caso de estudio 1.

Dimensiones		P.C.N.	P.C.E.	S.E.
Accesibilidad (Acces)				
<i>(CR: 0,844; AVE: 0,475; Alfa de Cronbach: 0,838)</i>				
B5	Facilidad de conexión con otros medios de transporte	1,000	0,625	-
B6	Facilidad de acceso desde la calle a las estaciones y andenes	1,033	0,768	0,03
B7	Funcionamiento de ascensores, escaleras mecánicas, etc.	0,996	0,718	0,03
B8	Facilidad de acceso para las personas con movilidad reducida	1,089	0,717	0,033
B9	Funcionamiento de los lectores de Tickets	1,053	0,658	0,034
B10	Facilidad de uso de las máquinas de venta de billetes	1,073	0,638	0,036
Atención al Cliente (At_Cl)				
<i>(CR: 0,895; AVE: 0,680; Alfa de Cronbach: 0,892)</i>				
B18	Apariencia del personal	1,000	0,795	-
B19	Amabilidad del personal	1,199	0,852	0,023
B20	Eficacia y rapidez del personal para informar, atender, resolver problemas cotidianos	1,291	0,852	0,024
B21	Funcionamiento del "Servicio de atención al cliente"	1,209	0,798	0,025
Disponibilidad (Disp)				
<i>(CR: 0,810; AVE: 0,463; Alfa de Cronbach: 0,803)</i>				
B2	Frecuencia del servicio	1,000	0,632	-
B4	Regularidad del servicio	0,851	0,566	0,031
B15	Puntualidad	0,842	0,739	0,025
B16	Rapidez en la realización del viaje	0,971	0,735	0,029
B17	Tiempo de espera en el andén	1,110	0,713	0,034
Equipamiento Tangible (Eq_Tan)				
<i>(CR: 0,857; AVE: 0,504; Alfa de Cronbach: 0,834)</i>				
B22	Limpieza de las estaciones	1,000	0,748	-
B23	Limpieza del vehículo	1,164	0,744	0,028
B24	Iluminación de las estaciones	1,061	0,79	0,024
B25	Iluminación del vehículo	1,081	0,789	0,024
B28	Temperatura y ventilación de las estaciones y a bordo	1,100	0,554	0,036
B29	Conducción adecuada y no brusca	1,166	0,598	0,035
Espacio Individual (Esp_Ind)				
<i>(CR: 0,698; AVE: 0,537; Alfa de Cronbach: 0,694)</i>				
B26	Disponibilidad de asientos en estaciones y andenes	1,000	0,682	-
B27	Nivel de comodidad a bordo de los vehículos	1,057	0,78	0,037

Información (Info)				
<i>(CR: 0,825; AVE: 0,544; Alfa de Cronbach: 0,809)</i>				
B11	Información actualizada, precisa y fiable a bordo	1,000	0,784	-
B12	Información actualizada, precisa y fiable en las estaciones	1,019	0,8	0,022
B13	Información disponible a través de otros sistemas de comunicación	0,916	0,591	0,028
B14	Paneles informativos y direccionales simples y claros en las estaciones	0,904	0,757	0,021
Medio Ambiente (Med_Am)				
<i>(CR: 0,871; AVE: 0,693; Alfa de Cronbach: 0,869)</i>				
B35	Niveles de ruido en estaciones	1,000	0,777	-
B36	Niveles de ruido a bordo	1,094	0,867	0,022
B37	Niveles de vibraciones a bordo	1,065	0,851	0,022
Seguridad (Seg)				
<i>(CR: 0,829; AVE: 0,549; Alfa de Cronbach: 0,828)</i>				
B31	Sensación de seguridad frente a accidentes cuando viaja en el metro	1,000	0,763	-
B32	Sensación de seguridad frente a robos y violencia en las estaciones y a bordo	0,931	0,716	0,024
B33	Sensación de seguridad frente a resbalones, caídas, atrapamientos por puertas y escaleras	1,018	0,743	0,025
B34	Señalización de las "Salidas de Emergencia" y extintores	0,867	0,74	0,021
Índices de ajuste del modelo				
	χ^2	3.701,96		
	Grados de libertad	499		
	GFI	0,927		
	AGFI	0,913		
	RMR	0,037		
	RMSEA	0,045		
	NFI	0,934		
	CFI	0,942		
	PGFI	0,778		
	PNFI	0,831		

Nota: CR: Fiabilidad del constructo; AVE: Media de la varianza extraída; S.E.: Error Estándar de la covarianza estimada (en inglés, Standard Error, SE); P.C.N.: Peso de carga no estandarizado; P.C.E.: Peso de carga estandarizados, todos fueron significantes ($p < 0,001$). Las dimensiones están en negrita.

Previo paso a analizar las dimensiones y sus resultados en el AFC, se evaluó el modelo SEM de acuerdo a los umbrales recomendados para los diferentes índices de ajuste (sección 4.1.4.). Se puede observar que cumple satisfactoriamente dichos umbrales (ver Tabla 11). La aplicación en esta tesis doctoral de la técnica estadística SEM fue llevada a cabo con el software *Amos Graphics v.22*.

X^2 es significativo al nivel 0,05, lo que indica que la magnitud de discrepancia entre la muestra y la matriz de covarianza estimada es insignificante. CFI es superior a 0,9 (0,942) y RMSEA es menor de 0,07 (0,045). Por tanto, todos ellos justifican que el modelo de medición establecido en SEM para realizar el AFC es adecuado.

Se confirma la validez de convergencia, ya que todos los pesos de carga estandarizados (P.C.) fueron superiores a 0,5, estando la mayoría de ellos sobre el umbral ideal de 0,7. En relación al AVE, se observa que todas las dimensiones obtuvieron un valor superior a 0,5, excepto las dimensiones Accesibilidad y Disponibilidad que, aunque obtuvieron valores inferiores a 0,5, son muy cercanos a dicho umbral recomendado (0,475 y 0,463, respectivamente). Además, para dichas dimensiones, sus Alfas de Cronbach son mayores de 0,791, sus valores de CR muy superiores a 0,7, y todos los pesos de carga estandarizados de sus atributos son superiores a 0,5, por lo que fueron considerados aceptables. El resto de dimensiones mostraron valores adecuados de Alfa de Cronbach y de CR superiores o iguales al umbral recomendado de 0,7.

En cuanto a la validez discriminante, el AVE de cada dimensión fue comparada con el cuadrado de la inter-correlación del resto de dimensiones (Tabla 12). Las siguientes inter-correlaciones al cuadrado entre Accesibilidad con Disponibilidad (0,50), con Información (0,54) y con Equipamiento Tangible (0,54) superaron al AVE de Acces (0,48); y las inter-correlaciones al cuadrado de Disp con Información (0,49) y con Equipamiento Tangible (0,48) superaron al AVE de Disponibilidad (0,46).

Debido a esta razón, se llevaron a cabo Test de diferencias Chi-Cuadrado para comprobar la validez discriminante de las dimensiones anteriormente mencionadas, dos a dos. Los resultados de dichos test mostraron evidencias de que el modelo sin correlación ajustaba mejor a los datos, y por tanto, indicaba que cada dimensión medía un aspecto diferente. Sus resultados pueden ser consultados en el Anexo 2. En consecuencia, quedó comprobada la validez discriminante de las dimensiones.

Tabla 12. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 1.

	Acces	At_Cl	Disp	Eq_Tan	Esp_Ind	Info	Med_Am	Seg
Acces	0,48	0,43	0,50	0,54	0,30	0,54	0,15	0,43
At_Cl		0,68	0,34	0,45	0,26	0,38	0,12	0,37
Disp			0,46	0,48	0,30	0,49	0,18	0,35
Eq_Tan				0,50	0,39	0,45	0,18	0,49
Esp_Ind					0,54	0,29	0,17	0,43
Info						0,54	0,13	0,39
Med_Am							0,69	0,24
Seg								0,55

Nota: La diagonal de la matriz muestra los valores del AVE de cada dimensión y las celdas sombreadas son valores de inter-correlación al cuadrado entre dos variables superiores a los valores AVE.

Por último, su validez nomológica fue comprobada ya que todas las dimensiones y los atributos agrupados tenían un sentido teórico dentro del marco de este caso de estudio.

Antes de la aplicación de la metodología de dos etapas, se procedió a calcular los valores de las dimensiones a partir de los atributos que las representan. Para ello, se hace uso de los coeficientes de puntuación factorial de la Tabla 10, que proporciona el software *SPSS v.20*. Las dimensiones con sus valores calculados, establecidos también en las mismas 5 categorías que los atributos, fueron utilizados en las RB para aprender cómo se relacionaban entre ellas y así construir el modelo de calidad.

5.1.2. 1ª ETAPA. EXTRACCIÓN DE CONOCIMIENTO DE LOS DATOS

En este paso, se calibran las RB para encontrar las posibles relaciones entre las dimensiones que conforman la calidad del servicio de este caso de estudio. Para ello, las 9 dimensiones extraídas en el pre-procesamiento son usadas para aprender la RB.

Este paso está basado en el proceso explicado en el trabajo de Cugnata et al. (2016) y descrito en la sección 4.1.3.2. Dicho proceso tiene como objetivo conseguir la estructura de la RB más robusta. Del mismo modo, los algoritmos y funciones de puntuación utilizados en esta tesis doctoral han sido desarrollados a través del *lenguaje estadístico R*, con la librería *bnlearn* de Scutari (2010) y Nagarajan et al. (2013) y la librería *gRain* de Hojsgaard (2012).

En primer lugar, 17 algoritmos diferentes fueron aplicados sobre los valores de las 9 dimensiones de calidad (8 dimensiones que agrupan las principales características del servicio y la dimensión que representa la calidad global del servicio) para el aprendizaje de las diferentes RBs. Se destaca que dichos valores ya habían sido calculados tras el AFE y AFC.

En este caso de estudio, antes de realizar la calibración de las redes, se impuso una condición o hipótesis sobre la relación existente entre las diferentes dimensiones que conforman la calidad del servicio (Accesibilidad, Atención al Cliente, Disponibilidad, Equipamiento Tangible, Espacio Individual, Información, Medio Ambiente y Seguridad) con la dimensión que representa la calidad global del servicio en sí misma (CS), lo que repercute directamente en el aprendizaje de las RBs. Dicha condición indica que la Calidad del Servicio está siempre directamente relacionada con todas las dimensiones que conforman esta calidad. Aunque en la literatura pueden encontrarse ejemplos donde la influencia de estas dimensiones sobre la Calidad del Servicio es indirecta (de Oña et al., 2017), la mayoría de las investigaciones muestran una influencia directa entre ellas (p.e., de Oña et al., 2013a; Eboli and Mazzulla, 2008; Redman et al., 2013; etc.). Su validez se comprueba a través de los resultados del SEM (2ª Etapa).

Una vez simuladas las RB, se reportó y puntuó la ocurrencia de los arcos (Anexo 3), lo que representa la fuerza de cada relación con un valor entre 0 y el número total de algoritmos aplicados. Las puntuaciones se mantuvieron en 1 si era un arco dirigido, 0 si no existía arco, y 0,5, en caso de que el algoritmo no haya podido establecer la dirección de dicho arco, pero sí la existencia de una relación entre dos dimensiones. Se estableció en 11 el valor del umbral para considerar un arco con alta ocurrencia, lo que corresponde a que tiene presencia en más de 2/3 del total de RB aprendidas, de acuerdo con lo propuesto por Cugnata et al. (2016).

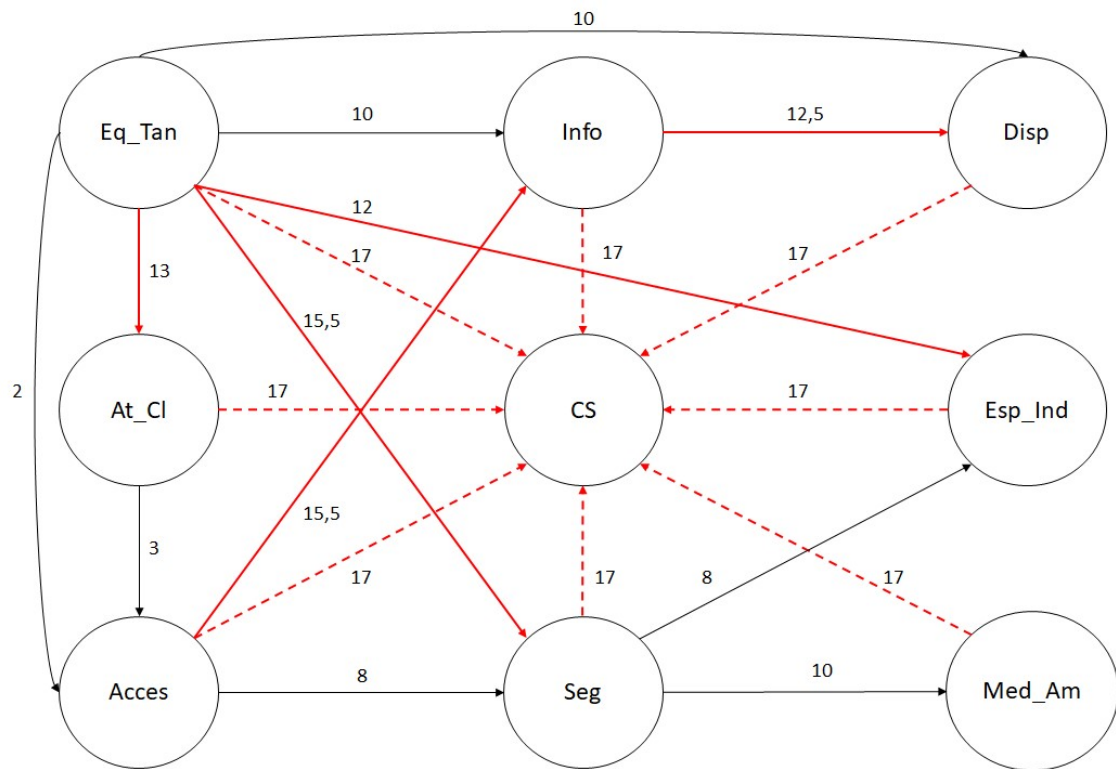
Dado que hubo varias RB con el mismo número de arcos con valores mayores al umbral establecido (la producida por los algoritmos Hill Climbing y TABU, con 13 arcos), se recurrió a utilizar la ratio de missclassification (Tabla 13).

Tabla 13. Ratios de missclassification de las RB con el mismo número de arcos con valores mayores al umbral establecido para el caso de estudio 1.

Algoritmo	Métrica	Ratio de missclassification
<i>Hill Climbing</i>	<i>AIC</i>	3,4595
	<i>BIC</i>	3,3731
	<i>Bde</i>	3,4286
	<i>mBde</i>	3,1321
	<i>K2</i>	3,3735
	<i>LogLik</i>	3,3522
<i>TABU</i>	<i>AIC</i>	3,2416
	<i>BIC</i>	3,6970
	<i>Bde</i>	3,5652
	<i>mBde</i>	2,9819
	<i>K2</i>	3,3128
	<i>LogLik</i>	3,3514

En base a este criterio, la RB conseguida con el algoritmo TABU y métrica mBDe fue la seleccionada (Figura 15).

Los valores de cada arco fueron reportados encima de los mismos, estando en rojo aquellos cuya puntuación es ≥ 11 . Destacan los arcos que conforman las relaciones entre Equipamiento Tangible con Seguridad, Accesibilidad con Información, Equipamiento Tangible con Atención al Cliente, Información con Disponibilidad y Equipamiento Tangible con Espacio Individual.



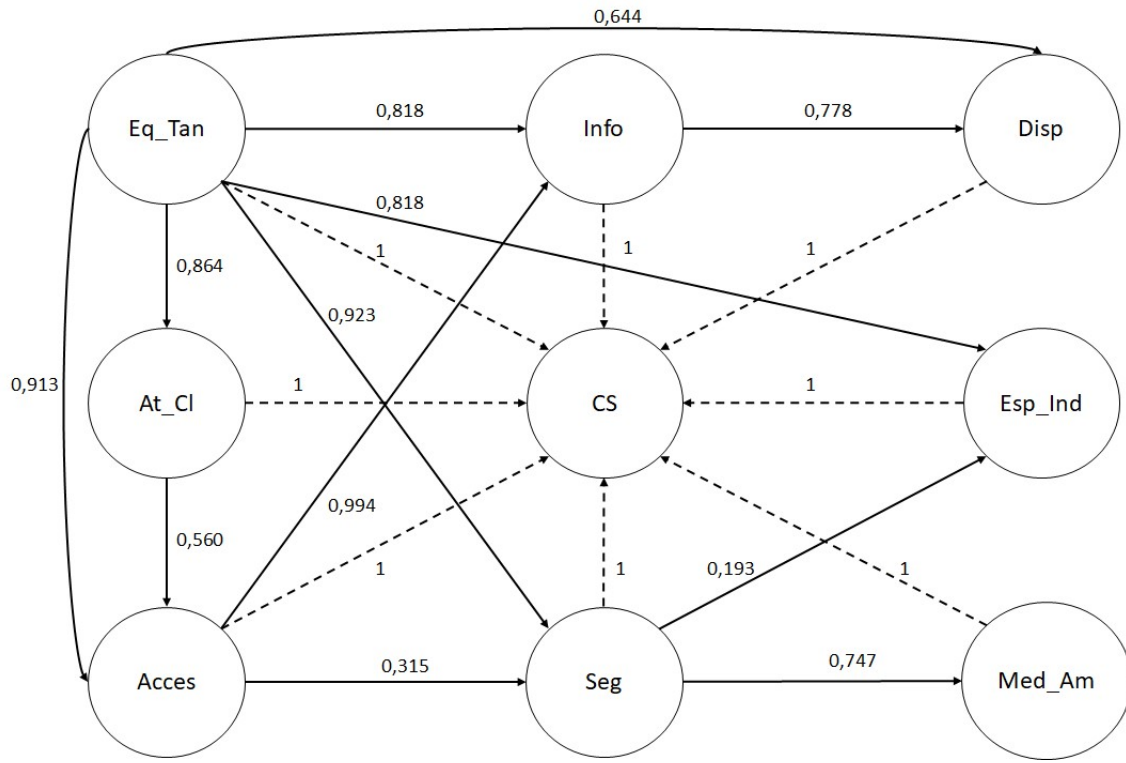
Nota: Las relaciones impuestas se representan con flechas discontinuas.

Figura 15. RB más robusta del caso de estudio 1.

Finalmente, se llevó a cabo el análisis de la importancia de las dimensiones mediante el proceso de re-muestreo iterativo sobre la base de datos inicial y el cálculo de la métrica DWI.

La Figura 16 muestra la proporción de ocurrencia de cada arco en el proceso iterativo de replicación (bootstrap). Se considera que los arcos que obtienen proporciones de ocurrencia cercanas a 1 pueden ser considerados más significantes, de acuerdo con Cugnata et al. (2016).

Se puede observar que casi todos los arcos rojos tuvieron proporciones elevadas y cercanas a 1, lo que evidencia que dicho modelo y sus relaciones eran robustos. Tras el cálculo de la métrica DWI se puede determinar el ranking de importancia de las dimensiones sobre la valoración global del servicio.



Nota: Las relaciones impuestas se representan con flechas discontinuas.

Figura 16. RB más robusta del caso de estudio 1, con la proporción de ocurrencia de cada arco en el proceso de re-muestreo iterativo.

En la Tabla 14 se puede observar que Equipamiento Tangible es la dimensión más importante, seguida de Accesibilidad. En los últimos puestos se encuentran Disponibilidad, Espacio Individual y Medio Ambiente.

Tabla 14. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 1.

Dimensión	Importancia	Ranking
Equipamiento Tangible (Eq_Tan)	6,97	1
Accesibilidad (Acces)	2,56	2
Información (Info)	1,61	3
Seguridad (Seg)	1,60	4
Atención al Cliente (At_Cl)	1,53	5
Disponibilidad (Disp)	1	6
Espacio Individual (Esp_Ind)	1	6
Medio Ambiente (Med_Am)	1	6

5.1.3. 2ª ETAPA: VALIDACIÓN DE LAS RELACIONES

En este paso las relaciones entre las dimensiones, que han sido identificadas en las RBs, son modeladas y validadas utilizando la técnica SEM. Para este propósito los parámetros de bondad de ajuste del modelo estructural fueron analizados según lo expuesto en la sección 4.1.4. de esta tesis doctoral. Además, todas las relaciones en el modelo son chequeadas de forma que sean significativas a un nivel de confianza de 0,01. En caso contrario, son eliminadas.

El modelo SEM fue construido considerando las relaciones entre las dimensiones extraídas de la RB más robusta lograda en el paso anterior. Las 8 dimensiones de calidad (Accesibilidad, Atención al Cliente, Disponibilidad, Equipamiento Tangible, Espacio Individual, Información, Medio Ambiente y Seguridad) representan las diferentes variables latentes del modelo, de forma que se dejó de considerar sus valores previamente calculados, y fueron explicadas con los 34 atributos de calidad (del B1 al B37, sin considerar B1, B3 y B30), de acuerdo a las agrupaciones obtenidas en el AFE y confirmadas por el AFC. De forma similar, la dimensión que representa la calidad global del servicio (CS) fue otra variable latente del modelo y explicada con las variables observadas “Calidad Global del Servicio” (SQ1) y “Grado de Satisfacción” (SQ2). Por lo que el modelo estuvo compuesto por 9 variables latentes y 36 variables observadas o atributos.

Tras una primera modelización, se obtuvo que las relaciones en el modelo estructural de Información y Accesibilidad con Calidad del Servicio no eran significativas al nivel de confianza de 0,01. Por tanto, se procedió a su eliminación y se re-calculó de nuevo el modelo, obteniéndose finalmente unos resultados satisfactorios de significancia de todas las relaciones. Dichos resultados son reportados en las Tablas 15, y 16 y en la Figura 17.

Tabla 15. Pesos de carga y Errores del modelo de medición para el SEM del caso de estudio 1.

Relaciones del modelo de medición			No Est.	S.E.	Est.	p
B5	<---	Acces	1	-	0,625	
B6	<---	Acces	1,021	0,03	0,76	***
B7	<---	Acces	0,986	0,03	0,712	***
B8	<---	Acces	1,088	0,033	0,717	***
B9	<---	Acces	1,054	0,034	0,659	***
B10	<---	Acces	1,076	0,035	0,64	***
B18	<---	At_Cl	1	-	0,795	
B19	<---	At_Cl	1,2	0,023	0,853	***
B20	<---	At_Cl	1,293	0,024	0,852	***
B21	<---	At_Cl	1,207	0,025	0,797	***
SQ1	<---	CS	1	-	0,891	
SQ2	<---	CS	0,872	0,02	0,773	***
B2	<---	Disp	1	-	0,639	
B4	<---	Disp	0,835	0,031	0,561	***
B15	<---	Disp	0,827	0,025	0,733	***
B16	<---	Disp	0,966	0,029	0,739	***
B17	<---	Disp	1,101	0,033	0,715	***
B22	<---	Eq_Tan	1	-	0,74	
B23	<---	Eq_Tan	1,171	0,028	0,741	***
B24	<---	Eq_Tan	1,06	0,024	0,782	***
B25	<---	Eq_Tan	1,082	0,025	0,782	***
B28	<---	Eq_Tan	1,118	0,037	0,558	***
B29	<---	Eq_Tan	1,192	0,036	0,605	***
B26	<---	Esp_Ind	1	-	0,673	
B27	<---	Esp_Ind	1,087	0,036	0,791	***
B11	<---	Info	1	-	0,782	
B12	<---	Info	1,019	0,022	0,797	***
B13	<---	Info	0,914	0,028	0,588	***
B14	<---	Info	0,906	0,021	0,757	***
B35	<---	Med_Am	1	-	0,775	
B36	<---	Med_Am	1,095	0,022	0,866	***
B37	<---	Med_Am	1,069	0,022	0,853	***
B31	<---	Seg	1	-	0,764	
B32	<---	Seg	0,925	0,024	0,711	***
B33	<---	Seg	1,013	0,025	0,741	***
B34	<---	Seg	0,858	0,021	0,733	***

Nota: No Est., No estandarizados; Est. Estandarizados; S.E., Error estándar de la covarianza estimada; ***, $p < 0,001$.

Tabla 16. Pesos de carga y Errores del modelo estructural para el SEM del caso de estudio 1.

Relaciones Estructurales			No Est.	S.E.	Est.	P
Info	<---	Acces	0,661	0,036	0,557	***
Seg	<---	Acces	0,402	0,037	0,305	***
Acces	<---	At_Cl	0,278	0,022	0,306	***
CS	<---	At_Cl	0,139	0,023	0,135	***
CS	<---	Disp	0,530	0,030	0,507	***
Acces	<---	Eq_Tan	0,608	0,031	0,534	***
At_Cl	<---	Eq_Tan	0,862	0,027	0,690	***
CS	<---	Eq_Tan	-0,213	0,050	-0,165	***
Disp	<---	Eq_Tan	0,525	0,033	0,427	***
Esp_Ind	<---	Eq_Tan	0,537	0,052	0,340	***
Info	<---	Eq_Tan	0,371	0,037	0,275	***
Seg	<---	Eq_Tan	0,757	0,044	0,506	***
CS	<---	Esp_Ind	0,284	0,024	0,349	***
Disp	<---	Info	0,381	0,024	0,418	***
CS	<---	Med_Am	0,058	0,014	0,075	***
CS	<---	Seg	0,078	0,028	0,090	0,006
Esp_Ind	<---	Seg	0,437	0,036	0,414	***
Med_Am	<---	Seg	0,581	0,024	0,518	***

Nota: No Est., No estandarizados; Est. Estandarizados; S.E., Error estándar de la covarianza estimada; ***, $p < 0,001$.

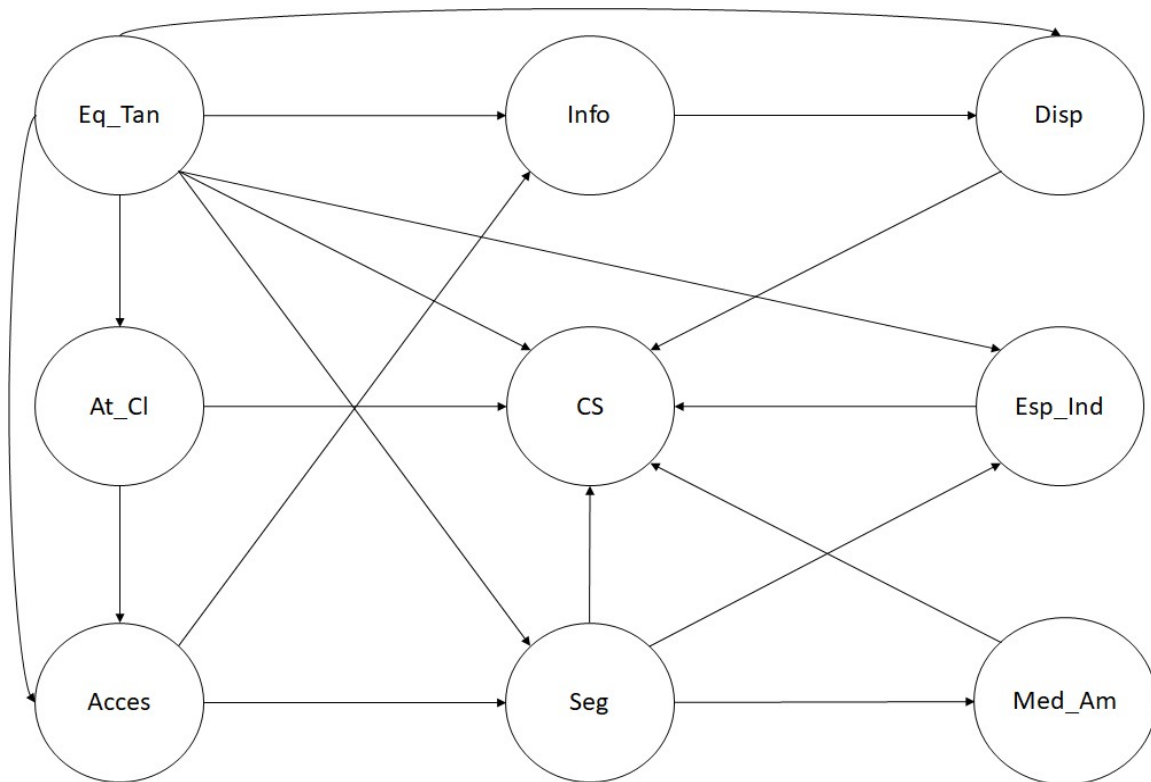


Figura 17. Modelo Estructural del SEM del caso estudio 1, con relaciones significativas.

Se destaca que en la Tabla 16 la relación Seguridad con Calidad del Servicio no es significativa al nivel 0,001, pero sí al 0,01, por tanto, se mantiene en el modelo.

El modelo de medición está formado por las relaciones entre las variables latentes y las observadas, es decir, por las relaciones entre los 34 atributos de la calidad (variables con terminología B_i) y los atributos SQ1 y SQ2, con las 8 dimensiones que conforman la calidad y la dimensión que representa la calidad global del servicio, establecidas previamente (Tabla 10). Se observa que todas las relaciones fueron significativas al nivel de significancia de 0,001, y todos los pesos de carga estandarizados considerablemente altos (Est.>0,5). Para el caso de la dimensión latente que representa la Calidad del Servicio (CS) es posible destacar que los dos atributos observados (SQ1 y SQ2) explicaron de forma muy satisfactoria esta dimensión, ya que mostraron unos pesos de carga estandarizados de 0,891 y 0,773, respectivamente.

Para el caso del modelo estructural, que está formado por las diferentes relaciones entre las dimensiones latentes (Tabla 16), se observa que todas ellas, tras depurar las dos relaciones entre dimensiones no significativas (Información con Calidad del Servicio y Accesibilidad con Calidad del Servicio, respectivamente), fueron significativas al nivel 0,001, salvo la relación de Seguridad con Calidad del Servicio, que fue significativa al nivel 0,01. Incluso fue posible contrastar que las dos relaciones, cuya proporción en la RB estaba por debajo de 0,5 (Accesibilidad con Seguridad y Seguridad con Espacio Individual), en el caso del SEM fueron significativas. Además, se observa que casi todas tenían efecto positivo sobre Calidad del Servicio, salvo la relación de Equipamiento Tangible, que tenía sentido negativo. Dicho hecho encuentra su justificación analizando los efectos totales de esta variable sobre Calidad del Servicio.

La Tabla 17 muestra los diferentes efectos (p.e., totales, directos e indirectos) de las diferentes dimensiones que conforman la calidad sobre la dimensión Calidad del Servicio. Se observa que todas las dimensiones, salvo Accesibilidad e Información, tienen efecto directo sobre Calidad del Servicio, es decir, que influyen directamente sobre la valoración de la calidad global del servicio. Por su parte, las dimensiones Accesibilidad, Atención al Cliente, Equipamiento Tangible, Información y Seguridad poseen efecto indirecto sobre Calidad del Servicio, lo que significa que van a influir sobre la valoración global de la calidad a través de otras dimensiones.

Para el caso de Equipamiento Tangible, también se observa que la dimensión ejerce un pequeño efecto directo negativo sobre Calidad del Servicio (-0,165), que se considera poco relevante si se compara con su efecto indirecto positivo (0,776). Por tanto, se puede afirmar que es necesario la consideración de los efectos totales, como resultado de sumar los efectos directos e indirectos, a la hora de analizar y validar modelos sobre calidad del servicio mediante SEM.

Pruebas adicionales de ello es que, si solo hubieran sido considerados los efectos directos en este estudio, dimensiones como Info no tendrían efecto sobre Calidad del Servicio y, por supuesto, Equipamiento Tangible tendría un efecto negativo. Esto sería contrario a las hipótesis de este estudio sobre las relaciones de las dimensiones con Calidad del Servicio y no sería posible validar este modelo.

Tabla 17. Efectos directos, indirectos y totales estandarizados sobre la dimensión “Calidad del Servicio” (CS) del caso de estudio 1.

	Acces	At_Cl	Disp	Eq_Tan	Esp_Ind	Info	Med_Am	Seg
Efectos Directos	-	0,135	0,507	-0,165	0,349	-	0,075	0,090
Efectos Indirectos	0,202	0,062	0,000	0,776	0,000	0,212	0,000	0,184
Efectos totales	0,202	0,197	0,507	0,611	0,349	0,212	0,075	0,274

Si se analiza en más detalle la Tabla 17, las dimensiones Equipamiento Tangible y Disponibilidad son las que mayor efecto total ejercen sobre la Calidad del Servicio y, por tanto, que mayor influencia tienen en la percepción general de los pasajeros sobre dicha calidad global del servicio. Caso opuesto es la dimensión Medio Ambiente, que apenas tiene influencia sobre Calidad del Servicio. Este hecho puede ser explicado mediante la consideración de que el Metro de Sevilla se trata de un transporte que utiliza vehículos eléctricos modernos, cuya única tipología de polución posible es la acústica (p.e., ruido y vibraciones) y, en consecuencia, se puede considerar bastante respetuoso con el medio ambiente. Se debe hacer mención que estos resultados son absolutamente coherentes con otros análisis que han utilizado la misma base de datos (p.e., De Oña et al., 2015b; 2016c; Machado-Leon et al., 2016; etc.).

Si se comparan estos resultados con el ranking de importancias establecido con el DWI en la RB, también se identificó el Equipamiento Tangible como la dimensión que mayor influencia ejercía sobre la valoración global del servicio. Sin embargo, esta métrica obtenida directamente de la RB no coincide con el SEM en identificar el orden de importancias del resto de dimensiones, siendo esta una debilidad de estas RB (por lo que se rechaza la hipótesis H3 de esta tesis doctoral).

En cuanto a los índices de ajuste de este modelo, se puede observar un resumen de los más importantes en la Tabla 18. Adicionalmente, en dicha tabla se han recogido los mismos parámetros de ajuste, pero de un modelo básico de calidad del servicio, que sirve como modelo de referencia para este estudio.

Para la construcción de dicho modelo básico, se han utilizado las mismas dimensiones que conforman la calidad (Accesibilidad, Atención al Cliente, Disponibilidad, Equipamiento Tangible, Espacio Individual, Información, Medio Ambiente y Seguridad), la misma que representa a la calidad global del servicio (Calidad del Servicio), y los mismos atributos que los utilizados para el modelo conseguido con la metodología de dos etapas. La diferencia radica en que solamente se han establecido relaciones directas entre las dimensiones que conforman la calidad y la dimensión que representa la calidad global del servicio.

Tabla 18. Medidas de bondad de ajuste del modelo SEM del caso de estudio 1.

	Modelo conseguido con la metodología de dos etapas	Modelo básico sin relaciones entre las dimensiones que conforman la calidad
Índices de ajuste		
Tamaño de la muestra	3.198	3.198
χ^2	4.273,87	15.112,00
Grados de libertad	576	586
Nivel de Probabilidad	0.000	0.000
Nº de parámetros distintos estimados.	90	80
Índices de ajuste absolutos		
GFI	0,921	0,687
AGFI	0,909	0,643
RMSEA	0,045	0,088
RMR	0,044	0,281
Índices de ajuste incrementales		
NFI	0,929	0,750
CFI	0,938	0,757
Índices de ajuste de parsimonia		
PGFI	0,796	0,604
PNFI	0,850	0,698

Una comparación entre ambos modelos pone de manifiesto que el modelo conseguido a través de la metodología de dos etapas mejora de manera muy notable al modelo básico utilizado como referencia. Como ejemplo, podemos destacar el valor del RMSEA, que para el modelo sin relaciones entre las dimensiones de calidad es 0,088, no cumpliendo con el umbral recomendado (inferior a 0,05) y siendo notablemente superior al modelo conseguido con la metodología (0,045). Lo mismo para el GFI, cuyo valor para el modelo sin relaciones es 0,687, por lo que no cumpliría tampoco con el umbral recomendado (superior a 0,90), y su diferencia con el del modelo conseguido con la metodología es bastante importante (0,921).

Como principales consecuencias cabe resaltar que:

- se verifican las hipótesis H1 y H2, que establecen la existencia de relaciones entre las dimensiones que conforman la calidad del servicio, siendo necesario su descubrimiento y establecimiento para una adecuada modelización de la calidad del servicio.
- se destaca la idoneidad de aplicar esta metodología para estudiar y analizar modelos directamente de una base de datos, como son el caso de los modelos de calidad del servicio en el transporte público.

Si se analizan los parámetros de ajuste del modelo conseguido con la nueva metodología, se puede observar la idoneidad de estos valores de ajuste. El valor X^2 indica que la magnitud de discrepancia entre la muestra y la matriz de covarianza ajustada es insignificante al nivel de 0,05, que es un umbral sugerido por varios autores (p.e., Golob, 2003; Hair et al., 2010; Hooper et al., 2008; Mulaik et al., 1989, etc.).

En relación a los índices de ajuste absolutos, como GFI (0,921) y AGFI (0,909), se ve que tiene valores superiores al valor recomendado (0,90), y que también, el valor del índice RMSEA (0,045) es inferior a 0,08, valor que significa un buen ajuste. Por último, el valor del índice RMR (0,044) es inferior a 0,05, cumpliendo de manera satisfactoria con su recomendación.

Los índices de ajuste incrementales obtenidos presentan valores similares (NFI=0,929 y CFI=0,938) y cercanos a 1, indicando un buen ajuste. Por su parte, los valores de los índices de ajuste de parsimonia, PGFI (0,796) y PNFI (0,850), respectivamente, son consistentes con las recomendaciones establecidas por Mulaik et al. (1989).

Como conclusión principal basada en los parámetros de ajuste del modelo conseguido en este estudio, el nivel de bondad de ajuste del modelo se puede considerar como muy satisfactorio, mejorando incluso otros estudios en la literatura que han realizado estudios solo con SEM, en este campo y en otros (p.e., De Oña et al., 2013a; Dimitrov, 2006; Philip et al., 2003; etc.).

5.1.4. DISCUSIÓN SOBRE EL MODELO DE CALIDAD DEL SERVICIO

La Figura 15 muestra las relaciones que son aprendidas a partir de los datos directamente usando la RB. Por su parte, la Figura 17 muestra las relaciones significativas definitivas del modelo sobre la calidad del servicio que han sido validadas por la técnica SEM. Se puede comprobar que han sido eliminadas las relaciones de Accesibilidad e Información con Calidad del Servicio que fueron aprendidas con la RB, pero se mostraron no significativas al nivel 0,01 en el modelado con SEM.

Si se analizan las relaciones que finalmente son obtenidas, se destacan una serie de ellas que son relevantes para el modelo, y que de no haber aplicado esta metodología de dos etapas, no habrían sido identificadas:

- Equipamiento Tangible muestra conexión con el resto de las dimensiones salvo con Medio Ambiente. Dicha dimensión cubre todos los aspectos relacionados con iluminación, limpieza (tanto en el interior de los vehículos como en las estaciones), temperatura, confort, ventilación, o incluso conducción apropiada y no brusca. Por lo que se puede concluir que la forma en que sea percibida esta dimensión por parte de los pasajeros va a tener una influencia directa en la percepción que tengan sobre el resto de las dimensiones. A saber:
 - a) Un nivel adecuado de temperatura, limpieza y/o ventilación permitirá a los pasajeros sentir un mayor nivel de comodidad en el vehículo (relación con Espacio Individual).
 - b) Una buena iluminación de los ascensores, escaleras mecánicas, entradas a las estaciones, plataformas, etc. ayuda a un fácil acceso por parte de los pasajeros, sobre todo desde el exterior. Además, la limpieza y mantenimiento de dichas instalaciones ayuda al buen funcionamiento de los elementos mecánicos de las mismas. Por lo que la relación con la dimensión Accesibilidad está justificada.
 - c) En el ámbito de la seguridad, se sabe que la iluminación es un factor que influye directamente sobre el número de atracos o agresiones, o sobre la sensación del pasajero de sentirse seguro sobre estos aspectos. Del mismo modo, una buena iluminación también es un hecho que ayuda a evitar y prevenir caídas, atrapamientos, resbalones, etc. Todos estos hechos justifican la relación de Equipamiento Tangible con Seguridad.
 - d) Además, no solo la iluminación ayuda en temas de seguridad, sino también en el ámbito informativo, al poder localizar las diferentes instalaciones si se utilizan las señales luminosas adecuadas, así como permite identificar tabloneros de anuncios con mucha más facilidad (relación con Información).
 - e) Una mala conducción del metro y/o brusca (p.e., frenazos, interrupciones, largas paradas, exceso de velocidad, etc.) va a influir en la forma en que se presta dicho servicio desde el punto de vista de la frecuencia, cumplimiento de horarios, puntualidad, etc. Aspectos que son cubiertos por la dimensión Disponibilidad y que, en consecuencia, demuestran la relación existente entre estas dos dimensiones.

- f) Por último, la dimensión Equipamiento Tangible cubre aspectos básicos del servicio, como son la iluminación, limpieza, temperatura, ventilación, etc., que se dan por hecho su buen funcionamiento a la hora de prestar cualquier servicio público. Un no adecuado funcionamiento de los mismos va a producir que los pasajeros los presenten como quejas o peticiones de mejora, cuya resolución va a tener impacto directo en la percepción de los pasajeros sobre la efectividad y rapidez de los empleados para tratarlos. En consecuencia, se puede soportar su relación con Atención al Cliente.
- La dimensión Accesibilidad se encuentra conectada con Información y Seguridad:
 - a) Por una parte, la relación con Información puede ser explicada a través del hecho de que un mal funcionamiento de las máquinas expendedoras o validadoras de tickets, producirá una sensación de confusión ante el desconocimiento de si es porque no funcionan adecuadamente o porque los pasajeros no saben utilizarlas adecuadamente, ambos hechos causados por falta de información.
 - b) Por otra parte, la relación con Seguridad se comprueba considerando el hecho de que si se proporciona un acceso fácil, tanto físicamente como a través de instalaciones, evitará problemas de caídas, resbalones, sensación de inseguridad frente a robos, atracos, etc.
 - Informaicón está relacionada con Disponibilidad. Una información actualizada, precisa y fiable, permite a los pasajeros la organización y planificación de sus viajes con cierta antelación. Este hecho produce reducción de tiempos de espera en la plataforma, las aglomeraciones, y la mejora de la sensación de puntualidad y regularidad del servicio.
 - La dimensión Atención al Cliente está relacionada con Accesibilidad, debido a que posibles problemas relacionados con el acceso (p.e., averías de los ascensores, máquinas de venta de billetes, escaleras mecánicas, etc.) o información de conexión con otros medios de transporte son resueltos o provistos a través de los diferentes trabajadores del metro, y según su amabilidad, su apariencia o incluso su eficacia y rapidez, van a influir en la percepción sobre la facilidad de acceso y buen funcionamiento de las instalaciones.

Por último, Seguridad se encuentra conectada con Espacio Individual y Medio Ambiente:

- a) Una buena sensación de seguridad puede permitir una mejor percepción del espacio individual disponible. Situaciones de inseguridad producirán que la sensación de poco espacio sea mayor.
- b) Y con Medio Ambiente se explica por el hecho de que una mala sensación de seguridad de los pasajeros, hacen que perciban un mayor nivel de vibración y ruido. Es decir, si los pasajeros se sienten inseguros frente a resbalones o caídas a bordo, aspectos cubiertos por la dimensión Seguridad, este hecho va a influir en que las vibraciones que ocurren en el metro sean percibidas en una mayor medida, y por consiguiente, se tenga una peor percepción de la dimensión Medio Ambiente.

En resumen, estos ejemplos muestran que las relaciones extraídas de la RB son apropiadas y en coherencia con la realidad, mostrando un modelo de calidad del servicio de metro de Sevilla preciso. Esta conclusión apoya, por un lado, lo apropiado que es usar las RB para encontrar potenciales relaciones entre las variables de un estudio, de forma automática a partir de los datos. Y por otro lado, la utilidad de SEM como validador de estas relaciones, chequeándolas y confirmando en última instancia el marco del modelo de calidad del servicio, extraído de las RB. Debido a estas razones, se afirma que esta metodología de dos pasos puede ser útil y relevante para desarrollar nuevas teorías, identificar nuevas relaciones entre variables de manera automática sin necesidad de hipótesis, y confirmar su validez en éste y otros ámbitos de conocimiento.

5.2. CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO

Para el segundo caso de estudio se analizó las actitudes de los pasajeros hacia el servicio de metro en Sevilla (España). Los resultados confirmaron la utilidad de dicha metodología en otros campos diferentes a la calidad del servicio. Además, proporcionó un modelo sobre dichas actitudes en el metro de Sevilla, que permitió entender mejor la forma de proceder y actuar de los pasajeros hacia este transporte.

Previamente y de manera similar al caso anterior, fue necesario llevar a cabo un pre-procesamiento de los datos que iban a ser utilizados.

5.2.1. PRE-PROCESAMIENTO DE LOS DATOS

Para el estudio de las actitudes se utilizaron los datos relacionados con las actitudes hacia el transporte público, intenciones, beneficios percibidos, costes percibidos, etc. Estos corresponden con los datos procedentes de la Parte A de la encuesta. Así mismo, se consideraron las cuestiones sobre satisfacción general y valoración de la calidad del servicio, y las dimensiones calculadas en el caso de estudio 1 a través del AFE. En total, 36 atributos fueron considerados.

- 26 cuestiones sobre actitudes hacia el servicio (Parte A: A1-A26)
- 8 dimensiones que conforman la calidad del servicio (Accesibilidad, Atención al Cliente, Disponibilidad, Equipamiento Tangible, Espacio Individual, Información, Medio Ambiente y Seguridad).
- Cuestión sobre la Percepción General de la calidad del Servicio (Parte B: SQ1).
- Cuestión sobre la Satisfacción General de los Pasajeros (Parte A: SQ2).

Las cuestiones procedentes de la Parte A (A1-A26) fueron preguntadas siguiendo una escala de 11 puntos Likert (0 -Total desacuerdo y 10-Total acuerdo). Similarmente, la pregunta SQ1 siguió una escala de 11 puntos Likert (0- Muy baja calidad y 10-muy alta calidad) y la cuestión SQ2 (Parte A) fue preguntada siguiendo una escala de 5 puntos Likert (1- Muy insatisfecho y 5-Muy satisfecho). Por último, las 8 dimensiones sobre aspectos de la calidad del servicio fueron calculadas a partir del AFE en el caso de estudio 1, en una escala de 5 puntos Likert, donde 1 representaba una calidad muy baja y 5 una calidad muy alta.

El pre-procesamiento se dividió en tres etapas que a continuación son detalladas.

5.2.1.1. RE-CATEGORIZACIÓN DE LOS ATRIBUTOS

En esta primera fase y siguiendo los argumentos que se expusieron en el caso de estudio 1, los diferentes atributos considerados para este caso de estudio fueron re-categorizados. Se pasó de una escala de 11 puntos Likert a una de 5 puntos Likert. En el caso de la variable SQ2, que fue preguntada originalmente en una escala de 5 puntos Likert, no fue necesaria su re-categorización, así como las 8 dimensiones que conforman la calidad (Accesibilidad, Atención al Cliente, Disponibilidad, Equipamiento Tangible, Espacio Individual, Información, Medio Ambiente y Seguridad), que una vez obtenidas a partir del AFE y AFC en el caso de estudio 1, ya habían sido re-categorizadas en una escala de 5 puntos Likert.

5.2.1.2. ANÁLISIS FACTORIAL EXPLORATORIO

Debido a que se deseó reducir el número de variables mediante la determinación de constructos, y que se desconocen a priori el número de variables que componen dichos constructos latentes, se realizaron tres AFE sobre diferentes configuraciones de atributos.

Un primer AFE fue realizado con 23 cuestiones de actitudes procedentes de la Parte A de la encuesta (variables de la A4 a A26), con el que se esperaba conseguir diferentes constructos latentes sobre las actitudes de los pasajeros, que fueran similares a otros constructos que ya hubieran sido previamente reconocidos en la literatura (p.e., costes percibidos, intención de recompra, beneficios percibidos, etc.) (p.e., De Oña et al., 2016b; Jen and Hu, 2003; Lai and Chen, 2011, Machado-Leon et al., 2016; etc.).

Un segundo AFE fue realizado sobre las cuestiones ligadas directamente a la satisfacción general, es decir, con las cuestiones de A1 a A3 y “Grado de Satisfacción” (SQ2). Con este AFE se esperaba conseguir un único constructo latente que representara la satisfacción general de los pasajeros.

Por último, el tercer AFE fue realizado con las 8 dimensiones de calidad extraídas en el caso de estudio 1 (Accesibilidad, Atención al Cliente, Disponibilidad, Información, Espacio Individual, Seguridad, Equipamiento Tangible y Medio Ambiente) y con el atributo “Valoración Global del Servicio” (SQ1). De este AFE se esperaba conseguir un único constructo que representara la calidad general del servicio.

En la Tabla 19 se muestran los constructos latentes obtenidos de los tres AFE realizados, previa depuración y mejora de los mismos.

Tabla 19. Primeros resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 2.

Constructos		Peso Factorial	Coef. P.F
Actitud General hacia el Transporte Público (Act_Tp)			
A26	Me gusta que la gente sepa que utilizo el transporte público	0,765	0,263
A21	Siento que usar el transporte público es acorde con mi estilo de vida	0,764	0,263
A25	Me gusta que la gente sepa que utilizo el transporte público	0,764	0,259
A22	No importa cuál sea el motivo de mi viaje, que siempre prefiero coger el transporte público	0,698	0,235
A23	Siento que viajando en transporte público colaboro a proteger el medio ambiente	0,645	0,216
A24	Creo que usar el transporte público influye en la opinión de la gente sobre mí	0,582	0,194

Beneficios Percibidos (Ben)			
A9	El servicio es bueno	0,718	0,342
A11	El horario satisface mis necesidades	0,710	0,430
A12	La atención al cliente es buena	0,630	0,320
A13	Me gusta el metro porque me lleva rápido	0,595	0,269
A10	Me parece adecuada la relación calidad-precio	0,318	0,078
Costes Económicos Percibidos (Cost_Eco_Per)			
A4	El precio me parece alto	0,869	0,370
A5	Creo que el precio del billete supera ampliamente los costes del metro	0,839	0,349
A8	Considero altos los costes de viajar en metro (tiempo, dinero y confort)	0,750	0,277
Costes Físicos Percibidos (Cost_Fis_Per)			
A6	Las estaciones se encuentran alejadas de mi origen y/o destino	0,736	0,642
A7	Considero que el tiempo de espera en los andenes es excesivo	0,573	0,464
Alternativas Atractivas (Alt_Atr)			
A15	Considero que existen buenas alternativas de transporte público al metro (como el autobús, taxi, etc.)	0,754	0,498
A16	Da igual el modo de transporte empleado, mientras cumpla mis necesidades	0,652	0,449
A17	Pienso que otros modos de transporte (como el coche, bicicleta, autobús, taxi, etc.) ofrecen más ventajas que el metro	0,619	0,384
Intención de Recompra (Int_Rec)			
A20	Seguro que volveré a utilizar el servicio de metro	0,785	0,414
A18	Volveré a utilizar el metro con las mismas condiciones (dinero, tiempo y confort)	0,749	0,378
A19	Suelo recomendar a otros que usen el metro	0,654	0,286
A14	Prefiero el metro antes que otros transportes públicos (como el autobús, el taxi o el préstamo de bicicletas)	0,548	0,253
Satisfacción (Sat)			
A2	Me siento cómodo viajando en metro	0,818	0,333
SQ2	Grado de satisfacción con el Servicio	0,792	0,322
A3	El servicio de metro suministrado cumple con mis expectativas	0,790	0,321
A1	Me atrae viajar en metro	0,733	0,298

Calidad del Servicio (CdS)			
Eq_Tan	Equipamiento Tangible	0,759	0,169
Acces	Accesibilidad	0,756	0,168
Info	Información	0,739	0,164
SQ1	Valoración Global del Servicio	0,739	0,164
Seg	Seguridad	0,733	0,163
Disp	Disponibilidad	0,721	0,160
At_Cl	Atención al Cliente	0,718	0,159
Esp_Ind	Espacio Individual	0,633	0,141
Med_Am	Medio Ambiente	0,539	0,120

Los resultados obtenidos con los AFE, que son comentados en este apartado, son los conseguidos inicialmente, previa depuración de constructos/atributos no significativos. Tras realizar el AFC en el siguiente apartado, se comentarán los resultados finales, permitiendo observar la mejora de los mismos con respecto a los iniciales.

En primer lugar, se evaluó la adecuación de los datos para realizar el primer AFE con las 23 cuestiones de actitudes (A4-A26). El ratio de observaciones por atributo que se consiguió fue de 139,04, superior a la recomendación de 5. Además, el test de Bartlett (Bartlett, 1950) obtuvo un valor de 253 significativo al nivel 0,001, mientras que el Medida de Adecuación de la Muestra (MSA) mostró un valor aceptable de 0,864 (superior al ideal de 0,8). En consecuencia, se pudo afirmar que los datos eran adecuados para dicho análisis.

Mediante la realización del AFE, se pudieron extraer 6 constructos latentes con valores propios superiores a 1 y que explican una varianza de 58,68%, lo cual, era suficientemente adecuada (Hair et al., 2010). Los constructos conseguidos fueron nombrados como “Actitud hacia el Transporte Público” (Act_Tp), “Beneficios Percibidos” (Ben), “Costes Económico Percibidos” (Cost_Eco_Per), “Costes Físicos Percibidos” (Cost_Fis_Per), “Alternativas Atractivas “ (Alt_Atr) y “Intención de Recompra” (Int_Rec).

Sin embargo, la cuestión A10 “Me parece adecuada la relación calidad-precio” presentó un peso factorial inferior al umbral recomendado de 0,4 (0,318). Su eliminación produjo que el constructo Costes Físico Percibidos dejará de aparecer y que las variables A6 “Las estaciones se encuentran alejadas de mi origen y/o destino” y A7 “Considero que el tiempo de espera en los andenes es excesivo” se agruparán en el constructo Costes Económicos Percibidos con un peso factorial inferior al umbral recomendado de 0,4 (0,314 y 0,315, respectivamente) (Brons et al., 2009). Por este motivo también fueron eliminadas de forma secuencial, comenzando por la que presentó un menor peso factorial.

Debido a que el constructo Costes Físicos Percibidos dejó de obtenerse, se decidió cambiar el nombre del constructo Costes Económicos Percibidos por Costes Percibidos (Cost_Per), que es el habitual usado en la literatura para designar a los constructos que agrupan este tipo de variables p.e., De Oña et al., 2016b; Jen and Hu, 2003; Lai and Chen, 2011, Machado-Leon et al., 2016; etc.) y que hace más fácil la comparación de los resultados aquí obtenidos con otros trabajos.

Finalmente, tras volver a realizar el AFE sobre las 20 cuestiones restantes, el número de factores extraídos se redujo a 5 constructos latentes. Todas las variables obtuvieron pesos de carga superiores a 0,5 en los constructos que explicaban, salvo A14 “Prefiero el metro antes que otros transporte públicos” (0,497) que decidió mantenerse y analizarla con más detalle con el AFC.

La base de datos para el segundo AFE también fue considerada adecuada. Su ratio de observaciones por atributo fue 799,5 (>5), el test de Bartlett fue de 3.824,043, significativo al nivel 0,001, y el MSA fue de 0,746 (superior a 0,5).

En cuanto a los constructos conseguidos, se agruparon todos estos atributos en solo un constructo latente que pudo ser identificado como la satisfacción general de los pasajeros. Su valor propio fue superior a 1 y explicaba una varianza del 61,43%. Además, no hizo falta eliminar ningún atributo, pues todos presentaron un peso factorial superior a 0,4. Dicho constructo fue nombrado como “Satisfacción” (Sat)

Como resultados del último AFE, la muestra se volvió a presentar como adecuada para dicho análisis. Es decir, el ratio de observaciones por atributo fue de 399,75 (>5), el test de Bartlett obtuvo un valor de 10.566,065, significativo al nivel 0,001, y el MSA fue de 0,926 (superior al ideal de 0,8). Se consiguió un solo constructo que representaba la calidad general del servicio, con un valor propio superior a 1, y que explicaba una varianza del 50,10%. Este constructo fue nombrado como “Calidad del Servicio” (CdS).

Se debe destacar que la dimensión Calidad del Servicio en el caso de estudio 1 y en el caso de estudio 2 representan el mismo concepto. No obstante, en el caso de estudio 1 se explica a través de la valoración global sobre el servicio y el grado de satisfacción con el mismo, mientras que en el caso de estudio 2, se explica tanto por la valoración global del servicio, como por sus dimensiones de calidad. Por tanto, con objeto de diferenciarlas, en el caso de estudio 1 se utilizó el acrónimo CS, y en el caso de estudio 2, el acrónimo CdS.

En la Tabla 20 se muestran todos los constructos conseguidos por los 3 AFE y habiendo eliminado aquellas cuestiones que presentaban un peso factorial inferior al umbral recomendado.

Tabla 20. Segundos resultados del Análisis Factorial Exploratorio previa depuración y mejora de las dimensiones, caso de estudio 2.

Constructos		Peso Factorial	Coef. P.F
Actitud General hacia el Transporte Público (Act_Tp)			
A25	Me gusta que la gente sepa que utilizo el transporte público	0,769	0,264
A26	Me gusta que la gente sepa que utilizo el transporte público	0,756	0,254
A21	Siento que usar el transporte público es acorde con mi estilo de vida	0,753	0,253
A22	No importa cuál sea el motivo de mi viaje, que siempre prefiero coger el transporte público	0,699	0,237
A23	Siento que viajando en transporte público colaboro a proteger el medio ambiente	0,630	0,201
A24	Creo que usar el transporte público influye en la opinión de la gente sobre mí	0,615	0,228
Beneficios Percibidos (Ben)			
A9	El servicio es bueno	0,762	0,397
A12	La atención al cliente es buena	0,720	0,413
A11	El horario satisface mis necesidades	0,684	0,430
A13	Me gusta el metro porque me lleva rápido	0,588	0,271
Costes Percibidos (Cost_Per)			
A4	El precio me parece alto	0,876	0,417
A5	Creo que el precio del billete supera ampliamente los costes del metro	0,670	0,407
A8	Considero altos los costes de viajar en metro (tiempo, dinero y confort)	0,785	0,348
Alternativas Atractivas (Alt_Atr)			
A15	Considero que existen buenas alternativas de transporte público al metro (como el autobús, taxi, etc.)	0,755	0,499
A16	Da igual el modo de transporte empleado, mientras cumpla mis necesidades	0,647	0,444
A17	Pienso que otros modos de transporte (como el coche, bicicleta, autobús, taxi, etc.) ofrecen más ventajas que el metro	0,632	0,396
Intención de Recompra (Int_Rec)			
A20	Seguro que volveré a utilizar el servicio de metro	0,826	0,461
A18	Volveré a utilizar el metro con las mismas condiciones (dinero, tiempo y confort)	0,767	0,400
A19	Suelo recomendar a otros que usen el metro	0,648	0,283
A14	Prefiero el metro antes que otros transportes públicos (como el autobús, el taxi o el préstamo de bicicletas)	0,497	0,128

Satisfacción (Sat)			
A2	Me siento cómodo viajando en metro	0,818	0,333
SQ2	Grado de satisfacción con el Servicio	0,792	0,322
A3	El servicio de metro suministrado cumple con mis expectativas	0,790	0,321
A1	Me atrae viajar en metro	0,733	0,298
Calidad del Servicio (CdS)			
Eq_Tan	Equipamiento Tangible	0,759	0,169
Acces	Accesibilidad	0,756	0,168
Info	Información	0,739	0,164
SQ1	Valoración Global del Servicio	0,739	0,164
Seg	Seguridad	0,733	0,163
Disp	Disponibilidad	0,721	0,160
At_Cl	Atención al Cliente	0,718	0,159
Esp_Ind	Espacio Individual	0,633	0,141
Med_Am	Medio Ambiente	0,539	0,120

5.2.1.3. ANÁLISIS FACTORIAL CONFIRMATORIO

Como última fase del pre-procesamiento, se realizó el AFC que permitió testear las diferentes agrupaciones de los atributos en los 7 constructos de actitudes extraídos del AFE.

Durante un primer AFC, se eliminó el constructo Alternativas Atractivas compuesto por los atributos A15 a A17 debido a que, a pesar de que los atributos que lo componían tenían pesos factoriales significativos, su alfa de Cronbach fue muy bajo (0,482) y no se pudo comprobar su validez discriminante ni de convergencia. Así mismo, fue necesario eliminar secuencialmente el atributo Med_Am del constructo CdS, la cuestión A11 del constructo Beneficios Percibidos, A14 de Intención de Recompra y A24 de Actitud General hacia el Transporte público, pues, aunque su peso factorial fue superior a 0,4, , no cumplieron la validez de convergencia.

Finalmente, sobre los 6 constructos resultantes, se volvió a realizar un AFE, cuyos resultados pueden observarse en la Tabla 21.

Tabla 21. Resultados del Análisis Factorial Exploratorio tras la depuración y mejora de las dimensiones, caso de estudio 2.

Constructos		Peso Factorial	Coef. P.F
Actitud General hacia el Transporte Público (Act_Tp)			
A21	Siento que usar el transporte público es acorde con mi estilo de vida	0,786	0,293
A26	Me gusta que la gente sepa que utilizo el transporte público	0,786	0,292
A25	Me gusta que la gente sepa que utilizo el transporte público	0,751	0,283
A22	No importa cuál sea el motivo de mi viaje, que siempre prefiero coger el transporte público	0,709	0,27
A23	Siento que viajando en transporte público colaboro a proteger el medio ambiente	0,669	0,236
Beneficios Percibidos (Ben)			
A12	La atención al cliente es buena	0,824	0,564
A9	El servicio es bueno	0,805	0,49
A13	Me gusta el metro porque me lleva rápido	0,621	0,341
Costos Percibidos (Cost_Per)			
A4	El precio me parece alto	0,880	0,421
A5	Creo que el precio del billete supera ampliamente los costes del metro	0,873	0,415
A8	Considero altos los costes de viajar en metro (tiempo, dinero y confort)	0,797	0,357
Intención de Recompra (Int_Rec)			
A20	Seguro que volveré a utilizar el servicio de metro	0,857	0,529
A18	Volveré a utilizar el metro con las mismas condiciones (dinero, tiempo y confort)	0,816	0,476
A19	Suelo recomendar a otros que usen el metro	0,645	0,318
Satisfacción (Sat)			
A2	Me siento cómodo viajando en metro	0,818	0,333
SQ2	Grado de satisfacción con el Servicio	0,792	0,322
A3	El servicio de metro suministrado cumple con mis expectativas	0,790	0,321
A1	Me atrae viajar en metro	0,733	0,298
Calidad del Servicio (CdS)			
Eq_Tan	Equipamiento Tangible	0,765	0,179
Acces	Accesibilidad	0,764	0,179
Info	Información	0,749	0,176
SQ1	Valoración Global del Servicio	0,739	0,173
Seg	Seguridad	0,729	0,171
At_Cl	Atención al Cliente	0,726	0,17
Disp	Disponibilidad	0,724	0,17
Esp_Ind	Espacio Individual	0,635	0,149

Seguidamente, se definió un modelo de medición en SEM basado en los resultados del AFE realizado previamente, que contaba con 6 variables exógenas: Act_Tp, Ben, Cost_Per, Int_Rec, Sat y CdS. De esta forma, se pudo verificar los atributos y los constructos que resultaron del AFE en base a los resultados del AFC, que son resumidos en la Tabla 22.

Tabla 22. Resultados del Análisis Factorial Confirmatorio, caso de estudio 2.

Constructos		P.C.N.	P.C.E.	S.E.
Actitud General hacia el Transporte Público (Act_Tp)				
<i>(CR: 0,810; AVE: 0,462; Alfa de Cronbach: 0,803)</i>				
A21	Siento que usar el Transporte público es acorde con mi estilo de vida	1,073	0,786	0,293
A26	Me gusta la gente que utiliza el Transporte público	0,987	0,786	0,292
A25	Me gusta que la gente sepa que utilizo el Transporte público	1,000	0,751	0,283
A22	No importa cuál sea el motivo de mi viaje, que siempre prefiero coger el Transporte público	1,049	0,709	0,27
A23	Siento que viajando en Transporte público colaboro a proteger el medio ambiente	0,805	0,669	0,236
Beneficios Percibidos (Ben)				
<i>(CR: 0,735; AVE: 0,484; Alfa de Cronbach: 0,713)</i>				
A12	La atención al cliente es buena	0,939	0,824	0,564
A9	El servicio es bueno	1,000	0,805	0,49
A13	Me gusta el metro porque me lleva rápido	0,766	0,621	0,341
Calidad del Servicio (CdS)				
<i>(CR: 0,874; AVE: 0,465; Alfa de Cronbach: 0,866)</i>				
Eq_Tan	Equipamiento Tangible	0,823	0,765	0,179
Acces	Accesibilidad	0,910	0,764	0,179
Info	Información	0,976	0,749	0,176
SQ1	Valoración Global del Servicio	1,000	0,739	0,173
Seg	Seguridad	1,024	0,729	0,171
At_Cl	Atención al Cliente	1,045	0,726	0,17
Disp	Disponibilidad	0,948	0,724	0,17
Esp_Ind	Espacio Individual	1,101	0,635	0,149
Costos Percibidos (Cost_Per)				
<i>(CR: 0,837; AVE: 0,631; Alfa de Cronbach: 0,833)</i>				
A4	El precio me parece alto	1,000	0,88	0,421
A5	Creo que el precio del billete supera ampliamente los costes del metro	0,993	0,873	0,415
A8	Considero altos los costes de viajar en metro (tiempo, dinero y confort)	1,033	0,797	0,357

Intención de Recompra (Int_Rec)				
<i>(CR: 0,782; AVE: 0,545; Alfa de Cronbach: 0,760)</i>				
A20	Seguro que volveré a utilizar el servicio de metro	0,712	0,857	0,529
A18	Volveré a utilizar el metro con las mismas condiciones (dinero, tiempo y confort)	1,000	0,816	0,476
A19	Suelo recomendar a otros que usen el metro	1,007	0,645	0,318
Satisfacción (Sat)				
<i>(CR: 0,788; AVE: 0,484; Alfa de Cronbach: 0,781)</i>				
A2	Me siento cómodo viajando en metro	0,777	0,818	0,333
SQ2	Grado de satisfacción con el Servicio	0,817	0,792	0,322
A3	El servicio de metro suministrado cumple con mis expectativas	1,000	0,79	0,321
A1	Me atrae viajar en metro	0,779	0,733	0,298
Índices de bondad de ajuste del modelo				
X^2		4293,842		
Grados de Libertad		284		
GFI		0,896		
AGFI		0,871		
RMSEA		0,066		
RMR		0,047		
NFI		0,885		
CFI		0,892		
PGFI		0,725		

Nota: CR: Fiabilidad del constructo; AVE: Media de la varianza extraída; S.E.: Error Estándar de la covarianza estimada (en inglés, Standard Error, SE); P.C.N.: Peso de carga no estandarizado, P.C.E: Peso de carga estandarizados, todos fueron significantes ($p < 0,001$). Los constructos están en negrita.

Se observa que el modelo SEM cumple satisfactoriamente con los umbrales recomendados para los diferentes índices de ajuste (sección 4.2.1). X^2 es significativo al nivel 0,05, indicando que la magnitud de discrepancia entre la muestra y la matriz de covarianza estimada es insignificante. El CFI es muy cercano a 0,9 (0,892), de forma que se puede considerar aceptable, y el RMSEA es menor de 0,07 (0,066). Por tanto, todos ellos justifican que el modelo de medición establecido en SEM para realizar el AFC es adecuado.

La validez de convergencia es confirmada porque todos los pesos de carga estandarizados, para cada uno de los atributos considerados, son superiores a 0,5, estando la mayoría de ellos sobre el umbral ideal de 0,7. En relación al AVE, se observa que solamente los constructos Cost_Per e Int_Rec presentaron un valor superior a 0,5 (0,631 y 0,545, respectivamente).

El resto de constructos mostraron valores inferiores a este umbral, aunque muy próximos a 0,5 (el más pequeño es de 0,462), lo que permite considerarlos como aceptables. Además, los valores de sus Alfas de Cronbach fueron superiores a 0,7, y los valores de CR en todos los constructos estuvieron por encima del umbral de 0,7. Por tanto, se consideraron como aceptables dichos resultados.

En cuanto a la validez discriminante, el AVE de cada constructo fue comparado con el cuadrado de la inter-correlación del resto de constructos (Tabla 23).

Tabla 23. Matriz comparativa del AVE y del cuadrado de las correlaciones de cada dimensión, caso de estudio 2.

	<i>Act_Tp</i>	<i>Ben</i>	<i>CdS</i>	<i>Cost_Per</i>	<i>Int_Rec</i>	<i>Sat</i>
<i>Act_Tp</i>	0,462	0,138	0,194	0,024	0,118	0,186
<i>Ben</i>		0,484	0,712	0,099	0,428	0,677
<i>CdS</i>			0,465	0,131	0,441	0,637
<i>Cost_Per</i>				0,631	0,167	0,181
<i>Int_Rec</i>					0,545	0,493
<i>Sat</i>						0,484

Nota: La diagonal de la matriz muestra los valores del AVE de cada constructo y las celdas sombreadas son valores de inter-correlación al cuadrado entre dos variables superiores a los valores AVE.

Las siguientes inter-correlaciones al cuadrado entre Ben con CdS (0,712) y Sat (0,677) superaron el AVE de Ben (0,484); las inter-correlaciones al cuadrado de CdS con Sat (0,637) superaron el AVE de CdS (0,465); y las inter-correlaciones al cuadrado de Sat con Int_Rec (0,493) que superaron el AVE de Sat (0,484). Debido a esta razón, se llevaron a cabo Test de diferencias Chi-Cuadrado para comprobar la validez discriminante de los constructos anteriormente mencionadas, dos a dos. Todos mostraron evidencias de que el modelo sin correlación ajustaba mejor a los datos, y por tanto, indicaban que cada constructo medía un aspecto diferente. Sus resultados se pueden ver en el Anexo 2.

Por último, su validez nomológica fue comprobada ya que todas los constructos y los atributos agrupados tenían un sentido teórico dentro del marco de este caso de estudio.

Se puede destacar que los resultados conseguidos están en línea con los de otros trabajos en la literatura que han utilizado la misma base de datos y han estudiado las actitudes en el metro de Sevilla (De Oña et al., 2016b; Machado-Leon et al., 2016; 2018). Aunque aparecen similares constructos en dichos trabajos, son ligeramente diferentes debido a que, por ejemplo, en De Oña et al. (2016b), la agrupación de las cuestiones en constructos se basa en lo recogido en la literatura existente y en la forma en que se agruparon en el cuestionario, no como en esta tesis doctoral que se ha utilizado el AFE y AFC para conseguir dicho objetivo.

Machado-Leon et al. (2016) utiliza simplemente las cuestiones relacionadas con Intención de Recompra, Actitud General hacia el Transporte Público, Satisfacción y Calidad del Servicio, por lo que, aunque agrupa las mismas cuestiones en los mismos constructos, sus resultados numéricos son diferentes.

Y por último Machado-Leon et al. (2018) presenta los mismos constructos agrupando las mismas cuestiones que en esta tesis doctoral, salvo la cuestión A10 que es mantenida en dicho trabajo. La principal razón yace en que en esta tesis doctoral se realiza una re-categorización de los datos que puede incurrir en que cambien ligeramente los resultados. No obstante, el hecho de que sean iguales los constructos y sus agrupaciones demuestra que la pérdida de información por la re-categorización es insignificante y soportan la adecuada realización del AFE y AFC en esta tesis doctoral.

Previo a la aplicación de la metodología de dos etapas, y al igual que ocurría en el caso de estudio 1, se procedió a calcular los valores de los constructos a través de los coeficientes de puntuación factorial que proporcionó el software *SPSS v20* (Tabla 21). Los constructos con sus valores calculados a partir de las atributos que los conforman (en una escala de 5 puntos Likert) fueron utilizados para construir las RB.

5.2.2. 1º ETAPA. EXTRACCIÓN DE CONOCIMIENTO DE LOS DATOS

En este paso se trabaja con las RB. Para ello, se consideraron los 6 constructos extraídos del AFE y AFC, calculados a partir de los atributos que los conforman.

Al igual que en los casos de estudio anteriores, se siguió el proceso explicado y una adaptación a este caso de estudio del código en *R* suministrado en Cugnata et al. (2016), para obtener la RB más robusta. Se consideraron para el aprendizaje de las RB los mismos 17 algoritmos que en el caso de estudio 1, así como los valores 1, 0,5 y 0 para puntuar los arcos. El umbral se volvió a establecer en 11 para determinar los arcos considerados con alta probabilidad de ocurrencia. La puntuación de ocurrencia de los arcos (o fuerza) se encuentra representada en una tabla recogida en el Anexo 3.

En esta ocasión no se estableció ninguna condición sobre las relaciones que debían cumplir los constructos considerados, dejando, por tanto, al algoritmo identificar a la RB más robusta sin ningún tipo de restricción.

Tras aprender las diferentes RB y puntuar los arcos, se identificó la red más robusta entre las que obtuvieron el mismo número de arcos con valores mayores que el umbral establecido (11). Esto fue realizado a través del ratio de casos mal clasificados (missclassification rate) (Tabla 24).

Tabla 24. Ratios de missclassification de las RB con el mismo número de arcos con valores mayores al umbral establecido para el caso de estudio 2.

Algoritmo	Métrica	Ratio de missclassification
Hill Climbing	AIC	3,7344
	BIC	3,8594
	Bde	4,0938
	mBde	3,6620
	K2	4,2031
	LogLik	3,9951
TABU	BIC	4,1406
	mBde	4,0000
	K2	3,7304
	LogLik	3,8438

Finalmente, la RB más robusta es mostrada en la Figura 18 que fue aprendida mediante el algoritmo Hill Climbing y puntuación mBDe. Como relaciones relevantes, que se encuentran en rojo en dicha figura, se destacan las relaciones de CdS con Act_Tp, Ben e Int_Rec, y la relación de Int_Rec con Cost_Per.

No obstante, se hace mención a que las relaciones restantes presentan valores cercanos a 11, indicando que el modelo es bastante robusto, ya que la estructura suele repetirse en casi todas las redes aprendidas con los diferentes algoritmos.

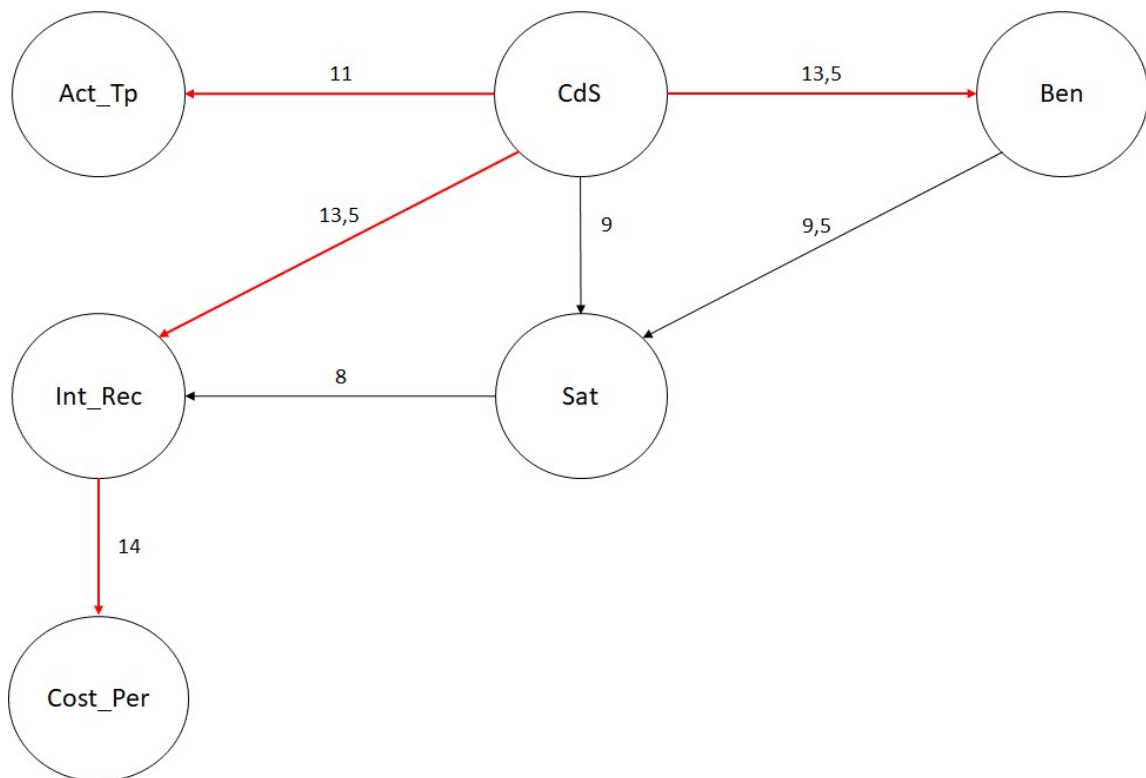


Figura 18. RB más robusta para el caso de estudio 2.

Finalmente, se llevó a cabo el análisis de la fuerza de los arcos mediante el proceso de re-muestreo iterativo sobre la base de datos inicial.

La Figura 19 muestra la proporción de ocurrencia de cada arco en el proceso iterativo de replicación.

Se puede observar que casi todos los arcos rojos de la Figura 18 tuvieron proporciones elevadas y cercanas a 1, evidenciando que dicho modelo y sus relaciones eran robustos.

A partir de esta fuerza de los arcos se calculó el DWI considerando como el nodo objetivo las Int_Rec. Sólo tres constructos latentes ejercían influencia sobre las Int_Rec: CdS con un DWI=0,79, Sat con un DWI=0,33 y Ben con un DWI=0,11, por lo que se pudo establecer que CdS es el constructo más importante para explicar estas Int_Rec.

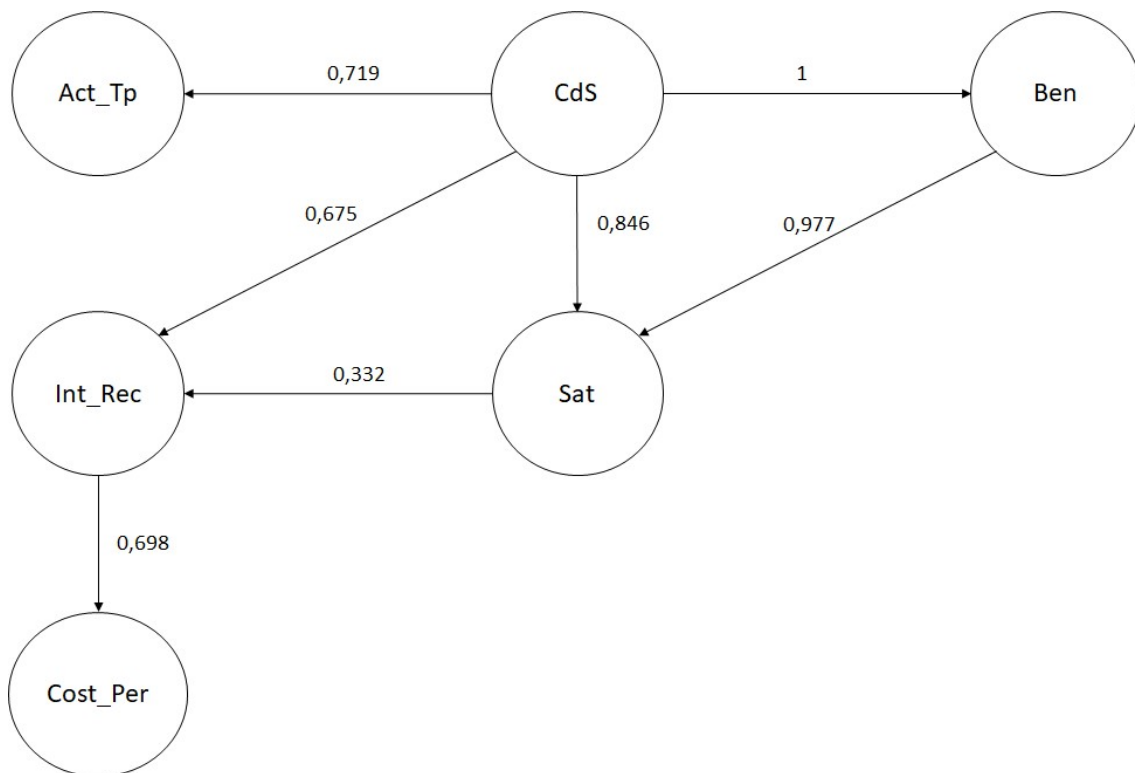


Figura 19. RB más robusta para el caso de estudio 2, con la proporción de ocurrencia de cada arco en el proceso de re-muestreo iterativo.

5.2.3. 2ª ETAPA: VALIDACIÓN DE LAS RELACIONES

Una vez obtenidas las relaciones y el modelo que rige las actitudes hacia el servicio del metro de Sevilla, se procede a su validación. Para tal fin, se utilizó SEM mediante el software *Amos Graphics v.22*. Los parámetros de bondad de ajuste del modelo estructural fueron analizados según lo expuesto en la sección 4.1.4.

Todas las relaciones del modelo son chequeadas de forma que sean superiores al nivel de confianza de 0,01. En caso contrario son eliminadas.

En este caso el modelo SEM fue construido considerando las relaciones de los constructos extraídas en la etapa anterior. Como variables latentes se utilizaron los 6 constructos sobre actitudes (Act_Tp, Ben, CdS, Cost_Per, Int_Rec y Sat), y como variables observadas que explicaban dichas variables latentes, fueron los 18 atributos sobre las actitudes (del A1 al A26 y SQ2, sin considerar A6, A7, A10, A11, A14–A17 y A24), las 7 dimensiones (consideradas como atributos observados a partir del AFE/AFC) sobre calidad del servicio (Acces, At_Cl, Disp, Eq_Tan, Esp_Ind, Info y Seg), , y la variable “Calidad Global del Servicio” (SQ1). Por lo que en total el modelo estuvo compuesto por 6 variables latentes y 26 variables observadas o atributos.

Tras un primer modelado, se pudieron confirmar que todas las relaciones eran significativas al nivel 0,001, incluso la relación de Sat con Int_Rec, cuya proporción de ocurrencia estaba por debajo del 0,5 (0,332) en la Figura 19.

Los resultados son reportados en la Tabla 25 y 26, así como el modelo definitivo en la Figura 20. Se puede observar que todas las relaciones obtenidas con la RB fueron confirmadas por el SEM.

Tabla 25. Pesos de carga y Errores del modelo de medición para el SEM del caso de estudio 2.

Relaciones del Modelo de Medición			No Est.	Est.	S.E.	P
A21	<---	Act_Tp	1	0,736		
A22	<---	Act_Tp	0,976	0,611	0,032	***
A23	<---	Act_Tp	0,754	0,625	0,024	***
A25	<---	Act_Tp	0,933	0,672	0,028	***
A26	<---	Act_Tp	0,924	0,743	0,025	***
A9	<---	Ben	1	0,811		
A12	<---	Ben	0,94	0,653	0,026	***
A13	<---	Ben	0,751	0,606	0,023	***
Acces	<---	CdS	1	0,706		
At_Cl	<---	CdS	1,149	0,679	0,032	***
Disp	<---	CdS	1,045	0,698	0,028	***
Esp_Ind	<---	CdS	1,207	0,573	0,040	***
Eq_Tan	<---	CdS	0,906	0,694	0,025	***
Info	<---	CdS	1,073	0,678	0,030	***
SQ1	<---	CdS	1,101	0,748	0,028	***
Seg	<---	CdS	1,124	0,66	0,032	***
A4	<---	Cost_Per	1	0,819		
A5	<---	Cost_Per	0,997	0,816	0,023	***
A8	<---	Cost_Per	1	0,747	0,024	***
A18	<---	Int_Rec	1	0,775		
A19	<---	Int_Rec	1,023	0,702	0,028	***
A20	<---	Int_Rec	0,709	0,721	0,019	***
A1	<---	Sat	1	0,592		
A2	<---	Sat	1,01	0,699	0,033	***
A3	<---	Sat	1,285	0,704	0,042	***
SQ2	<---	Sat	1,045	0,775	0,032	***

Nota: No Est., No estandarizados; Est. Estandarizados; S.E., Error estándar de la covarianza estimada; ***, $p < 0,001$.

Tabla 26. Pesos de carga y Errores del modelo estructural para el SEM del caso de estudio 2.

Relaciones Estructurales			No Est.	Est.	S.E.	P
Sat	<---	Ben	0,443	0,511	0,040	***
Act_Tp	<---	CdS	0,715	0,451	0,035	***
Ben	<---	CdS	1,068	0,845	0,029	***
Int_Rec	<---	CdS	0,384	0,282	0,050	***
Sat	<---	CdS	0,405	0,371	0,046	***
Cost_Per	<---	Int_Rec	-0,612	-0,443	0,030	***
Int_Rec	<---	Sat	0,622	0,499	0,050	***

Nota: No Est., No estandarizados; Est. Estandarizados; S.E., Error estándar de la covarianza estimada; ***, $p < 0,001$.

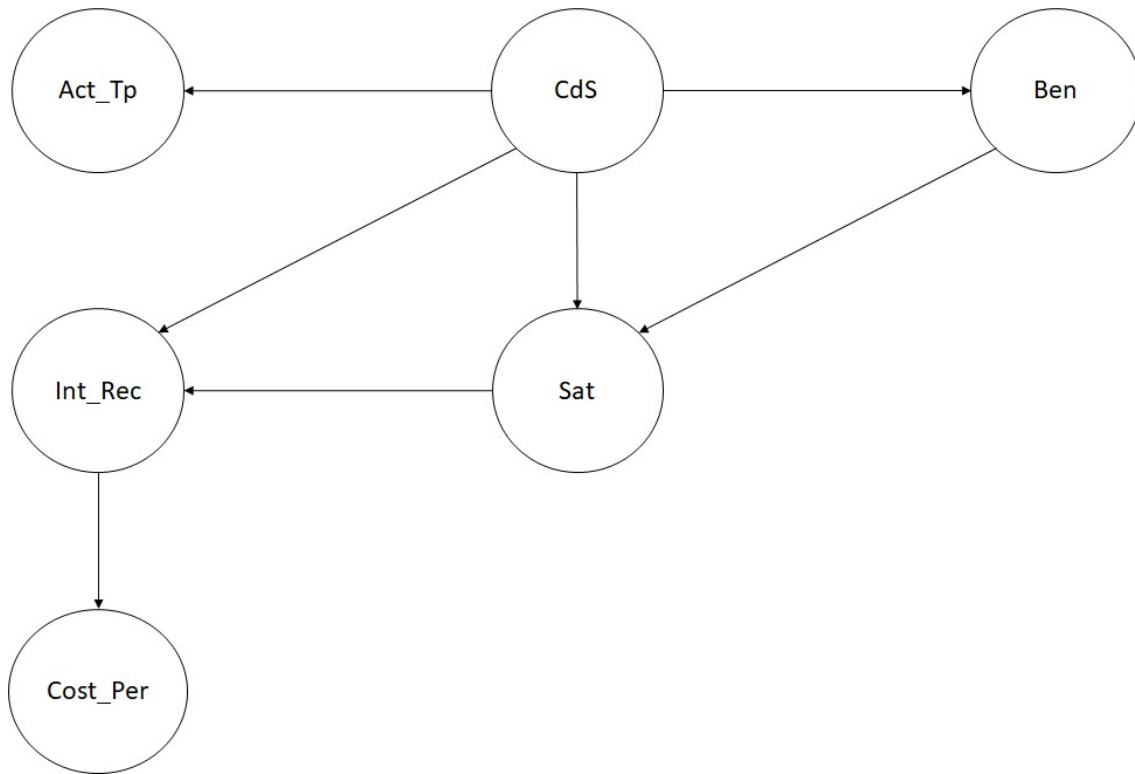


Figura 20 Modelo estructural del SEM del caso de estudio 2, con relaciones significativas.

Los resultados del modelo de medición (Tabla 25) muestran las relaciones entre las variables latentes (constructos) y las observadas (atributos). Se observó que todas las relaciones fueron significativas al nivel 0,001, y todos los pesos de carga estandarizados fueron considerablemente altos (Est.>0,5). Por tanto, no fue necesario depurar el modelo mediante la eliminación de variables.

Por su parte, analizando los resultados del modelo estructural (Tabla 26) formado por las relaciones entre los constructos latentes, se observa que todas las relaciones son significativas al nivel 0,001, incluso la relación entre Sat e Int_Rec. Se debe resaltar el sentido negativo de la relación Int_Rec con Cost_Per, lo que es esperable dado que este constructo tiene connotaciones negativas para los pasajeros (ie., una intención de recompra alta por parte del pasajero induce a que los costes derivados del uso del transporte público sean percibidos en un menor grado).

El hecho de que todas las variables y relaciones sean significativas se traduce en que han sido aplicados correctamente los AFE y AFC, que la RB ha obtenido un buen esquema sobre las actitudes hacia el servicio, y que sus resultados son relevantes.

Las Tablas 27, 28 y 29 muestran los diferentes tipos de efectos (p.e., totales, directos e indirectos) que se producen en el modelo estructural de los diferentes constructos.

De forma general, se observa que todos los efectos, tanto directos como indirectos o totales, de los constructos Ben, CdS, Int_Rec y Sat sobre Cost_Per son negativos. Como ha sido comentado previamente, este hecho es lógico, esperable y coherente con la realidad.

Tabla 27. Efectos directos estandarizados en el modelo de Actitudes.

	Act_Tp	Ben	CdS	Cost_Per	Int_Rec	Sat
Act_Tp	0	0	0,451	0	0	0
Ben	0	0	0,845	0	0	0
Cost_Per	0	0	0	0	-0,443	0
Int_Rec	0	0	0,282	0	0	0,499
Sat	0	0,511	0,371	0	0	0

Tabla 28. Efectos indirectos estandarizados en el modelo de Actitudes.

	Act_Tp	Ben	CdS	Cost_Per	Int_Rec	Sat
Act_Tp	0	0	0	0	0	0
Ben	0	0	0	0	0	0
Cost_Per	0	-0,113	-0,302	0	0	-0,221
Int_Rec	0	0,255	0,401	0	0	0
Sat	0	0	0,432	0	0	0

Tabla 29. Efectos totales estandarizados en el modelo de Actitudes.

	Act_Tp	Ben	CdS	Cost_Per	Int_Rec	Sat
Act_Tp	0	0	0,451	0	0	0
Ben	0	0	0,845	0	0	0
Cost_Per	0	-0,113	-0,302	0	-0,443	-0,221
Int_Rec	0	0,255	0,683	0	0	0,499
Sat	0	0,511	0,803	0	0	0

Los diferentes constructos que tienen efecto sobre los costes percibidos mantienen en común una “actitud” positiva hacia el servicio (ie., influyen de forma positiva en los sentimientos y/o actitudes que un pasajero pueda tener hacia el servicio). Sin embargo, los costes totales percibidos son completamente lo opuesto. Por tanto, es esperable que de producirse algún tipo de relación entre ellos sea de forma inversa. Dicho de otro modo, a medida que haya mayor satisfacción, percepción de la calidad, intención de recompra o beneficios percibidos, menor será la percepción de los costes en los que se incurre al utilizar el servicio de metro.

Otro hecho resaltable es que la actitud que se disponga hacia el transporte público (Act_Tp) no ejerce ningún efecto sobre la intención de recompra (Int_Rec). En este caso, el hecho de que una pasajero tenga una actitud pro-activa hacia el uso del transporte público en general, no implica que se sienta más tentado a volver a utilizar ese mismo servicio de metro, ya sea porque le resulte caro, no le convenga su utilización, tenga otras alternativas de viaje, etc.

Este mismo hecho puede ser extendido al resto de constructos, pues solamente la calidad del servicio percibida producirá un efecto, en sentido positivo, hacia dichas actitudes. Se podría explicar como que el resto de constructos (Beneficios Percibidos, Costes Percibidos, Intención de Recompra, Satisfacción) van asociados al servicio de manera específica, es decir, son sentimientos producidos de forma singular por dicho servicio.

Por otro lado, la calidad del servicio puede ser asociada, por parte del pasajero, a un concepto más amplio, de forma que su actitud hacia el transporte público en general, mejore o empeore, dependiendo directamente de esa calidad percibida.

Por último, cabe destacar que la calidad del servicio (CdS) es el único constructo que tiene influencia sobre todos los otros constructos (de forma directa y/o indirecta), siendo por tanto un factor clave en este modelo de actitudes.

En cuanto a los índices de ajuste de este modelo, se puede observar un resumen de los mismos en la Tabla 30. Adicionalmente, se han recogido índices de otros trabajos, que han utilizado la misma base de datos y los mismos constructos (Machado-Leon et al., 2018); así como de otros trabajos similares sobre actitudes como son el de Irtema et al. (2018) y Li et al. (2018). Se han considerado estos trabajos por sus semejanzas con los atributos y constructos empleados en esta tesis doctoral.

Por otro lado, se puede observar que la Calidad del Servicio influye en la Intención de Recompra de forma directa y, a su vez, influye sobre ésta de forma indirecta a través del constructo de Satisfacción (tal y como se detalla en el popular paradigma de la Calidad del Servicio–Satisfacción-Intención de Recompra), lo que concuerda con los resultados de diversos trabajos previos (Chou et al., 2014; de Oña et al., 2015b; Kuo and Tang, 2013; Lai and Chen, 2011; Sumaedi et al., 2012; etc.).

De hecho, la CdS es el constructo que mayor efecto total ejerce sobre las Int_Rec, seguido de la Sat y por último de los Ben. Esto concuerda con el ranking de importancia establecido a partir de la estructura de la RB y calculado con la métrica DWI. Por lo que en este caso de estudio, sí se comprueba la validez de esta métrica propuesta por Albrecht et al. (2014) y Cugnata et al. (2016). En este caso de estudio 2 se ha permitido un aprendizaje de la red completamente automático, frente a un aprendizaje mixto (parte de la red establecida con restricciones y la otra parte aprendida automáticamente) en el caso de estudio 1. Este motivo ha podido condicionar la validez de la métrica propuesta para determinar el orden de importancia de los constructos, ya que la fuerza de los arcos depende de la probabilidad de ocurrencia de los mismos cuando se realiza el bootstrap.

Por último, los modelos de los trabajos propuestos como referencia han sido contruidos apoyándose en la literatura, mientras que en el caso de estudio de esta tesis doctoral, el modelo ha sido extraído directamente de los datos sin ningún tipo de imposición. Por lo que una mejora o similitud en los parámetros, así como cierta igualdad en las relaciones de los modelos, apoyarán la idoneidad de la metodología de dos etapas para extraer y validar conocimiento directamente de los datos. En conclusión, ambos casos representan buenas referencias para comparar los resultados de este caso de estudio.

Tabla 30. Medidas de bondad de ajuste del modelo SEM del caso de estudio 2.

	Modelo Actitudes hacia el servicio	Machado-Leon et al. (2018)	Machado-Leon et al. (2016)	Irtema et al. (2018)	Li et al. (2018)
Índices de Ajuste					
Tamaño de la muestra	3.198	3.198	3.211	412	337
Chi-Cuadrado	4.404,44	16.685,00	3.084,30	543,191	-
Grados de libertad	292	1415	147	258	-
Nivel de Probabilidad	0.000	-	-	0,000	-
Nº de parámetros distintos estimados	59	-	-	-	-
Índices de Ajuste Absolutos					
GFI	0,893	0,820	0,892	0,902	-
AGFI	0,872	-	0,861	0,880	-
RMSEA	0,066	0,060	0,079	0,047	0,045
RMR	0,051	-	0,173	-	0,06
Índices de Ajuste Incrementales					
NFI	0,882	-	0,902	-	0,904
CFI	0,889	0,850	0,906	-	0,911
Índices de Ajuste Parsimoniosos					
PGFI	0,743	-	0,690	-	0,751
PNFI	0,792	-	-	-	0,750

Una comparación entre los índices de ajuste pone de manifiesto que el modelo conseguido a través de la metodología de dos etapas, propuesta en esta tesis para aplicarla al campo del transporte público, mejora ligeramente al modelo en Machado-Leon et al. (2018) y al modelo en Machado-Leon et al. (2016). Este último trabajo (Machado-Leon et al., 2016) analizaba diferentes modelos de actitudes, seleccionando el que mejores índices de ajuste presentaba. Se ha seleccionado el modelo que más se parece al de esta tesis doctoral para realizar la comparación, el cual, comparte las siguientes relaciones: Calidad del Servicio con Satisfacción, Satisfacción con Intención de Recompra, Calidad del Servicio con Intención de Recompra y Calidad del Servicio con Actitud General hacia el transporte público.

Si se considera el modelo de Machado-Leon et al. (2018) (Figura 21), se observa que existen características comunes: el constructo de Calidad del Servicio influye en la Satisfacción, en los Beneficios Percibidos y en la Intención de Recompra en ambos modelos, y los Beneficios Percibidos influyen en la Satisfacción. Por tanto, se puede decir que el núcleo de ambos modelos es el mismo (Calidad del Servicio-Satisfacción-Lealtad/Intención de Recompra), lo que simplifica y da más sentido a la comparación de ambos estudios. Si se observan los resultados mostrados en la Tabla 30, vemos como los valores obtenidos para los índices de ajuste son muy parecidos, incluso un poco mejor los obtenidos en esta tesis doctoral. Por ejemplo, el GFI y CFI son mayores en los valores conseguido en esta tesis (0,893 frente a 0,820, y 0,889 frente a 0,850, respectivamente), y el RMSEA es similar en ambos modelos (0,066 y 0,06, para el modelo de la tesis y de Machado et al., 2018).

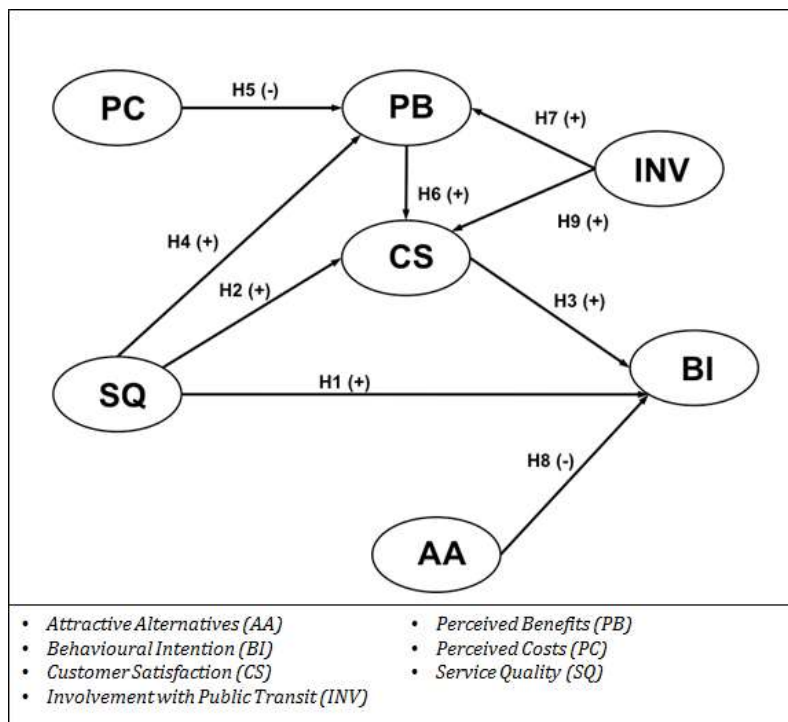


Figura 21 Modelo estructural del SEM de Machado et al. (2018).

Por otro lado, con respecto al trabajo desarrollado por Irtema et al. (2018), éste posee en común con los resultados de esta tesis doctoral las relaciones verificadas entre el constructo de Calidad del Servicio con los constructos Actitud General hacia el Transporte Público, Intención de Recompra y Satisfacción, así como la relación existente entre Satisfacción con Intención de Recompra.

Del mismo modo, con el estudio de Li et al. (2018), también se sustentan las relaciones obtenidas en esta investigación entre la Calidad del Servicio con Intención de Recompra (Lealtad) y Satisfacción, y la relación entre Satisfacción con Lealtad.

Con respecto a estos dos últimos trabajos, aunque los resultados conseguidos en esta tesis doctoral no mejoran los de éstos, sí que se acercan lo suficiente como para considerarlos como satisfactorios.

A raíz de esta comparación se pueden realizar dos lecturas. La primera es la idoneidad de aplicación de esta metodología, desde la perspectiva de ajuste del modelo, para estudiar y analizar modelos de actitudes hacia el transporte público. Y la segunda, es la constatación de que dicha metodología no queda reñida al estudio y análisis de modelos en el campo de la calidad del transporte público, al haberse aplicado de manera satisfactoria esta metodología en otra base de datos con variables diferentes, así como sus objetivos.

Si se analizan los parámetros de ajuste del modelo conseguido con la nueva metodología, se puede observar que el valor X^2 indica que la magnitud de discrepancia entre la muestra y la matriz de covarianza ajustada es insignificante al nivel de 0,05, que es un umbral sugerido por varios autores (p.e., Golob, 2003; Hooper et al., 2008; Mulaik et al., 1989, etc.).

Los índices de ajuste absolutos (GFI y AGFI) muestran valores muy cercanos al valor umbral recomendado de 0,9 (0,893 y 0,872, respectivamente), por lo que pueden ser aceptados. Por su parte, el valor del RMSEA (0,066) se encuentra por debajo de 0,08, y el RMR es igual al valor recomendado de 0,05, lo que indican ambos un buen ajuste del modelo.

Los índices de ajuste incrementales obtenidos presentan valores similares (NFI=0,882 y CFI=0,889) y cercanos a 1, de forma que, aunque no superan el valor recomendado de 0,9, se pueden considerar aceptables. Por último, los valores de los índices de ajuste de parsimonia PGFI (0,743) y PNFI (0,792), son consistentes con las recomendaciones establecidas expresadas por Mulaik et al. (1989).

En conclusión, el modelo presenta valores apropiados, superando, o por lo menos siendo suficientemente cercanos, a los umbrales recomendados por diversos autores (Golob, 2003; Hooper et al., 2008; Mulaik et al., 1989). Comparándose con otros modelos en la literatura, como los propuestos como referencia, constatan lo adecuado que es este modelo, poniendo de manifiesto el interés de esta metodología propuesta para explorar nuevos campos, extrayendo, desarrollando y validando modelos.

5.2.4. DISCUSIÓN SOBRE EL MODELO DE ACTITUDES HACIA EL SERVICIO CONSEGUIDO

Las Figuras 18 y 19 muestran las relaciones que son aprendidas a partir de los datos directamente usando las RB. Por su parte, la Figura 20 muestra las relaciones significativas definitivas del modelo sobre actitudes hacia el servicio que han sido validadas por la técnica SEM. Se puede comprobar que todas las relaciones han sido significativas al nivel 0,001, y no ha sido necesario eliminar ninguna.

Al compararse estas relaciones con las utilizadas en otros trabajos expuestos como referencias (Machado et al., 2016; 2018), se puede observar que el modelo aquí propuesto, que ha sido extraído directamente de los datos sin conocimiento previo a diferencia de los mencionados anteriormente, presenta un gran número de relaciones que son iguales a las validadas en otros trabajos, siendo lo más característico del modelo extraído las relaciones existentes entre la Calidad del Servicio-Satisfacción-Lealtad/Intención de Recompra (núcleo de este modelo, que representa un axioma en cualquier trabajo en este campo).

Por tanto, desde el punto de vista de obtener/extraer las relaciones existentes entre constructos, esta metodología de dos etapas propuesta también se puede considerar adecuada.

Se pueden analizar y explicar las diferentes relaciones conseguidas, mostradas en las Figuras 18, 19 y 20, como sigue:

- La Calidad del Servicio (CdS) muestra un efecto directo sobre el resto de los constructos salvo con Costes Percibidos (Cost_Per) que, como ya se ha visto en la Tabla 29, lo realiza de forma indirecta. Esto pone de relevancia la importancia de dicho constructo en la actitud de los pasajeros hacia dicho servicio, y hacia el transporte público en general. Dicho constructo cubre todos los aspectos relacionados con el servicio (p.e., instalaciones, seguridad, accesos, etc.). Por lo que una buena calidad percibida por el pasajero inducirá a que sus sentimientos y/o actitudes hacia el servicio sean positivos, mostrando una mejor pre-disposición hacia el mismo.

A continuación se analizan las relaciones obtenidas una a una:

- a) Un nivel adecuado de la calidad general del servicio de metro hará que los pasajeros tengan un mejor concepto del transporte público en general (p.e., ambiental o social), ya que muchos de ellos son personas con predisposición al vehículo privado, debido a que piensan que el transporte público no puede presentar las mismas prestaciones y/o calidad que su propio vehículo, o porque no es consistente con su estilo de vida. De esta forma, si se consigue que perciban este servicio como un competidor claro del vehículo privado, en cuanto a calidad y prestaciones se refiere, su actitud hacia el resto de modos de transporte público puede verse mejorada (relación con Actitud General hacia el Transporte Público, Act_Tp). Esta relación ya ha sido testada en otros trabajos como Lai and Chen (2011), Machado-Leon et al. (2016), Tsiotsou (2006), etc.
 - b) Se puede observar que los diferentes beneficios que perciban los pasajeros, ya sea en forma de servicio, económicos, ambientales, de seguridad, etc., van a depender directamente de la calidad con la que dicho servicio se ha prestado, justificando de esta forma su relación con el constructo Beneficios Percibidos (Ben). Similarmente, esta relación se ha testado también en trabajos como el de Doods et al. (1991), Sumaedi et al. (2012), Zeithaml (1988), etc.
 - c) Si la calidad con la que se presta el servicio contribuye a que se cumplan las expectativas del pasajero, se puede observar que su relación con la Satisfacción (Sat) es totalmente coherente y necesaria. Esta relación se encuentra de forma muy frecuente en la literatura y es la corriente que actualmente se encuentra aceptada (Farooq et al., 2018; Li et al., 2018; Mugion et al., 2018).
 - d) Por último, dicha calidad del servicio va a establecer unas condiciones de dinero, tiempo, confort, que de ser buenas, influyen directamente en la intención de volver a utilizar dicho servicio (lealtad) y de su recomendación a otros (relación con Intención de Recompra, Int_Rec). Al igual que ocurre con la satisfacción, esta relación ha sido ampliamente estudiada (Caruana, 2002; Hapsari et al., 2017; Minser and Webb, 2010)
- El constructo Beneficios Percibidos (Ben) se encuentra conectado con Satisfacción (Sat). La explicación a este hecho se encuentra en que los beneficios percibidos cubren aspectos relacionados con la percepción del confort (p.e., atención al cliente, relación calidad-precio, etc.) y con el cumplimiento de expectativas (p.e., rapidez, satisfacción de necesidades, etc.), cuya influencia es directa en la satisfacción de los pasajeros con el servicio.

Varios trabajos han testeado esta relación (De Oña et al., 2016b; Machado et al., 2018, Sumaedi et al., 2012)

- A su vez, Satisfacción (Sat) está relacionada con Intenciones de Recompra (Int_Rec). La atracción hacia utilizar el metro, la comodidad sentida y las expectativas cumplidas, son aspectos que cubre el constructo Sat y que influyen directamente en que un pasajero vuelva a utilizar el servicio en las mismas condiciones o que lo recomienden a otras personas (Int_Rec). Esta dimensión también es conocida en la literatura como Lealtad y conforma junto con satisfacción y calidad del servicio el ampliamente estudiado paradigma de Calidad del Servicio-Satisfacción-Intención de recompra/Lealtad (Chou et al., 2014; De Oña et al., 2015b; Jen et al., 2011)
- Por último, Intención de Recompra (Int_Rec) se encuentra conectada con Costes Percibidos (Cost_Per). Si se tiene una intención de recompra baja (ie., que el servicio le resulte poco atractivo como para volver a reutilizarlo o recomendarlo), influirá en que los costes de utilizar el servicio se perciban como muy altos, incluso alejados de la realidad. Se debe destacar que en la literatura dicha relación es ampliamente recogida, aunque su sentido es inverso (Cost_Per sobre Int_Rec) (Forgas et al., 2010; Hapsari et al., 2017; Lai and Chen, 2011)

En resumen, estos ejemplos muestran que las relaciones extraídas de la RB son apropiadas, mostrando un modelo de actitudes hacia el servicio del metro de Sevilla preciso. Además, las relaciones encontradas automáticamente con la RB muestran una gran similitud con las relaciones establecidas en otros trabajos mediante conocimiento previo o mediante una revisión de la literatura.. Si se le suma el hecho de que las relaciones extraídas mediante la RB son validadas con SEM, se puede afirmar que esta metodología de dos pasos mostrada en esta tesis doctoral puede ser útil y relevante para desarrollar nuevas teorías, identificar nuevas relaciones entre variables y confirmar su validez en otros campos.



CAPÍTULO 6. CONCLUSIONES

Capítulo 6.

CONCLUSIONES

En este capítulo se exponen las principales conclusiones de esta tesis doctoral.

El objetivo principal de este trabajo es desarrollar, implementar y validar una metodología de dos etapas, basada en el uso conjunto de dos técnicas complementarias, para descubrir y confirmar las relaciones existentes entre diferentes dimensiones o constructos que conforman un fenómeno bajo estudio. En esta metodología se aplican las Redes Bayesianas en su primera etapa (como fase exploratoria), y los Modelos de Ecuaciones Estructurales en su segunda etapa (como fase confirmatoria).

La utilidad y validez de esta metodología propuesta han sido verificadas mediante su aplicación a dos bases de datos diferentes, ambas enmarcadas dentro del campo de las percepciones hacia el transporte público.

Por un lado, se ha desarrollado un modelo sobre la Calidad del servicio, y por el otro lado, un modelo sobre Actitudes hacia el servicio, ambos focalizados en el servicio de Metro de Sevilla. Los datos utilizados provienen de una encuesta de calidad/actitudes diseñada y realizada por el grupo de investigación TRYSE en el año 2014 en este servicio de metro. A partir de esta información se modelizaron los dos casos de estudio presentados en esta tesis doctoral, siendo las principales conclusiones obtenidas tras su aplicación las siguientes:

- El Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC) llevados a cabo en una fase previa al empleo de la metodología propuesta de dos etapas, han permitido agrupar de una forma satisfactoria el elevado número de atributos que describen la calidad del servicio (36 atributos) en un número de dimensiones latentes más reducido (8 dimensiones). Estos factores son Equipamiento Tangible, Accesibilidad, Disponibilidad, Atención al Cliente, Seguridad, Información, Medio Ambiente y Espacio Individual. Esta reducción de dimensiones resulta fundamental para reducir posteriormente la complejidad del modelo aprendido y simplificar la interpretación de los resultados obtenidos.
- Se ha determinado la Red Bayesiana más robusta calibrando diferentes algoritmos de aprendizaje y seleccionando aquel algoritmo que presenta un mayor número de arcos (relaciones entre dimensiones) que tienen una fuerza de ocurrencia por encima del umbral establecido, y en caso de empate, el que además cuenta con un menor ratio de casos mal clasificados. Esta metodología para seleccionar la red más robusta se basa en el estudio de Cugnata et al. (2016), y ha sido constatada su efectividad en esta tesis doctoral.
- La Red Bayesiana aprendida ofrece un modelo de calidad del servicio nunca antes estudiado, en el que se identifican relaciones directas e indirectas entre las dimensiones latentes de calidad y la calidad global del servicio (lo que confirma las hipótesis H1 y H2). Este hecho concuerda con las hipótesis previamente planteadas en otros trabajos de investigación recientes (Chou and Kim, 2009; de Oña et al., 2017; Rahman et al., 2016; etc.). En estos trabajos, se rompe con el modelo tradicional de medida de la calidad del servicio en el transporte público, en el que todos los atributos y/o dimensiones influyen de forma directa sobre la calidad global del servicio, y se propone que la relación de dichos atributos y/o dimensiones con la calidad global no siempre tiene que hacerse de forma directa, sino que también puede existir una influencia indirecta. Este modelo pone de manifiesto la necesidad de combinar diferentes técnicas y/o metodologías para llevar a cabo un análisis completo y adecuado de la calidad del servicio, especialmente, en el ámbito del transporte público.

- El análisis del modelo extraído ha servido para identificar qué dimensiones son fundamentales para este servicio mediante el Modelo de Ecuaciones Estructurales, así si el operador desea mejorar la percepción global sobre la calidad de su servicio, debe centrar sus esfuerzos en éstas. En este caso de estudio el Equipamiento Tangible y la Disponibilidad son las dimensiones que obtienen un mayor efecto total. Si se comparan los resultados obtenidos con el ranking propuesto por la métrica DWI de las Redes Bayesianas, sólo coincide la dimensión Equipamiento Tangible en primer lugar, rechazando de esta forma la hipótesis H3. El hecho de que no coincida el ranking establecido a través de la estructura de la RB con los efectos totales obtenidos de los Modelos de Ecuaciones Estructurales puede deberse a dos motivos: 1. La falta de validez de la métrica propuesta por Albrecht et al. (2014), o 2. Esta métrica no puede emplearse cuando la forma de aprendizaje de la red Bayesiana sea mixta (parte de la estructura de la red se establece con restricciones/imposiciones, y la otra es aprendida de forma automática).
- Así mismo, se han identificado relaciones entre las dimensiones bastante interesantes, como por ejemplo, que Equipamiento Tangible tiene relación directa con todas las otras dimensiones que conforman la calidad, menos Medio Ambiente, por lo que una mejora en la misma producirá una mejora sustancial en el resto. Igualmente las dimensiones Accesibilidad e Información no influyen directamente sobre la percepción global del servicio, sino que lo realizan de forma indirecta.
- En base a estos resultados queda demostrado que en la mayor parte de los estudios de calidad del transporte público, se tiende a obviar ciertas relaciones importantes existentes entre los atributos o dimensiones, ya que se recurre a la literatura o conocimiento previo para establecer las relaciones a testear, las cuales en este campo de estudio sólo siguen la pauta de la influencia directa sobre la calidad global del servicio. Las relaciones extraídas y posteriormente validadas en esta tesis doctoral sólo han podido obtenerse gracias al empleo de una técnica exploratoria como son las redes bayesianas. En el trabajo de Oña et al. (2017) se pudo determinar la forma en que influían dos dimensiones latentes sobre la calidad global del servicio a través de la calibración de cinco modelos de ecuaciones estructurales distintos, siguiendo una estrategia de modelos competitivos. Si el número de dimensiones latentes analizadas se incrementa (en esta tesis doctoral se han identificado 8 dimensiones), el número de modelos a calibrar se incrementaría exponencialmente para poder trabajar con la misma estrategia de modelos competitivos que en este trabajo.

- Los parámetros de ajuste obtenidos en el Modelo de Ecuaciones Estructurales han sido muy satisfactorios, mejorando los parámetros obtenidos en el modelo de calidad tradicional (sin relaciones entre las dimensiones latentes, por lo tanto sin efectos indirectos sobre la calidad global). Esto confirma la hipótesis H4.
- En el modelo de actitudes aprendido y validado con la metodología de dos etapas, se confirma el Paradigma de Calidad-Satisfacción-Intenciones de recompra ampliamente aceptado en la literatura (Chou et al., 2014; de Oña et al., 2015b; Kuo and Tang, 2013; Lai and Chen, 2011; Sumaedi et al., 2012; etc.), en el que la Calidad del servicio influye sobre la Satisfacción y la Satisfacción sobre las Intenciones de recompra (confirmando la hipótesis H6). Así mismo, se confirman otras relaciones bien establecidas en la literatura, como es el efecto también directo de la Calidad sobre las Intenciones de Recompra y la influencia de los Beneficios percibidos sobre la Satisfacción. Similarmente, se observa que las Actitudes hacia el transporte público y el Beneficio percibido se encuentran influenciados directamente y positivamente por la percepción de la Calidad global del servicio, de forma que una mejora en este último constructo puede ayudar al cambio de mentalidad de los pasajeros hacia el transporte público.
- Se confirma la hipótesis H5, por la que se calibra un modelo de actitudes hacia el transporte público de forma satisfactoria. De este modo, y a través de los dos casos de estudio planteados en esta tesis doctoral, se comprueba que la metodología de dos etapas propuesta, es una forma robusta y adecuada de modelizar un fenómeno a partir simplemente de una base de datos. En ambos casos de estudio, tras modelizar la Red Bayesiana aprendida con el Modelo de Ecuaciones Estructurales, se consiguen unos parámetros de ajuste satisfactorios y unas relaciones entre las dimensiones o constructos acordes a la realidad. Este modo de proceder ha puesto de relieve la capacidad de extraer conocimiento por parte de las Redes Bayesianas, y de cómo los Modelos de Ecuaciones Estructurales han podido depurar y confirmar dicho conocimiento. Esta metodología solventa el problema de necesidad de conocimiento previo sobre las relaciones que deben existir entre los constructos que conforman un modelo a evaluar, y permite identificar nuevas e importantes relaciones en cualquier ámbito de conocimiento nunca antes explorado.

- Un gran ventaja de estas Redes Bayesianas reside en que éstas permiten descubrir conocimiento de forma completamente automática, por lo que no es necesario poseer ningún conocimiento sobre el campo a explorar (esta modalidad de aprendizaje ha sido aplicada al caso de estudio 2 para el modelo de Actitudes hacia el transporte público), pero también permiten imponer ciertas relaciones causales entre las variables y posteriormente, aprender el resto de la estructura de forma automática (cómo se ha realizado en el caso de estudio 1, sobre el modelo de Calidad del servicio). Esta última modalidad de construcción del modelo bayesiano resulta muy interesante cuando se posee conocimiento experto y se quieren establecer ciertas hipótesis fijas que han sido ampliamente aceptadas en la comunidad científica, y por lo tanto se deben de cumplir.

En resumen, esta tesis doctoral, muestra la eficacia y utilidad de la metodología de dos etapas propuesta para modelizar diferentes fenómenos relacionados con el transporte público, extrayendo de una forma sencilla modelos muy intuitivos y fácilmente interpretables, y que son de gran ayuda para los operadores y gestores del transporte público. Además, su aplicación no está reñida a éste ámbito de estudio, si no que puede ser extrapolada a cualquier área de conocimiento, permitiendo extraer conocimiento en ámbitos nunca antes explorados, así como confirmar y/o mejorar modelos existentes.



CHAPTER 6.

CONCLUSIONS

Chapter 6.

CONCLUSIONS

This chapter presents the main conclusions of this doctoral thesis.

The target of this study is to develop, implement and validate a two-stage methodology, based on the joint use of two complementary techniques. Its main function is to discover and confirm the existing relationships between different dimensions or constructs that make up a phenomenon under study. In this methodology Bayesian Networks are applied in the first stage (as exploratory phase), and the Structural Equation Models in the second stage (as confirmatory phase).

The usefulness and validity of this proposed methodology has been verified through its application into two different databases, both framed within the field of perceptions towards public transport. On the one hand, a service quality model has been obtained, and on the other hand, a model of attitudes towards the public transport has been derived, both of them focused on the Light Rail Transit of Seville.

The data used were gathered from a quality/attitude survey which was designed and carried out by the TRYSE research group in 2014, in this metro service. The two case studies presented in this doctoral thesis were modelled from this information and the main conclusions reached are the following:

- The Exploratory Factor Analysis (EFA) and the Confirmatory Factor Analysis (CFA) carried out in a phase prior to the use of the proposed two-stage methodology, have allowed to satisfactorily group the high number of attributes which describe the service quality (36 attributes) in a smaller number of latent dimensions (8 dimensions). These factors are Tangible Equipment Service, Accessibility, Availability, Customer Service, Security, Information, Environment and Individual Space. This reduction is essential to subsequently reduce the complexity of the learned model and simplify the interpretation of the obtained results.
- The most robust Bayesian Network (BN) has been determined by calibrating different learning algorithms and selecting the one which has a greater number of arcs (relationships between dimensions) that have an occurrence force above the established threshold. In case of a tie, the BN with the lowest ratio of misclassified cases is selected. This methodology to select the most robust network is based on the Cugnata et al. (2016) study, and its effectiveness has been verified in this doctoral thesis.
- The learned BN offers a service quality model never before studied, in which direct and indirect relationships between the latent dimensions of quality and the overall service quality are identified (which confirms hypotheses H1 and H2). This fact is consistent with the hypotheses previously raised in other recent research studies (Chou and Kim, 2009; de Oña et al., 2017; Rahman et al., 2016; etc.). In these works, it is broken with the traditional model of measurement of service quality in public transport, in which all the attributes and/or dimensions influence directly on the overall service quality. Moreover, it is proposed that the relationship between the attributes and/or dimensions with the overall service quality do not always have to be done directly, but indirectly also. This model shows the necessity of combine different techniques and/or methodologies to carry out a complete and suitable analysis of the service quality, especially in the field of public transport.

- The analysis of the extracted model has served to identify which are the fundamental dimensions for this service through the Structural Equations Model (SEM). It allows to identify, for example, if the operator wishes to improve the overall service quality perception, where he must focus his efforts and investments. In this case study, Tangible Equipment Service and Availability are the dimensions that obtain the greatest overall effect. If the results obtained with the SEM are compared with the proposed ranking derived from the DWI metric at the BN, only the Tangible Equipment Service dimension coincides as the most important. Thus, it is rejected the H3 hypothesis. That no coincidence on the established ranking from the BN structure and the total effects obtained from the SEM, may be due to two reasons: 1. The lack of validity of the metric proposed by Albrecht et al. (2014), or 2. This metric cannot be used when the learning method of the BN is mixed (part of the network structure is established with restrictions/impositions, and the other is automatically learned).
- Likewise, it has been identified quite interesting relationships among the dimensions. For example, the Tangible Equipment Service dimension has direct relationship with all the other dimensions that conform the quality, less to the Environment dimension. Therefore, it can be highlighted that an improvement in it will produce a substantial improvement in the rest. Similarly, the Accessibility and Information dimensions do not directly influence the overall perception of the service, but indirectly.
- Based on these results, it is demonstrated that in the majority of the service quality studies in public transport, certain important relationships between attributes or dimensions tend not to be considered. This fact is due to literature or prior expert knowledge is used to establish the relationships to test, which in this field of study only follow the pattern of direct influence on the overall service quality. In this doctoral thesis, the extracted and subsequently validated relationships have only been obtained thanks to the use of an exploratory technique such as BN. In the study of de Oña et al. (2017), it was possible to determine the way in which two latent dimensions influenced the overall service quality. It was achieved by the calibration of five different structural equation models and following a strategy of competitive models. If the number of analysed latent dimensions increase (in this doctoral thesis 8 dimensions have been identified), the number of models to be calibrated would increase exponentially in order to work with the same strategy of competitive models, as in this study.
- The adjustment parameters obtained in the SEM have been very satisfactory. They improve the parameters obtained by the traditional quality model (without relationships between latent dimensions, and therefore without indirect effects on the overall quality). This confirms the H4 hypothesis.

- At the attitude model learned and validated with the two-stage methodology, it is confirmed the Service Quality-Satisfaction-Loyalty Paradigm widely accepted in the literature (Chou et al., 2014; de Oña et al., 2015b; Kuo and Tang, 2013; Lai and Chen, 2011; Sumaedi et al., 2012; etc.). In this paradigm the Service Quality influences on Satisfaction and, at the same time, Satisfaction on Loyalty (it supports the H6 hypothesis). Likewise, other well established relationships in the literature are confirmed, such as, the direct effect of Service Quality on Loyalty and the influence of the Benefits Perceived on Satisfaction. Similarly, it is observed that the Attitudes Towards Public Transport and Benefits Perceived, both are directly and positively influenced by the Overall Service Quality. This fact means that an improvement in the Overall Service Quality construct can help to change the mentality of passengers towards this public transport.
- The H5 hypothesis is confirmed due to an attitude model at a public transport context has been satisfactorily calibrated. Consequently, it is checked that the proposed two-stage methodology is a robust and adequate way to model a phenomenon from a database. In both case studies, after modeling the SEM structure learned with the BN, a satisfactory adjustment of the parameters is achieved and different realistic relationships between dimensions were obtained. This way of proceeding has highlighted the ability to extract knowledge from BN, and how the SEM has been able to refine and confirm this knowledge. This methodology solves the necessity of prior expert knowledge about the relationships between the constructs which make up a model to be evaluated. Besides, it allows to identify new and important relationships in any field of knowledge which has never been explored before.
- A great advantage of the BN is that it allows to discover knowledge automatically. In this way, not only it is not necessary to have any knowledge about the field which is going to be explored (this learning modality has been applied to case study 2), but also it allows to impose certain causal relationships between the variables and, subsequently, learn the rest of the structure automatically (it was done in case study 1). This last mode of learning BN is quite interesting when expert knowledge is available and when certain fixed hypotheses which have been widely accepted in the scientific community wants to be established, and therefore must be met.

In summary, this doctoral thesis shows the effectiveness and usefulness of the two-stage methodology proposed to model different phenomena related to public transport. It allows to extract in a simple way intuitive and easily interpretable models, which are of great help for operators and public transport managers. In addition, its application is not limited to this field of study, but it can be extrapolated to any area of knowledge, allowing knowledge to be extracted in areas never explored before, as well as confirming and/or improving existing models.



**CAPÍTULO 7.
FUTURAS LÍNEAS DE
INVESTIGACIÓN**

Capítulo 7.

FUTURAS LÍNEAS DE INVESTIGACIÓN

A partir de la elaboración de esta tesis y analizando las principales conclusiones obtenidas en la investigación, se plantean algunos estudios para ser realizados en el futuro:

- Utilización de la metodología de dos etapas en otros campos de conocimiento (p.e., Medicina, Turismo, Economía, etc.), de forma que, al igual que ha quedado demostrado en esta tesis doctoral, pueden descubrirse relaciones que no son identificables por otras vías y que pueden cambiar la forma de entender diferentes casos de estudio.

- Aplicación conjunta de la metodología de dos etapas en el análisis de la Calidad del Servicio y de las Actitudes hacia el Servicio, conjuntamente con el Análisis Cluster. El hecho de que existan diferentes tipologías de pasajeros, con sus diferentes contextos socioeconómicos, hábitos y percepciones, hace totalmente comprensible que sus percepciones del transporte público y sus actitudes hacia el mismo sean radicalmente diferentes, lo que, de cara al operador, puede resultar muy útil la obtención de modelos específicos de cada perfil de pasajeros para el desarrollo de estrategias de marketing focalizadas.
- Aplicar la metodología de dos etapas para elaborar un modelo completo que considere las dimensiones de las actitudes y las dimensiones de la calidad del servicio. La obtención de un modelo completo de actitudes y dimensiones de calidad puede llevar al descubrimiento de que ciertas dimensiones de calidad se encuentren relacionadas de alguna forma específica con las actitudes o comportamientos de los pasajeros hacia el transporte público. Este hecho, desde el punto de vista del operador, puede servir para un mejor entendimiento de su servicio, así como para el desarrollo de nuevas teorías sobre la calidad del transporte y su relación con las actitudes de los pasajeros.
- Aplicar la técnica de Redes Bayesianas, en dicha metodología de dos etapas, considerando variables de tipo latente. La consideración de variables latentes en la fase de extracción de conocimiento, puede resultar fundamental para la obtención de modelos más precisos y con mejores parámetros de ajuste, lo cual, implicaría una mejor generalización de los modelos.
- Testear la validez de la métrica propuesta por Albrecht et al. (2014) para identificar las dimensiones que son fundamentales en un modelo de RB y su aplicación en Redes Bayesianas cuyo aprendizaje haya sido mixto (aprendizaje automático de ciertas relaciones combinado con imposición de relaciones basadas en conocimiento experto previo). El hecho de que no hayan coincidido el ranking de la estructura de la Red Bayesiana con los efectos totales obtenidos del Modelo de Ecuaciones Estructurales plantea la cuestión de la validez de esta métrica, y en especial, de su aplicación en este tipo de Redes Bayesianas aprendidas por procedimiento mixto. Su comprobación permitirá, en caso de no ser válida, exponer la necesidad de nuevas métricas adaptadas a estos casos para la identificación de la importancia de los atributos a partir de las Redes Bayesianas.
- Considerar más algoritmos de aprendizaje de las Redes Bayesianas, así como hacer un análisis de sensibilidad para determinar los umbrales a establecer para la selección de la red más robusta (en esta tesis doctoral se han empleado las recomendaciones establecidas por Cugnata et al., 2016).

Esta consideración se escapa del ámbito de estudio de esta tesis doctoral, pero, se considera de gran utilidad de cara a la mejora de la aplicación de la metodología de dos etapas, puesto que, la utilización de más algoritmos junto con el estudio de sensibilidad del umbral conllevaría a explorar más posibles soluciones y, por tanto, la RB seleccionada sería más robusta.

- Realización de inferencia con el modelo conseguido mediante la metodología de dos etapas. El hecho de introducir alguna o algunas evidencias en el modelo de Redes Bayesianas, y analizar cómo se comportarían el resto de variables, es de gran utilidad y relevancia sobre todo en el ámbito del transporte público, ya que serviría de soporte para la planificación de diferentes estrategias comerciales, así como para la localización de inversiones de mejora.



**CAPÍTULO 8.
REFERENCIAS
BIBLIOGRÁFICAS**

Capítulo 8.

REFERENCIAS BIBLIOGRÁFICAS

Abad, F., Olea, J., Ponsoda, V., Garcia, C. (2011). *Medición en Ciencias Sociales y de la Salud*. Madrid. Síntesis.

Abdlla, G., Mohamed, A.R., Mekawy, M.A. (2007). Managing Tourists' Needs and Expectations: An Empirical Analysis of the Egyptian Airline Sector. *Tourism*, 55(3), 277-296.

Abenzoza, R.F., Cats, O., Susilo, Y.O. (2017). Travel Satisfaction with Public Transport: Determinants, User Classes, Regional Disparities and their Evolution. *Transportation Research Part A: Policy and Practice*, 95, 64-84.

Agarwal, R. (2008). Public Transportation and Customer Satisfaction: The Case of Indian Railways. *Global Business Review*, 9(2), 254-272.

Agrawal, R., Imielinski, T., Swami, A., (1993). Mining Association Rules between Sets of Items in Large Database. In: Proceedings of ACM SIGMOD Conference on Management of Data, pp. 207-216.

Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6), 716-723.

Aksoy, S., Atilgan, E., Akinci, S. (2003). Airline Services Marketing by Domestic and Foreign Firms: Differences from the Customers' Viewpoint. Journal of Air Transport Management, 9, 343-351.

Alarcon-del-Amo, M.D., Lorenzo-Romero, C., Gomez-Borja, M.A. (2011). Classifying and Profiling Social Networking Site Users: A Latent Segmentation Approach. Cyberpsychology, Behavior, and Social Networking, 14(9), 547-553.

Albrecht, D., Nicholson, A.E., Whittle, C. (2014). "Structural Sensitivity for the Knowledge Engineering of Bayesian Networks. Probabilistic Graphical Models, 1-16. Basel, Switzerland: Springer International Publishing.

Allen, I.E., Seaman, C.A. (2007). Likert Scales and Data Analyses. Quality Progress, 40(7), 64-65.

Alonso, B., Barreda, R., Dell'Olivo, L., Ibeas, A. (2018). Modelling User Perception of Taxi Service Quality. Transport Policy, 63, 157-164.

Anable, J. (2005). "Complacent Car Addicts" or "Aspiring Environmentalists"? Identifying Travel Behaviour Segments using Attitudes Theory. Transport Policy, 12(1), 65-78.

Andreassen, T. (1995). (Dis)satisfaction with Public Service: The Case of Public Transportation. Journal of Service Marketing, 9(5), 30-41.

Andreassen, T., Lindestad, B. (1998). The Effect of Corporate Image in the Formation of Customer Loyalty. Journal of Service Research, 1(1), 82-92.

Awasthi, A., Chauhan, S.S., Omrani, H., Panahi, A. (2011). A Hybrid Approach based on SERVQUAL and Fuzzy TOPSIS for Evaluating Transportation Service Quality. Computers and Industrial Engineering, 61(3), 637-646.

Aydin, N. (2017). A Fuzzy-Based Multi-Dimensional and Multi-Period Service Quality Evaluation Outline for Rail Transit Systems. Transport Policy, 55, 87-98.

Babakus, E., Boller, G.W. (1992). An Empirical Assessment of the SERVQUAL Scale. Journal of Business Research, 24(3), 253-268.

- Bacon, D.R., Sauer, P.L., Young, M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement*, 55, 394-406.
- Bae, S., Lee, Y. (2018). Comparative Efficacy and Tolerability of Sarilumab 150 and 200 mg in Patient with Active Rheumatoid Arthritis. A Bayesian Network Meta-Analysis of Randomized Controlled Trials. *Zeitschrift Fur Rheumatologie*, 77(5), 421-428.
- Bagozzi, R.P. (1994). *Structural Equations Models in Marketing: Basic Principles*. Principles of Marketing Research. Ed. Cambridge: Blackwell, 317-385.
- Bamberg, S., Fujii, S., Friman, M., Gärling, T. (2011). Behaviour Theory and Soft Transport Measures. *Transport Policy*, 18(1), 228-235.
- Bansal, H.S., Taylor, S. (2015). Investigating the Relationship between Service Quality, Satisfaction and Switching Intentions. *Proceedings of the 1997 Academy of Marketing Science (AMS) Annual Conference*, 304-313.
- Barbero-García, M.I., Vila-Abad, E., Holgado-Tello, F.P. (2011). *Introducción Básica al Análisis Factorial*. Madrid: UNED.
- Barabino, B., Deiana, E., Tilocca, P. (2012). Measuring Service Quality in Urban Bus Transport: a Modified SERVQUAL Approach. *International Journal of Quality and Service Sciences*, 4(3), 238-252.
- Barabino, B., Di Francesco M. (2016). Characterizing, Measuring, and Managing Transit Service Quality. *Journal of Advanced Transportation*, 50, 818-840.
- Barrett, P. (2007). Structural Equation Modelling: Adjudging Model Fit. *Personality and Individual Differences*, 42(5), pp. 815-824.
- Bartlett, M.S. (1950). Tests of Significance in Factor Analysis. *British Journal of Psychology*, 3 (Part II), 77-85.
- Behara, R.S., Fisher, W.W., Lemmink, J.G.A.M. (2002). Modelling and Evaluating Service Quality Measurements using Neural Networks. *International Journal of Operations & Production Management*, 22(10), 1162-1185.
- Beirão, G., Cabral, J.A. (2007). Understanding Attitudes towards Public Transport and Private Car: A Qualitative Study. *Transport Policy*, 14(6), 478-489.
- Beirão, G., Cabral, J.A. (2008). Market Segmentation Analysis using Attitudes toward Transportation: Exploring the Differences between Men and Women. *Transportation Research Record: Journal of the Transportation Research Board*, 2067, 56-64.
- Bentler, P.M. (1990). Comparative Fit Indexes in Structural Models, *Psychological Bulletin*, 107(2), 238-246.

- Bentler, P.M., Bonnet, D.C. (1980). Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychological Bulletin*, 88(3), pp. 588-606.
- Berry, L.L., Zeithaml, V.A., Parasuraman, A. (1990). Five Imperatives for Improving Service Quality. *Sloan Management Review Summer*, 9-38.
- Birago, D., Mensah, S.O., Sharma, S. (2017). Level of Service Delivery of Public Transport and Mode Choice in Accra, Ghana. *Transportation Research Part F: Traffic Psychology and Behaviour*, 46, Part B, 284-300
- Bitner, M.J. (1990). Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses. *The Journal of Marketing*, 54(2), 69-82.
- Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Wiley, New York.
- Bollen, K.A. (1990). Overall Fit in Covariance Structure Models: Two Types of Sample Size Effects. *Psychological Bulletin*, 107(2), 256-259.
- Bolton, R.N., Drew, J.H. (1991). A Multistage Model of Customers' Assessment of Service Quality and Value. *Journal of Consumer Research*, 17(4), 375-384.
- Boomsma, A. (2000). Reporting Analyses of Covariance Structures. *Structural Equation Modeling*, 7(3), 461-483.
- Bordagaray, M., Ibeas, A., dell' Olio, L. (2012). Modeling User Perception of Public Bicycle Services. *Procedia- Social and Behavioral Sciences*, 54, 1308-1316.
- Bowen, N.K., Guo, S. (2011). *Structural Equation Modeling*. Oxford University Press.
- Brady, M.K., Cronin, J.J. (2001). Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach. *Journal of Marketing*, 65, 34-49.
- Brons, M., Givoni, M., Rietveld, P. (2009). Access to Railway Stations and its Potential in Increasing Rail Use. *Transportation Research Part A: Policy and Practice*, 43(2), 136-149.
- Brown, T.A. (2014). *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press.
- Brown, T.J., Churchill, J.G.A., Peter, J.P. (1993). Improving the Measurement of Service Quality. *Journal of Retailing*, 69(1), 127-139.
- Burguillo, M., Romero-Jordan, D., Sanz-Sanz, J.F. (2017). The New Public Transport Pricing in Madrid Metropolitan Area: A Welfare Analysis. *Research in Transportation Economics*, 62, 25-36.

Burns, A.C. (1986). Generating Marketing Strategy Priorities based on Relative Competitive Position. *Journal of Customer Marketing*, 13(4), 49-56.

Byrne, B.M. (1998). *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Cardamone, A.S., Eboli, L., Mazzulla, G. (2014). Drivers' Road Accident Risk Perception. A Comparison between Face-to-Face Interview and Web-Based Survey. *Advances in Transportation Studies. An International Journal, Section B*, 33, 59-72.

Carman, J. (1990). Consumer Perceptions of Service Quality: an Assessment of the SERVQUAL Dimensions. *Journal of Retailing*, 66, 33-55.

Carrillat, F.A., Jaramillo, F., Mulki, J.P. (2007). The Validity of the SERVQUAL and SERVPERF Scales: A Meta-Analytic View of 17 Years of Research across Five Continents. *International Journal of Service Industry Management*, 18(5), 472-490.

Carter, D.N., Lomax, T.J. (1992). Development and Application of Performance Measures for Rural Public Transportation Operators. *Transportation Research Record*, 1338, 28-36.

Caruana, A. (2002). Service Loyalty: The Effects of Service Quality and the Mediating Role of Customer Satisfaction. *European Journal of Marketing*, 36(7), 811-828.

Cascetta, E., Cartenì, A. (2014). A Quality-Based Approach to Public Transportation Planning: Theory and a Case Study. *International Journal of Sustainable Transportation*, 8(1), 84-106.

Castillo, J.M., Benitez, F.G. (2013). Determining a Public Transport Satisfaction Index from User Surveys. *Transportmetrica*, 9(8), 713-741.

Cavana, R.Y., Corbett, L.M., Lo, Y.L. (2007). Developing Zones of Tolerance for Managing Passenger Rail Service Quality. *International Journal of Quality & Reliability Management*, 24(1), 7-31.

Celik, E., Bilisik, O.N., Erdogan, M., Gumus, A.T., Baracli, H. (2013). An Integrated Novel Interval Type-2 Fuzzy MCDM Method to Improve Customer Satisfaction in Public Transportation for Istanbul. *Transportation Research Part E*, 58, 28-51.

Celik, E., Aydin, N., Gumus, A.T. (2014). A Multiattribute Customer Satisfaction Evaluation Approach for Rail Transit Network: A Real Case Study for Istanbul, Turkey. *Transport Policy*, 36, 283-293.

Chakrapani, C. (2004). *Statistics in Marketing Research*. Oxford University Press, New York.

- Chang, Y., Yeh, C. (2002). A Survey Analysis of Service Quality for Domestic Airlines. *European Journal of Operational Research*, 139(1), 166-177.
- Chen, C.F. (2008). Investigating Structural Relationship between Service Quality, Perceived Value, Satisfaction, and Behavioral Intentions for Air Passengers: Evidences from Taiwan. *Transportation Research Part A: Policy and Practice*, 42(4), 709-717.
- Chen, F.Y., Chang, Y.H. (2005). Examining Airline Service Quality from a Process Perspective. *Journal of Air Transport Management*, 11, 79-87.
- Chen, S.H., Pollino, C.A. (2012). Good Practice in Bayesian Network Modeling. *Environmental Modelling & Software*, 37, 134-145.
- Chen, Y.L., Tang, K., Shen, R.J., Hu, Y.H. (2005). Market Basket Analysis in a Multiple Store Environment. *Decision support systems*, 40(2), pp. 339-354.
- Chica-Olmo, J., Gachs-Sanchez, H., Lizarraga, C. (2018). Route Effect on the Perception of Public Transport Service Quality. *Transport Policy*, 67, 40-48.
- Chiou, Y.C., Chen, Y.H. (2012). Service Quality Effects on Air Passenger Intentions: A Service Chain Perspective. *Transportmetrica*, 8(6), 406-426.
- Choi, K.S., Cho, W.H., Lee, S., Lee, H., Kim, C. (2004). The Relationships among Quality, Value, Satisfaction and Behavioral Intention in Health Care Provider Choice: A South Korean Study. *Journal of Business Research*, 57(8), 913-921.
- Chou, J.S., Kim, C. (2009). A Structural Equation Analysis of the QSL Relationship with Passenger Riding Experience on High Speed Rail: An Empirical Study of Taiwan and Korea. *Expert Systems with Applications*, 36(3), part 2, 6945-6955.
- Chou, P.F., Lu, C.S., Chang, Y.H. (2014). Effects of Service Quality and Customer Satisfaction on Customer Loyalty in High-Speed Rail Services in Taiwan. *Transportmetrica A: Transport Science*, 10(10), 917-945.
- Chowdhury, S., Hadas, Y., Gonzalez, V., Schot, B. (2018). Public Transport Users' and Policy Makers' Perceptions of Integrated Public Transport Systems. *Transport Policy*, 61, 75-83.
- Christopher, M.K., Stuart, D., Foote, P.J. (1999). Structuring and Assessing Transit Management Response to Customer Satisfaction Surveys. *Transportation Research Record*, 1669, 99-108.
- Chu, X., Fielding, G.J., Lamar, B.W. (1992). Measuring Transit Performance Using Data Envelopment Analysis. *Transportation Research Part A: Policy and Practice*, 26(3), 223-230.

- Churchill, G.A. (1979). A Paradigm for Developing better Measure of Marketing Constructs. *Journal of Marketing Research*, 16, 64-73.
- Cobanoglu, C., Warde, B., Moreo, P.J. (2001). A Comparison of Mail, Fax, and Web Survey Methods. *International Journal of Market Research*, 43, 441- 452.
- Constantine, L., Ioannis, E., Chaniotakis, M.S. (2006). The Importance of Service Quality in Bank Selection for Mortgage Loans. *Managing Service Quality: An International Journal*, 16(4), 365-379.
- Cooper, G.F., Herskovits, E. (1992). A Bayesian Method for the Induction of Probabilistic Networks from Data. *Machine Learning*, 9(4), 309-347.
- Cortez, P., Embrechts, M.J. (2013). Using Sensitivity Analysis and Visualization Techniques to Open Black Box Data Mining Models. *Information Sciences*, 225, 1-17.
- Coscia, C., Fontana, R., Semeraro, P. (2018). Graphical Models for Complex Networks: an Application to Italian Museums. *Journal of Applied Statistics*, 45(11), pp. 2020-2038.
- Cronbach, L.J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3), 297-334.
- Cronin, J., Taylor, S. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing*, 56, 55-68.
- Crosby, P.B. (1979). *Quality is Free: The Art of Making Quality Certain*. New York: New American Library.
- Crowley, S.L., Fan, X. (1997). Structural Equation Modeling: Basic Concepts and Applications in Personality Assessment Research. *Journal of Personality Assessment*, 68(3), 508-531.
- Cudeck, R. (2000). Exploratory Factor Analysis. In H.E.A. Tinsley & S.D. Brown (Eds.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling*, (265-296). San Diego, C.A.: Academic Press.
- Cugnata, F., Kenett, R.S., Salini, S. (2016). Bayesian Networks in Survey Data: Robustness and Sensitivity Issues. *Journal of Quality Technology*, 48(3), 253-264.
- Cunningham, L., Young, C., Lee, M. (2000). Methodological Triangulation in Measuring Public Transportation Service Quality. *Transportation Journal*, 40(1), 35-47
- Czepiel, J.A. (1990). Service Encounters and Service Relationships: Implications for Research. *Journal of Business Research*, 20, 13-21.

d´Ovidio, F.D., Leogrande, D., Mancarella, R., Schinzano, A., Viola, D. (2014). A Multivariate Analysis of the Quality of Public Transport Services. *Innovation and Society-Statistical Methods for the Evaluation of Services*, 17, 238-247.

Dabholkar, P.A., Shepherd, C.D., Thorpe, D.I. (2000). A Comprehensive Framework for Service Quality: An Investigation of Critical Conceptual and Measurement Issues through a Longitudinal Study. *Journal of Retailing*, 76(2), 139-173.

Dabholkar, P.A., Thorp, D.I., Rentz, J.O. (1996). A Measure of Service Quality for Retail Stores: Scale Development and Validation. *J. Acad. Mark. Sci.*, 24(1), 3-16.

Daniels, R., Mulley, C. (2013). The Paradox of Public Transport Peak Spreading Universities and Travel Demand Management. *International Journal of Sustainable Transportation*, 7(2), 143-165.

Das, S., Pandit, D. (2014). Determination of Level-of-Service Scale Value for Quantitative Bus Transit Service Attributes Based on User Perception. *Transportmetrica A: Transport Science*, 11(1), 1-21.

De Oña, J., De Oña, R., (2014). Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches. *Transportation Science*, 49(3), 605-622.

De Oña, J., De Oña, R., Calvo, F.J. (2012). A Classification Tree Approach to Identify Key Factors of Transit Service Quality. *Expert Systems with Applications*, 39(12), 11164-11171.

De Oña, J., De Oña, R., Eboli, L., Mazzulla, G. (2013a). Perceived Service Quality in Bus Transit Service: A Structural Equation Approach. *Transport Policy*, 29, 219-226.

De Oña, J., De Oña, R., Eboli, L., Mazzulla, G. (2015a). Heterogeneity in Perceptions of Service Quality among Groups of Railway Passengers. *International Journal of Sustainable Transportation*, 9(8), 612-626.

De Oña, J., de Oña, R., Diez-Mesa, F., Eboli, L., Mazzulla, G. (2016a). A Composite Index for Evaluating Transit Service Quality across Different User Profiles. *Journal of Public Transportation*, 19(2), 128-153.

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: Transport Science*, 12(5), 385-412.

De Oña, J., De Oña, R., Lopez, G. (2016c). Transit Service Quality Analysis Using Cluster Analysis and Decision Trees: A Step Forward to Personalized Marketing in Public Transportation. *Transportation*, 43(5), 725-747.

De Oña, J., Lopez, G., Mujalli, R., Calvo, F.J. (2013b). Analysis of Traffic Accidents on Rural Highways using Latent Class Clustering and Bayesian Networks. *Accident Analysis & Prevention*, 51, 1-10.

De Oña, J., Mujalli, R.O., Calvo, F.J. (2011). Analysis of Traffic Accident Injury Severity on Spanish Rural Highways using Bayesian Networks. *Accident Analysis & Prevention*, 43(1), 402-411.

De Oña, R. (2014). Analysis of Service Quality in Public Transportation Using Decision Trees. Granada: Universidad de Granada 174 p. [<http://hdl.handle.net/10481/29935>].

De Oña, R., de Abreu, J., Muñoz-Monge, C., de Oña, J., (2017). Users' Satisfaction Evolution of a Metropolitan Transit System in a Context of Economic Downturn. *International Journal of Sustainable Transportation* (Aceptado)

De Oña, R., Machado, J.L., de Oña, J. (2015b). Perceived Service Quality, Customer Satisfaction, and Behavioral Intentions: Structural Equation Model for the Metro of Seville, Spain. *Transportation Research Record: Journal of the Transportation Research Board*, 2538, pp. 76-85.

Dell'Olio, L., Ibeas, A., Cecin, P. (2010). Modelling User Perception of Bus Transit Quality. *Transport Policy*, 17(6), 388-397.

Dell'Olio, L., Ibeas, A., Cecin, P. (2011). The Quality of Service Desired by Public Transport Users. *Transport Policy*, 18(1), 217-227.

Di Pietro, L., Mugion, R.G., Musella, F., Renzi, M.F., Vicard, P. (2017). Monitoring and Airport Check-In Process by Using Bayesian Networks. *Transportation Research Part A: Policy and Practice*, 106, pp. 235-247.

Diana, M. (2012). Measuring the Satisfaction of Multimodal Travelers for Local Transit Services in Different Urban Contexts. *Transportation Research Part A: Policy and Practice*, 46, 1-11.

Dimitrov, D.M., (2006). Comparing Groups on Latent Variables: A Structural Equation Modeling Approach. *Speaking of Research*, 26, 429-436.

Doods, W.B., Monroe, K.B., Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research*, 28, 307-319.

Duarte, C. W., Klimentidis, Y. C., Harris, J. J., Cardel, M. and Fernandez, J. R. (2011). A Hybrid Bayesian Network/Structural Equation Modelling (BN/SEM) Approach for Detecting Physiological Networks for Obesity-Related Genetic Variants. *Proceedings IEE International Conference Bioinformatics and Biomedicine*, pp. 696-702.

Duffy, B., Smith, K. (2005). Comparing Data from Online and Face-to-Face Surveys. *International Journal of Market Research*, 47(6), 615-639.

Eboli, L., Mazzulla, G. (2007). Service Quality Attributes Affecting Customer Satisfaction for Bus Transit. *Journal of Public Transportation*, 10(3), 21-34.

Eboli, L., Mazzulla, G., (2008). A Stated Preference Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology*, 31(5), 509-523.

Eboli, L., Mazzulla, G., (2009). A New Customer Satisfaction Index for Evaluating Transit Service Quality. *Journal of Public Transportation*, 12(3), 21-37.

Eboli, L., Mazzulla, G., (2010). How to Capture the Passengers' Point of View on a Transit Service through Rating and Choice Options. *Transport Review*, 30(4), 435-450.

Eboli, L., Mazzulla, G. (2011). A Methodology for Evaluating Transit Service Quality Based on Subjective and Objective Measures from the Passenger's Point of View. *Transport Policy*, 18(1), 172-181.

Eboli, L., Mazzulla, G. (2012). Structural Equation Modelling for Analyzing Passengers' Perceptions about Railway Services. *Procedia-Social and Behavioural Science*, 54, 96-106.

Eboli, L., Mazzulla, G. (2015). Relationships between Rail Passengers' Satisfaction and Service Quality: A Framework for Identifying Key Service Factors. *Public Transport*, 7(2), 185-201.

Eboli, L., Yanbing, F., Mazzulla, G. (2016). Multilevel Comprehensive Evaluation of the Railway Service Quality. *Procedia Engineering*, 137, 21-30.

Echaniz, E., Dell'Olio, L., Ibeas, A. (2018). Modelling Perceived Quality for Urban Public Transport Systems Using Weighted Variables and Random Parameters. *Transport Policy*, 67, 31-39.

Ennew, C.T., Binks, M.R. (1996). The Impact of Service Quality and Service Characteristics on Customer Retention: Small Businesses and their Banks in the UK. *British Journal of Management*, 7(3), 219-230.

European Commission (2001). White paper – European Transport Policy for 2010: Time to Decide. COM (370).

European Commission (2007). Green paper – Towards a New Culture for Urban Mobility. COM (551).

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Straham, E.J. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods*, 4, 272-299.

Fan, X., Thompson, B., Wang, L. (1999). Effects of Sample Size, Estimation Methods, and Model Specification on Structural Equation Modeling Fit Indexes, *Structural Equation Modeling*, 6(1), 56-83.

Farooq, M.S. (2016). Social Support and Entrepreneurial Skills as Antecedents of Entrepreneurial Behaviour. PhD Thesis. Universiti Malaysia Sarawak (UNIMAS), Malaysia.

Farooq, M.S., Salam, M., Fayolle, A., Jaafar, N., Ayupp, K. (2018). Impact of Service Quality on Customer Satisfaction in Malaysia Airlines: A PLS-SEM Approach. *Journal of Air Transport Management*, 67, 169-180.

Farzana, R., Das, T., Hadiuzzaman, M., Hossain, S. (2016). Perceived Service Quality of Paratransit in Developing Countries: A Structural Equation Approach. *Transportation Research Part A: Policy and Practice*, 93, 23-38.

Fazil, M., Abulaish, M. (2018). A Hybrid Approach for Detecting Automated Spammers in Twitter. *IEEE Transactions on Information Forensics and Security*, 13(11), 2707-2719.

Fernando, P.J. (1993). *Introducción al Análisis Factorial*. Barcelona: PPU.

Fielding, G.J. (1992). Transit Performance Evaluation in the USA. *Transportation Research*, 26A(6), 483-491.

Fielding, G.J., Babitsky, T.J., Brenner, M.E. (1985). Performance Evaluation for Bus Transit. *Transportation Research*, 19A(1), 73-82.

Fielding, G.J., Glauthier, R.E., Lave, C.A. (1978). Performance Indicators for Transit Management. *Transportation*, 7(4), 365-379.

Figler, S.A., Sriraj, P.S., Welch, E.W., Yavuz, N. (2011). Customer Loyalty and Chicago, Illinois, Transit Authority Buses. *Transportation Research Record*, 2216, 148-156.

Filipovic, S., Tica, S., Zivanovic, P., Milovanovic, B. (2009). Comparative Analysis of the Basic Features of the Expected and Perceived Quality of Mass Passenger Public Transport Service in Belgrade. *Transport*, 24(4), 265-273.

Fillone, A.M., Montalbo, C.M., Tiglaio, N.C. (2005). Assessing Urban Travel: A Structural Equation Modelling (SEM) Approach. *Proceedings of the Eastern Asia Society for Transportation Studies*, 5, 1050-1064.

Finn, D.W., Lamb, Jr. C.W. (1991). An Evaluation of the SERVQUAL Scale in a Retailing Setting. *Advances in Consumer Research*, 18, 312-322.

Fiorio, C.V., Florio, M., Perucca, G. (2013). User Satisfaction and the Organization of Local Public Transport: Evidence from European Cities. *Transport Policy*, 29, 209-218.

Fonseca, F., Pinto, S., Brito, C. (2010). Service Quality and Customer Satisfaction in Public Transport. *International Journal of Quality Research*, 4(2), 249-258.

Foot, P.J., Stuart, D.G., Elmore-Yalch, R. (2001). Exploring Customer Loyalty as a Transit Performance Measure. *Transportation Research Record*, 1753, 93-101.

Forgas, S., Moliner, M.A., Sanchez, J., Palau, R. (2010). Antecedents of Airline Passenger Loyalty: Low-Cost versus Traditional Airlines. *Journal of Air Transport Management*, 16(4), 229-233.

Fornell, C., Larcker, D.F. (1981). Evaluating Structural Equations Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18, 39-50.

Fox, J. (1997). *Applied Regression Analysis, Linear Models, and Related Methods*. Sage Publications, Inc.

Friedman, N., Goldszmidt, M., Wyner, A. (1999). Data Analysis with Bayesian Networks: A Bootstrap Approach. In *Proceedings of the Fifteenth Conference on Uncertainty in artificial intelligence*, 196-205.

Friman, M. (2004). Implementing Quality Improvements in Public Transport. *Journal of Public Transportation*, 7(4), 49-65.

Friman, M., Edvardsson, B., Gärling, T. (2001). Frequency of Negative Critical Incidents and Satisfaction with Public Transport Services. I. *Journal of Retailing and Consumer Services*, 8(2), 95-104.

Friman, M., Fellesson, M. (2009). Service Supply and Customer Satisfaction in Public Transportation: The Quality Paradox. *Journal of Public Transportation*, 12(4), 57-69.

Fu, X., Juan, Z. (2017). Understanding Public Transit Use Behavior: Integration of the Theory of Planned Behavior and the Customer Satisfaction Theory. *Transportation*, 44, 1021-1042.

Fu, X., Zhang, J., Chan, F. (2018). Determinants of Loyalty to Public Transit: A Model Integrating Satisfaction-Loyalty Theory and Expectation-Confirmation Theory. *Transportation Research Part A: Policy and Practice*, 113, 476-490.

Furrer, O., Liu, B.S.C., Sudharshan, D. (2000). The Relationship between Culture and Service Quality Perceptions: Basis for Cross-Cultural Market Segmentation and Resource Allocation. *Journal of Service Research*, 2(4), 355-371.

Garcia-Prats, A., Gonzalez-Sanchis, M., Del Campo, A., Lull, C. (2018). Hydrology-Oriented Forest Management Trade-offs. A Modeling Framework Coupling Field Data, Simulation Results and Bayesian Networks. *Science of the Total Environment*, 639, 725-741.

Garrido, C., De Oña, R., De Oña, J. (2014). Neural Networks for Analyzing Service Quality in Public Transportation. *Expert Systems with Applications*, 41(15), 6830-6838.

Garvin, D.A. (1983). "Paradigms Lost": On Theory and Method in Research in Marketing. *Journal of Marketing*, 47(4), 101-110.

Geetika, Nandan, S. (2010). Determinants of Customer Satisfaction on Service Quality: A Study of Railway Platforms in India. *Journal of Public Transportation*, 13(1), 97-113.

Geurts, K., Thomas, I., Wets, G. (2005). Understanding Spatial Concentrations of Road Accidents Using Frequent Item Sets. *Accident Analysis and Prevention*, 37(4), pp. 787-799.

Gilbert, D., Wong, R.K.C. (2003). Passenger Expectations and Airlines Services: A Hong Kong Based Study. *Tourism Management*, 24, 519-532.

Glover, F., Laguna, M. (1997). *Tabu Search*. Kluwer, Norwell, MA.

Golob, T.F. (2003). Structural Equations Modeling for Travel Behavior Research. *Transportation Research Part B*, 37(1), pp. 1-25.

Gordon, A.D. (1999). *Classification*, Chapman & Hall, CRC, Boca Raton, FL., London.

Green, P.E. (1977). A New Approach to Market Segmentation. *Business Horizons*, 20(1), 61-73.

Greene, D.L., Wegener, M. (1997). Sustainable Transport. *Journal of Transportation Geography*, 5(3), 177-190.

Grönroos, C. (1982). An Applied Service Marketing Theory. *European Journal of Marketing*, 16(7), 30-41.

Grönroos, C. (1984). A Service Quality Model and Its Marketing Implications. *European Journal of Marketing*, Arizona State University.

Grönroos, C. (1988). *Service Quality: The Six Criteria of Good Service Quality*. Review of Business. New York: St. John's University Press.

Grönroos, C. (1990). *Service Management and Marketing: Managing the Moments of Truth in Service Competition*. Free Press, Lexington Books, Lexington, MA.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2010). *Multivariate Data Analysis*. In P. Education Ed., 7th Edition. Prentice Hall, New Jersey.

Han, J., Kamber, M. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Amsterdam.

Han, X., Kwortnik, R.J., Wang, C. (2008). Service Loyalty an Integrative Model and Examintaion across Service Contexts. *Journal of Service Research*, 11(1), 22-42.

Hapsari, R., Clemes, M.D., Dean, D. (2017). The Impact of Service Quality, Customer Engagement and Selected Marketing Constructs on Airline Passenger Loyalty. *International Journal of Quality and Services*, 9(1), 21-40.

Hayduk, L., Cummings, G.G., Boadu, K., Pazderka-Robinson, H., Boulianne, S. (2007). Testing! Testing! One, Two, Three – Testing the Theory in Structural Equation Models!. *Personality and Individual Differences*, 42(2), 841-850.

Haywood-Farmer, J. (1988). A Conceptual Model of Service Quality. *International Journal of Operations & Production Management*, 8(6), 19-29.

Hassan, M.N., Hawas, J.E., Ahmed, K.A. (2013). A Multi-Dimensional Framework for Evaluating the Transit Service Performance. *Transportation Research Part A: Policy and Practice*, 50, 47-61.

Heckerman, D., Geiger, D., Chickering, D.M. (1995). Learning Bayesian Network: The Combination of Knowledge and Statistical Data. *Machine Learning*, 20(3), 197-243.

Hensher, D.A., Daniels, P. (1995). Productivity Measurement in the Urban Bus Sector. *Transport Policy*, 2(3), 179-194.

Hensher, D.A., Mulley, C., Yahya, N. (2010). Passenger Experience with Quality-Enhanced Bus Service: the Tyne and Wear “Superoute” Services. *Transportation* 37(2), 239-256.

Hensher, D.A., Prioni, P. (2002). A Service Quality Index for Area-Wide Contract Performance Assessment. *Journal of Transport Economics and Policy*, 36(1), 9-113.

Hensher, D.A., Stopher, P., Bullock, P. (2003). Service Quality- Developing a Service Quality Index in the Provision of Commercial Bus Contracts. *Transportation Research Part A: Policy and Practice*, 37(6), 499-517.

Hernandez, J., Ferrari, C., Ramirez, M. (2004). *Introducción a la Minería de Datos*. Pearson Educación S.A., España, ISBN: 9788420540917.

- Hernandez, S., Monzon, A., de Oña, R. (2016). Urban Transport Interchanges: A Methodology for Evaluating Perceived Quality. *Transportation Research Part A: Policy and Practice*, 84, 31-43.
- Hill, N., Brierley, G., MacDougall, R. (2003). *How to Measure Customer Satisfaction*. Gower Publishing, Hampshire
- Hine, J., Scott, J. (2000). Seamless, Accessible Travel: Customers' Views of the Public Transport Journey and Interchange. *Transport Policy*, 7(3), 217-226.
- Hojsgaard, S. (2012). Graphical Independence Networks with the gRain Package for R. *Journal of Statistical Software*, 46(19), 1-26.
- Holmbergh, U., Nilsson, G., Martenson, R., Ossiansson, E. (1991). Customers' Perceptions of Service in Retailing. 6th World Conference on Research in the Distributive Trades, The Hague, 60-67.
- Hooper, D., Coughlan, J., Mullen, M.R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods*, 6(1), pp. 53-60.
- Horowitz, J.L. (1991). Reconsidering the Multinomial Probit Model. *Transportation Research Part B: Methodological*, 25(6), 433-438.
- Hu, K. (2010). Evaluating City Bus Service Based on Zone of Tolerance of Expectation and Normalized Importance. *Transport Reviews*, 30(2), 195-217.
- Hu, K., Jen, W. (2006). Passengers' Perceived Service Quality of City Buses in Taipei: Scale Development and Measurement. *Transport Reviews*, 26(5), 645-662.
- Hu, L.T., Bentler, P.M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6(1), pp. 1-55.
- Imam, R. (2014). Measuring Public Transport Satisfaction from User Surveys. *International Journal of Business and Management*, 9(6), 106-114.
- Irtema, H., Ismail, A., Borhan, M., Das, A., Alshetwi, A. (2018). Case Study of the Behavioural Intentions of Public Transportation Passengers in Kuala Lumpur. *Cases Studies on Transport Policy*, In Press.
- Isikli, E., Aydin, N., Celik, E., Gumus, A.T. (2017). Identifying Key Factors of Rail Transit Service Quality: An Empirical Analysis for Istanbul. *Journal of Public Transportation*, 20(1), 63-90.

- Islam, M.R., Hadiuzzaman, M. Banik, R., Hasnat, M.M., Musabbir, S.R., Hossain, S. (2016). Bus Service Quality Prediction and Attribute Ranking: A Neural Network Approach. *Public Transport*, 8(2), 295-313.
- Jabnoun, N., Khalifa, A. (2005). A Customized Measure of Service Quality in the UAE. *Managing Service Quality*, 15(4), 374-388.
- Jen, W., Hu, K.C. (2003). Application of Perceived Value Model to Identify Factors Affecting Passenger's Repurchase Intentions on City Bus: A Case Study of the Taipei Metropolitan Area. *Transportation*, 30, 307-327.
- Jen, W., Tu, R., Lu, T. (2011). Managing Passenger Behavioral Intention: An Integrated Framework for Service Quality, Satisfaction, Perceived Value, and Switching Barriers. *Transportation*, 38, 321-342.
- Jensen, F.V. (2001). *Bayesian Networks and Decision Graphs*. Springer, New York.
- Jiang, H., Zhang, Y. (2016). An Investigation of Service Quality, Customer Satisfaction and Loyalty in China's Airline Market. *Journal of Air Transport Management*, 57, 80-88.
- Jo, N.Y., Park, B. (2017). Empirical Analysis of Online Game Players' Loyalty. *The Journal of Internet Electronic Commerce Research*, 17(2), pp. 169-185.
- Joewono, T.B., Kubota, H. (2007). User Perception of Private Paratransit Operation in Indonesia. *Journal of Public Transportation*, 10(4), 99-118.
- Jöreskog, K.G. (1973). Analysis of Covariance Structures. In *Multivariate Analysis-III*, P.R. Krishnaiah, ed. New York: Academic Press, 263-285.
- Jöreskog, K., Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.
- Kamaruddin, R., Osman, I., Pei, C.A.C. (2012). Public Transport Services in Klang Valley: Customer Expectations and its Relationships Using SEM. *Procedia-Social and Behavioral Sciences*, 36, 431-438.
- Kaplowitz, M.D., Hadlock, T.D., Levine, R. (2004). A Comparison of Web and Mail Survey Response Rates. *Public Opinion Quarterly*, 68(1), 94-101.
- Karlaftis, M.G. (2004). A DEA Approach for Evaluating the Efficiency and Effectiveness of Urban Transit Systems. *European Journal of Operational Research*, 152(2), 354-364.
- Karlaftis, M.G., Golias, J., Papadimitriou, E. (2001). Transit Quality as an Integrated Traffic Management Strategy: Measuring Perceived Service. *Journal of Public Transportation*, 4(1), 27-44.

- Kashani, A.T., Mohaymany, A.S. (2011). Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. *Safety Science*, 49(10), 1314-1320.
- Kaufman, L., Rousseeuw, P.J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Mathematical Statistics.
- Kaur, M., Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Minig. *Procedia Computer Science*, 85, pp. 78-85.
- Kenett, R.S., Salini, S. (2011). Modern Analysis of Customer Satisfaction Surveys: Comparison of Models and Integrated Analysis. *Applied Stochastic Models in Business and Industry*, 27, 465-475.
- Kenett, R., Salini, S. (2012). *Modern Analysis of Customer Surveys: With Applications Using R*. John Wiley & Sons.
- Kenny, D.A., McCoach, D.B. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modelling. *Structutral Equation Modeling*, 10(3), 333-351.
- Kiatcharoenpol, T., Laosirihongthong, T. (2006). Innovations in Service Strategy: An Evaluation of Quality in Airline Service Operations by Using SERVQUAL Model. In: *IEEE International Conference on Management of Innovation and Technology*, Singapore, China, 748-752.
- Kim, N., Kim, J., Lee, C., Know, Y. (2005). Critical Measures of Transit Service Quality in Various City Types. *Proceedings of the Eastern Asia Society for Transportation Studies*, 5, 414-427.
- Kjaerulff, U.B., Madsen, A.L. (2013). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis (2nd Ed.)*. Springer Science+Business Media, New York.
- Kline, R.B. (2005). *Principles and Practice of Structural Equation Modeling (2nd Edition Ed.)*. New York: The Guilford Press.
- Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling (3^a Ed.)*. Guilford Publishing, New York.
- Kuo C., Tang, M. (2013). Relationships among Service Quality, Corporative Imagen, Customer Satisfaction, and Behavioral Intention for the Elderly in High Speed Rail Services. *Journal of Advanced Transportation*, 47, 512-525.
- Kuo, M. (2011). A Novel Interval-Valued Fuzzy MCDM Method for Improving Airlines' Service Quality in Chinese Cross-Strait Airlines. *Transportation Research Part E*, 47(6), 1177-1193.

Lai, J. (2006). Assessment of Employees' Perceptions of Service Quality and Satisfaction with e-business. *International Journal of Human-Computer Studies*, 64(9), 926-938.

Lai, W., Chen, C. (2011). Behavioral Intentions of Public Transit Passengers – The Roles of Service Quality, Perceived Value, Satisfaction and Involvement. *Transport Policy*, 18(2), 318-325.

Landrum, H., Prybutok, V.R. (2004). A Service Quality and Success Model for the Information Service Industry. *European Journal of Operational Research*, 156, 628-642.

Lara, A. (2014). Introducción a la Ecuaciones Estructurales en Amos y R. http://masteres.ugr.es/moea/pages/curso201314/tfm1314/tfm-septiembre1314/memoriasterantonio_lara_hormigo/! (31/12/2016).

Lauritzen, S.L., Spiegelhalter, D.J. (1988). Local Computations with Probabilities on Graphical Structures and their Applications to Expert Systems. *Journal of the Royal Statistical Society, Series B*, 50(2), 157-224.

Lawley, D.N., Maxwell, A.E. (1971). *Factor Analysis as Statistical Method*. London: Butterworths.

Leong, L., Hew, T., Lee, V., Ooi, K. (2015). An SEM-Artificial-Neural-Network Analysis of the Relationships between SERVPERF, Customer Satisfaction and Loyalty among Low-Cost and Full-Service Airline. *Expert Systems with Applications*, 42(19), 6620-6634.

Lethinen, U., Lethinen, J. (1982). *Service Quality – A Study of Quality Dimensions*. Service Management Institute, Helsingfors.

Lewis, R.C., Bernard, H.B. (1983). The Marketing Aspects of Service Quality. *Emerging Perspectives on Services Marketing*, L. Berry, G. Shostack, and G. Upah, eds., Chicago: American Marketing, 99-107.

Li, L., Bai, Y., Song, Z., Chen, A., Wu, B. (2018). Public Transportation Competitiveness Analysis Based on Current Passenger Loyalty. *Transportation Research Part A: Policy and Practice*, 113, 213-226.

Lin, J.H., Lee, T.R., Jen, W. (2008). Assessing Asymmetric Response Effect of Behavioral Intention to Service Quality in an Integrated Psychological Decision-Making Process Model of Intercity Bus Passenger: A Case of Taiwan. *Transportation*, 35, 129-144.

Lin, S., Kernighan, B.W. (1973). An Effective Heuristic for the Traveling-Salesman Problem. *Operational Research*, 21, 498-516.

Liljander, V., Strandvik, T. (1993). Estimating Zones of Tolerance in Perceived Service Quality and Perceived Service Value. *International Journal of Service Industry Management*, 4(2), 6-28.

Liou, J.J.H., Tang, C., Yeh, W., Tsai, C. (2011). A Decision Rules Approach for Improvement of Airport Service Quality. *Expert Systems with Applications*, 38(11), 13723-13730.

Lopez, G. (2014). Análisis de la Severidad de los Accidentes de Tráfico Utilizando Técnicas de Minería de Datos. Granada: Universidad de Granada, 209 p. [<http://hdl.handle.net/10481/30355>].

Lopez-Bonilla, J.M., Lopez-Bonilla, L.M. (2008). Influence of the State-Owned Airlines on Passenger Satisfaction. *Journal of Air Transport Management*, 14(3), 143-145.

MacCallum, R.C., Browne, M.W., Sugawara, H.M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods*, 1(2), 130-149.

Machado-Leon, J.L., de Oña, R., Baouni, T., de Oña, J. (2017). Railway Transit Services in Algiers: Priority Improvement Actions Based on Users Perceptions. *Transport Policy*, 53, 175-185.

Machado-Leon, J.L., de Oña, R., de Oña, J. (2016). The Role of Involvement in Regards to Public Transit Riders' Perceptions of the Service. *Transport Policy*, 48, 34-44.

Machado-Leon, J.L., de Oña, R., Diez-Mesa, F., de Oña, J. (2018). Finding Service Quality Improvement Opportunities across Different Typologies of Public Transit Customers. *Transportmetrica A: Transport Science* (Just Accepted).

Magidson, J., Vermunt, J. (2002). Latent Class Models for Clustering: A Comparison with K-Means. *Canadian Journal of Marketing Research*, 20(1), 37-44.

Maimon, O., Rokach, L. (2010). *Data Mining and Knowledge Discovery Handbook*. Springer, New York.

Marcucci, E., Gatta, V. (2007). Quality and Public Transport Service Contracts. *European Transport*, 36, 92-106.

Margaritis, D. (2003). Learning Bayesian Network Model Structure from Data. Ph.D. thesis, School of Computer Science, Carnegie-Mellon University, Pittsburgh, PA. Available as Technical Report CMU-CS-03-153.

Martilla, J.A., James, J.C. (1977). Importance-Performance Analysis. *Journal of Marketing*, 14(1), 77-79.

Martin del Bio, B., Sanz, A. (2006). *Neural Networks and Fuzzy Systems*. Editorial RA-MA.

MasCallum, R.C., Browne, M.W., Sugawara, H.M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods*, 1(2), 130-149.

Mattsson, J. (1992). A Service Quality Model Based on an Ideal Value Standard. *International Journal of Service Industry Management*, 3(3), 18-33.

McDonald, R.P., Ho, M.H.R. (2002). Principles and Practice in Reporting Statistical Equation Analyses. *Psychological Methods*, 7(1), pp. 64-82

McDougall, G.H.G., Levesque, T. (2000). Customer Satisfaction with Services: Putting Perceived Value into the Equation. *Journal of Services Marketing*, 14(5), 392-410.

McIntosh, C. (2006). Rethinking Fit Assessment in Structural Equation Modelling: A Commentary and Elaboration on Barrett (2007). *Personality and Individual Differences*, 42(5), 859-867.

McLachlan, G.J., Peel, D. (2000). *Finite Mixture Models*. Wiley, New York.

Miles, J., Shevlin, M. (1998). Effects of Sample Size, Model Specification and Factor Loadings on the GFI in Confirmatory Factor Analysis. *Personality and Individual Differences*, 25, 85-90.

Miller, J.A. (1977). Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems, and Making Meaningful Measurements. *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*. Ed. H. Keith Hunt. Bloomington: School of Business, Indiana University, 72-91.

Miller, J.H. (1980). The Use of Performance – Based Methodologies for the Allocation of Transit Operating Funds. *Traffic Quarterly*, 34(4), 555-585.

Minser, J., Webb, V. (2010). Quantifying the Benefits: Applications of Customer Loyalty Modeling in Public Transportation Context. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2144. Transportation Research Board of the National Academies, Washington, D.C., 111-120.

Montella, A., Andreassen, D., Tarko, A., Turner, S., Mauriello, F., Imbriani, L., Singh, R., (2012). Critical Review of the International Crash Databases and Proposals for Improvement of the Italian National Databases. *Procedia – Social and Behavioral*, 53, pp. 49-61.

Morton, C., Caulfield, B., Anable, J. (2016). Customer Perceptions of Quality of Service in Public Transport: Evidence for Bus Transit in Scotland. *Case Studies on Transport Policy*, 4(3), 199-207.

- Mugion, R.G., Toni, M., Raharjo, H., Pietro, L., Sebathu, S.P. (2018). Does the Service Quality of Urban Public Transport Enhance Sustainable Mobility?. *Journal of Cleaner Production*, 174, 1566-1587.
- Mulaik, S.A., James, L.R., Van Alstine, J., Bennet, N., Lind, S., Stilwell, C.D. (1989). Evaluation of Goodness-of-fit Indices for Structural Equation Models. *Psychological Bulletin*, 105(3), 430-445.
- Murray, S.J., Walton, D., Thomas, J.A. (2010). Attitudes towards Public Transport in New Zealand. *Transportation*, 37(6), 915-929.
- Nagarajan, R., Scutari, M., Lebre, S. (2013). *Bayesian Networks in R with Applications in Systems Biology. Use R! Series*. New York, NY: Springer.
- Nandan, S. (2010). Determinants of Customer Satisfaction on Service Quality: A Study of Railway Platforms in India. *Journal of Public Transportation*, 13(1), 97-113.
- Nathanail, E. (2008). Measuring the Quality of Service for Passengers on the Hellenic Railways. *Transportation Research Part A*, 42(1), 48-66.
- Nesheli, M.M., Ceder, A., Brissaud, R. (2017). Public Transport Service-Quality Elements Based on Real-Time Operational Tactics. *Transportation*, 44(5), 957-975.
- Nesheli, M.M., Ceder, A., Estines, S. (2016). Public Transport User's Perception and Decision Assessment Using Tactic-Based Guidelines. *Transport Policy*, 49, 125-136.
- Nilsson, L., Johnson, M.D., Gustafsson, A. (2001). The Impact of Quality Practices on Customer Satisfaction and Business Results: Product versus Service Organizations. *Journal of Quality Management*, 6(1), 5-27.
- Nurul-Habib, K.M., Kattan, L., Islaam, T. (2011). Model of Personal Attitudes towards Transit Service Quality. *Journal of Advanced Transportation*, 45, 271-285.
- Oliver, R.L (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17(4), 460-469.
- Oliver, R.L. (1981). Measurement and Evaluation of Satisfaction Processes in Retail Settings. *Journal of Retailing*, 57, 25-48.
- Oliver, R.L. (1999). Whence Customer Loyalty?. *Journal of Marketing*, 63, 33-44.
- Oliver, R.L. (2010). *Satisfaction: A Behavioral Perspective on the Consumer*. M.E. Sharpe, Armonk, N.Y.
- Olson, J.C., Dover, P. (1979). Disconfirmation of Consumer Expectations Through Product Trial. *Journal of Applied Psychology*, 64, 179-189.

Ostrowski, P.L., O'Brien, T.V., Gordon, G.L. (1993). Service Quality and Customer Loyalty in the Commercial Airline Industry. *Journal of Travel Research*, 32, 16-24.

Ozment, J., Morash, E.A., (1998). Assessment of the Relationship between productivity and Performance Quality in the U.S. Domestic Airline Industry. *Transportation Research Board*, 1622, 22-30.

Parasuraman, A., Valarie, A., Zeithaml, V.A., Leonard, L.B. (1985). A Conceptual Model of Service Quality and its Implications for Future Research. *Journal of Marketing*, 49(4), 41-50.

Parasuraman, A., Zeithaml, V.A., Berry, L.L. (1988). SERVQUAL: A Multiple Item Scale for Measuring Customer Perceptions of Service Quality. *Journal of Retailing*, 64, 12-40.

Park, J., Robertson, R., Wu, C.L. (2004). The Effect of Airline Service Quality on Passengers' Behavioural Intentions: A Korean Case Study. *Journal of Air Transport Management*, 10, 435-439.

Park, J., Robertson, R., Wu, C.L. (2006). Modelling the Impact of Airline Service Quality and Marketing Variables on Passengers' Future Behavioural Intentions. *Transport Planning and Technology*, 29(5), 359-381.

Parkan, C. (2002). Measuring the Operational Performance of a Public Transit Company. *International Journal of Operations and Production Management*, 22(6), 693-720.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco, C.A. Morgan Kaufman.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pedersen, T., Friman, M., Kristensson, P. (2011). Affective Forecasting: Predicting and Experiencing Satisfaction with Public Transportation. *Journal of Applied Social Psychology*, 41(8), 1926-1946.

Pekel, E., Kara, S.S. (2017). A Comprehensive Review for Artificial Neural Network Application to Public Transportation. *Sigma Journal of Engineering and Natural Sciences*, 35(1), 157-179.

Perucca, G., Salini, S. (2014). Travellers' Satisfaction with Railway Transport: A Bayesian Network Approach. *Quality Technology & Quantitative Management*, 11(1), 71-84.

Perry, J.L., Babitsky, T.T. (1986). Comparative Performance in Urban Bus Transit: Assessing Privatization Strategies. *Public Administration Review*, 57-66.

- Philip, K., Gus, M., Rodney, A., John, A. (2003). Customer Repurchase Intention: A General Structural Equation Model. *European Journal of Marketing*, 37(11, 12), 1762-1800.
- Philip, G., Hazlett, S.A. (1997). The Measurement of Service Quality: A New P-C-P Attributes Model. *International Journal of Quality & Reliability Management*, 14(3), 260-286.
- Piatetsky-Shapiro, G., Frawley, W.J., Matheus, C.J. (1991). *Knowledge Discovery in Databases: An Overview*. AAAI-MIT Press, Menlo Park, California.
- Prakash, V. (1984). Validity and Reliability of the Confirmation of Expectations Paradigms as a Determinant of Consumer Satisfaction. *Journal of the Academy of Marketing Science*, 12, 63-76.
- Pullen, W.T. (1993). Definition and Measurement of Quality Service for Local Public Transport Management. *Transport Reviews*, 13(3), 247-264.
- Rahman, F., Das, T., Hadiuzzaman, M., Hossain, S. (2016). Perceived Service Quality of Paratransit in Developing Countries: A Structural Equation Approach. *Transportation Research Part A: Policy and Practice*, 93, pp. 23-38.
- Randheer, K., Al-Motawa, A.A. (2011). Measuring Commuters' Perception on Service Quality Using SERVQUAL in Public Transportation. *International Journal of Marketing Studies*, 3(1), 21-34.
- Redman, L., Friman, M., Gärling, T., Hartig, T. (2013). Quality Attributes of Public Transport that Attract Car Users: A Research Review. *Transport Policy*, 25, 119-127.
- Rezende, F., Ladeira, M. (2019). Market Basket Analysis in a Financial Institution. *Revista Singular-Engenharia, Tecnologia e Gestão*, 1(1), pp. 6-12.
- Rietveld, P. (2005). Six Reasons Why Supply Oriented Indicators Systematically Overestimate Service Quality in Public Transport. *Transport Reviews*, 25(3), 319-328.
- Rojo, M., Dell'Olivo, L., Moura, J.L., Gonzalo, H. (2008). User Satisfaction in Interurban Bus Transit: Modelling Relevant Variables and their Influence. In: *Proceedings of the 10th International Conference on Applications of Advanced Technologies in Transportation*, Athens, Greece.
- Rojo, M., Gonzalo, H., Dell'Olivo, L., Ibeas, A. (2011). Modelling Gender Perception of Quality in Interurban Bus Services. *Proceedings of the Institution of Civil Engineers, Transport*, 164(1), 43-53.

Rokach, L., Maimon, O. (2008). *Data Mining with Decision Trees: Theory and Applications* (Series in Machine Perception and Artificial Intelligence). Singapore: World Scientific Publishing Co. Pte. Ltd.

Sa, L. (2010). Evaluation Model of Passenger Satisfaction Degree of Service Quality for Urban Rail Transit. In: ICCTP 2010: Integrated Transportation Systems: Green, Intelligent, Reliable, 1464-1473.

Sachdew, S.B., Verman, H.V. (2004). Relative Importance of Service Quality. *Journal of Services Research*, 4(1), 93-116.

Saha, G.C., Theingi, (2009). Service Quality, Satisfaction, and Behavioural Intentions: A Study of Low-Cost Airline Carriers in Thailand. *Managing Service Quality*, 19(3), 350-372.

Sampson, S.E., Showalter, M.J. (1999). The Performance-Importance Response Function: Observations and Implications. *Service Industries Journal*, 19(3), 1-25.

Sanchez-Perez, M., Gazquez-Abad, J., Marin-Carrillo, G., Sanchez-Fernandez, R. (2007a). Effects of Service Quality Dimensions on Behavioural Purchase Intentions: A Study in Public-Sector Transport. *Managing Service Quality: An International Journal*, 17(2), 134-151.

Sanchez-Perez, M., Sanchez-Fernandez, R., Marin-Carrillo, G.M., Gazquez-Abad, J.C. (2007b). Service Quality in Public Services as a Segmentation Variable. *The Service Industries Journal*, 27(4), 355-369.

Sari, K.E., Sulisty, D.E., Utomo, D.M. (2017). Reduction of CO2 Emission from Transportation Activities in the Area of Pasar Besar in Malang City. *IOP Conference Series: Earth and Environmental Science*, 70(1), 12-18.

Sasser, W.E., Olsen, R.P., Wyckoff, D.D. (1978). *Understanding Service Operations. Management of Service Operations*, Ally & Bacon, Boston, MA.

Scheines, R., Hoijtink, H. and Boomsma, A. (1999). Bayesian Estimation and Testing of Structural Equation Models. *Psychometrika*, 64(1), 37-52.

Schermelleh-Engel, K., Moosbrugger, H., Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-fit Measures. *Methods of Psychological Research Online*, 8(2), pp. 23-74.

Schwarz, G.E. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6(2), 461-464.

Scutari, M. (2010). Learning Bayesian Network with the Bnlearn R Package. *Journal of Statistical Software*, 35(3), 1-22.

Selman, B., Gomes, C.P. (2006). Hill-Climbing Search. Encyclopedia of Cognitive Science.

Sharma, S. (1996). Applied Multivariate Techniques. John Wiley & Sons.

Sharma, S., Mukherjee, S., Kumar, A., Dillon, W.R. (2005). A Simulation Study to Investigate the Use of Cut-off Values for Assessing Model Fit in Covariance Structure Models. *Journal of Business Research*, 58(1), 935-943.

Sheehan, K.B. (2001). E-mail Survey Response Rates: A Review. *Journal of Computer-Mediated Communication*, (Online), 6(2).

Sheehan, K.B., Hoy, M.G. (1999). Using E-mail to Survey Internet Users in The United States: Methodology and Assessment. *Journal of Computer Mediated Communication*, (Online), 4(3).

Shen, J., Li, W. (2014). Discrete Hopfield Neural Networks for Evaluating Service Quality of Public Transit. *International Journal of Multimedia and Ubiquitous Engineering*, 9(2), 331-340.

Shen, W., Xiao, W., Wang, X. (2016). Passenger Satisfaction Evaluation Model for Urban Rail Transit: A Structural Equation Modeling Based on Partial Least Squares. *Transport Policy*, 46, 20-31.

Shiftan, Y., Outwater, M.L., Zhou, Y. (2008). Transit Market Research Using Structural Equation Modeling and Attitudinal Market Segmentation. *Transport Policy*, 15(3), 186-195.

Simester, D.I., Hauser, J.R., Wernerfelt, B., Rust, R.T. (2000). Implementing Quality Improvement Programs Designed to Enhance Customer Satisfaction: Quasi-Experiments in the United States and Spain. *Journal of Marketing Research*, 37(1), 102-112.

Spearman, C. (1904). "General Intelligence", Objectively Determined and Measured. *The American Journal of Psychology*, 15(2), 201-292.

Steg, L. (2003). Can Public Transport Compete the Private Car?. *International Association of Traffic and Safety Sciences Research*, 27 (2), 27-35.

Steiger, J.H. (2007). Understanding the Limitations of Global Fit Assessment in Structural Equation Modeling. *Personality and Individual Differences*, 42(5), 893-898.

Steiger, J.H., Lind, J.C. (1980). Statistically-Based Test for the Number of Common Factors. Handout for a Presentation Delivered at the Meeting of the Psychometric Society, Iowa City, IA.

Stuart, K.R., Mednick, M., Bockman, J. (2000). Structural Equation Model of Customer Satisfaction for the New York City Subway System. *Transportation Research Record*, 1735, 133-137.

Su, L., Swanson, S.R., Chen, X. (2016). The Effects of Perceived Service Quality on Repurchase Intentions and Subjective Well-being of Chinese Tourists: The Mediating Role of Relationship Quality. *Tourism Management*, 52, pp. 82-95.

Sultan, F., Simpson, M.C. (2000). International Service Variants: Airline Passenger Expectations and Perceptions of Service Quality. *Journal of Services Marketing*, 14(3), 188-216.

Sumaedi, S., Bakti, G., Yarmen, M. (2012). The Empirical Study of Public Transport Passengers' Behavioral Intentions: The Roles of Service Quality, Perceived Sacrifice, Perceived Value and Satisfaction (Case Study: Paratransit Passengers in Jakarta, Indonesia). *International Journal for Traffic and Transport Engineering*, 2(1), 83-97.

Suman, H.K., Bolia, N.B., Tiwari, G. (2017). Comparing Public Bus Transport Service Attributes in Delhi and Mumbai: Policy Implications for Improving Bus Services in Delhi. *Transport Policy*, 56, 63-74.

Swan, J.E., Trawik, F. (1980). Satisfaction Related to Predictive vs Desired Expectations. Refining Concepts and Measures of Consumer Satisfaction and Complaining Behavior. Eds. H. Keith Hunt and Ralph L. Day. Bloomington: School of Business, Indiana University, 7-12.

Tabachnick, B.G., Fidell, L.S. (2013). *Using Multivariate Statistics* (6th Ed.). Boston: Allyn & Bacon.

Talley, W.K., Becker, A.J. (1982). A Single Measure for Evaluating Public Transit Systems. *Traffic Quarterly*, 36(3), 423-431.

Teas, R.K. (1993). Consumer Expectations and the Measurement of Perceived Service Quality. *Journal of Professional Services Marketing*, 8, 33-54.

Teas, R.K. (1994). Expectations as a Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment. *Journal of Marketing*, 58(1), 132-139.

Too, L., Earl, G. (2010). Public Transport Service Quality and Sustainable Development: A Community Stakeholder Perspective. *Sustainable Development*, 18(1), 51-61.

Transportation Research Board. (1998). *A Handbook for Integrating Market Research into Transit Management*. TCRP Report 37, Transportation Research Board, Washington, DC.

Transportation Research Board. (1999). A Handbook for Measuring Customer Satisfaction and Service Quality. TCRP Report 47, Transportation Research Board, Washington, DC.

Transportation Research Board. (2013). Transit Capacity and Quality of Service Manual, Third Edition.

Trentini, F., Malgaroli, M., Camerini, A.L., Di Serio, C. and Schulz, P. (2015). Multivariate Determinants of Self-Management in Health Care: Assessing Health Empowerment Model by Comparison between Structural Equation and Graphical Models Approaches. *Epidemiology Biostatistics and Public Health*, 12 (1), 1-13.

Triplett, J.L., Yau, O.H.M., Neal, C. (1994). Assessing the Reliability and Validity of SERVQUAL in a Longitudinal Study: The Experience of an Australian Organization. *Asia Pacific Journal of Marketing and Logistics*, 6(12), 41-62.

Tripp, C., Drea, J.T. (2002). Selecting and Promoting Service Encounter Elements in Passenger Rail Transport. *The Journal of Services Marketing*, 16(5), 432-442.

Tsai, W., Hsu, W., Chou, W. (2011). A Gap Analysis Model for Improving Airport Service Quality. *Total Quality Management and Business Excellence*, 22(10), 1025-1040.

Tsamardinos, I., Aliferis, C.F., Statnikov, A. (2003). Algorithms for Large Scale Markov Blanket Discovery. In *Proceedings of the Sixteenth International Florida Artificial Intelligence Research Society Conference*, 376-381.

Tsamardinos, I., Brown, L.E., Aliferis, C.F. (2006). The Max-Min Hill Climbing Bayesian Network Structure Learning Algorithm. *Machine Learning*, 65(1), 31-78.

Tsami, M., Nathanail, E. (2017). Guidance Provision for Increasing Quality of Service Public Transport. *Procedia Engineering*, 178, 551-557.

Tsiotsou, R. (2006). The Role of Perceived Product Quality and Overall Satisfaction on Purchase Intentions. *International Journal of Consumer Studies*, 30(2), 207-217.

Tullis, J.A., Jensen, J.R. (2003). Expert System House Detection in High Spatial Resolution Imagery Using Size, Shape, and Context. *Geocarto International*, 18(1), 5-15.

Tyrinopoulos, Y., Antoniou, C. (2008). Public Transit User Satisfaction: Variability and Policy Implications. *Transport Policy*, 15(4), 260-272.

Ubillos, B.J., Sainz, F.A. (2004). The Influence of Quality and Price on the Demand for Urban Transport: The Case of University Students. *Transportation Research Part A: Policy and Practice*, 38(8), 607-614.

Ullman, J.B., Bentler, P.M. (2003). *Structural Equation Modeling*. John Wiley & Sons, Inc.

UNE-EN 13186. (2003). *Transportation. Logistics and Services. Public Passenger Transport. Service Quality Definition, Targeting and Measurement*. AENOR.

Uusitalo, L. (2007). Advantages and Challenges of Bayesian Networks in Environmental Modelling. *Ecological Modelling*, 203, 312-318.

Van Dyke, T.P., Kappelman, A.L., Prybutok, V.R. (1997). Measuring Information Systems Service Quality: Concerns on the Use of SERVQUAL Questionnaire. *MIS Quarterly*, 21(2), 195-208.

Van Lierop, D., El-Geneidy, A. (2016). Enjoying Loyalty: The Relationship between Service Quality, Customer Satisfaction, and Behavioral Intentions in Public Transit. *Research in Transportation Economics*, 59, 50-59.

Vermunt, J., Magidson, J. (2002). Latent Class Cluster Analysis. In Hagenaars, J.A., and A.L. McCutcheon, eds., *Applied Latent Class Analysis*. Cambridge: Cambridge University Press, 89-106.

Vermunt, J., Magidson, J. (2005). *Latent Gold 4.0 User's Guide*. Belmont, MA: Statistical Innovations, Inc.

Vicente, P., Reis, E. (2016). Profiling Public Transport Users through Perceptions about Public Transport Providers and Satisfaction with the Public Transport Service. *Public Transport*, 8(3), 387-403.

Wan, D., Kanga, C., Hao, W., Sugiura, A., Beaton, E. (2016). Customer Satisfaction with Bus Rapid Transit: a Study of New York City Select Bus Service Applying Structural Equation Modeling. *Public Transport*, 8, 497-520.

Wang, S., Feng, C., Hsieh, C. (2010). Stakeholder Perspective on Urban Transport System Service Quality. *Total Quality WManagement and Business Excellence*, 21(11), 1103-1119.

Wang, Y., Lo, H.P., Chi, R., Yang, Y. (2004). An Integrated Framework for Customer Value and Customer Relationship-Management Performance: a Customer-Based Perspective from China. *Managing Service Quality: An International Journal*, 14 (2/3), 169-182.

Weible, R., Wallace, J. (1998). The Impact of the Internet on Data Collection. *Marketing Research*, 10(3), 19-23.

- Wei-Jaw, D. (2008). Fuzzy Importance-Performance Analysis for Determining Critical Service Attributes. *International Journal of Service Industry Management*, 19(2), 252-270.
- Weijawa, D. (2007). Using a Revised Importance-Performance Analysis Approach: The Case of Taiwanese Hot Springs Tourism. *Tourism Management*, 28(5), 1274-1284.
- Weinstein, A. (2000). Customer Satisfaction among Transit Riders. How Customer Rank the Relative Importance of Various Service Attributes. *Transportation Research Record*, 1735, 123-132.
- Weiss, S.M., Indurkha, N. (1998). *Predictive Data Mining*. Morgan Kaufman Publishers. San Francisco.
- Wen, C.H., Lai, S.C. (2005). Structural Equation Modeling to Passengers Loyalty towards Intercity Bus Services. *Transportation Research Record*, 1927, 249-255.
- Wen, C.H., Lai, S.C. (2010). Latent Class Models of International Air Carrier Choice. *Transportation Research Part E*, 46, 211-221.
- Wen, C.H., Lai, S.C., Yen, W. (2008). Segmentation and Positioning Analysis Air Carrier Choice. *Transportation Research Part E*, 46, 211-221.
- Werner, C., Schermelleh-Engel, K. (2010). Deciding Between Competing Models: Chi-Square Difference Tests. *Introduction to Structural Equation Modeling with LISREL*.
- Wiley, D.E. (1973). The Identification Problem for Structural Equation Models with Unmeasured Variables. In *Structural Equation Models in the Social Science*, A.S. Goldberger and O.D. Duncan, Eds. New York: Seminar Press, 69-83.
- Williams, B., Onsmann, A., Brown, T. (2010). Exploratory Factor Analysis: A Five-Steps Guide for Novices. *Journal of Emergency Primary Health Care*, 8(3), 1-13.
- Wipper, L.R. (1993). Performance Measurement: Producing Results at the Oregon Department of Transportation. *Transportation Research Record*, 1395, 168-175.
- Witten, I.H., Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufman, Amsterdam.
- Wong, J., Chung, P. (2007). Managing Valuable Taiwanese Airline Passengers Using Knowledge Discovery in Database Technique. *Journal of Air Transport Management*, 13(6), 362-370.
- Woods, R., Masthoff, J. (2017). A Comparison of Car Driving, Public Transport and Cycling Experiences in Three European Cities. *Transportation Research Part A*, 103, 211-222.

Wu, J., Yang, M., Rasouli, S., Xu, C. (2016). Exploring Passenger Assessments of Bus Service Quality Using Bayesian Networks. *Journal of Public Transportation*, 19(3), 36-54.

Yang, J., Park, J., Choi, Y. (2015). Passengers' Expectations of Airport Service Quality: A Case Study of Jeju International Airport. *International Journal of Business and Social Research*, 5(7), 30-37.

Yang, K., Hsieh, T., Li, H., Yang, C. (2012). Assessing How Service Quality, Airline Image and Customer Value Affect the Intentions of Passengers Regarding Low Cost Carriers. *Journal of Air Transport Management*, 20, 52-53.

Yanik, S., Aktas, E., Topcu, Y.I. (2017). Traveler Satisfaction in Rapid Rail Systems: The Case of Istanbul Metro. *International Journal of Sustainable Transportation*, 11(9), 642-658.

Yao, Z.G., Ding, X.D. (2011). Measuring Passenger's Perceptions of Taxi Service Quality with Weighed SERVPERF. *Applied Mechanics and Materials*, 97-98, 1181-1184.

Yaramakala, S., Margaritis, D. (2005). Speculative Markov Blanket Discovery for Optimal Feature Selection. In *ICDM 05: Proceedings of the Fifth IEEE International Conference on Data Mining*, 809-812. IEEE Computer Society.

Yeh, C.H., Deng, H., Chang, Y.H. (2000). Fuzzy Multicriteria Analysis for Performance Evaluation of Bus Companies. *European Journal of Operational Research*, 126, 459-473.

Yilmaz, V., Ari, E. (2017). The Effects of Service Quality, Image, and Customer Satisfaction on Customer Complaints and Loyalty in High-Speed Rail Service in Turkey: A Proposal of the Structural Equation Model. *Transportmetrica A: Transport Science*, 13(1), 67-90.

Yuan, K.H. (2005). Fit Indices Versus Test Statistics. *Multivariate Behavioral Research*, 40(1), pp. 115-148.

Zakaria, Z., Hussin, Z.H., Batau, M.F.A., Zuriawati, Z. (2010). Service Quality of Malaysian Public Transport: A Case Study in Malaysia. *Cross-Cultural Communication*, 6(2), 84-92.

Zeithaml, V.A. (1988). Consumer Perceptions of Price, Quality and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, 52, 2-22.

Zeithaml, V.A., Berry, L.L., Parasuraman, A. (1993). The Nature and Determinants of Customer Expectations of Service. *Journal of the Academy of Marketing Science*, 21(1), 1-12.



CAPÍTULO 9.

ANEXOS

Anexo 1.

ENCUESTA CONDUCTIDA EN METRO DE SEVILLA

En este anexo está recogida la encuesta on-line del proyecto “Mejora de la calidad del transporte público para fomentar la movilidad sostenible: Metro de Sevilla (G-GI3002/IDIK).

Encuesta sobre la Línea 1 del Metro de Sevilla

La Junta de Andalucía, la Universidad de Granada, y la Unión Europea están trabajando juntos para mejorar su metro y la calidad de su viaje.

¡Su opinión es muy importante para nosotros!

Se realizará un **SORTEO** de dos iPad Mini (Wifi, 16Gb) entre todos los que cumplimenten la siguiente encuesta.

La fecha límite para realizar la encuesta y participar en el sorteo será hasta el **15 de Junio del 2014**.

Si tiene alguna duda póngase en contacto con nosotros a través de la siguiente dirección de correo electrónico:
transportes@ugr.es

¡Muchísimas gracias por su colaboración!



Encuesta sobre la Línea 1 del Metro de Sevilla

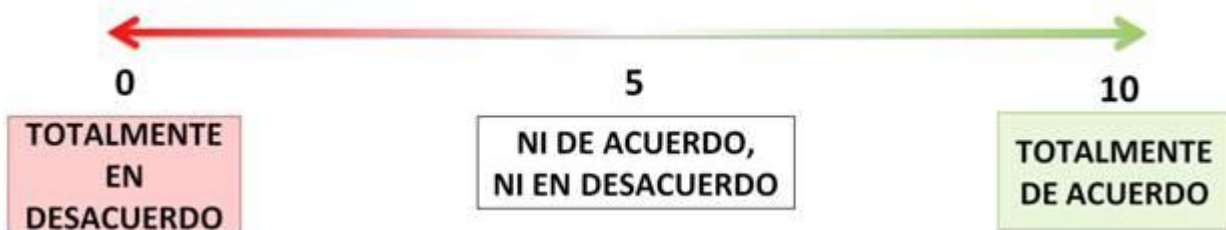
Por favor introduzca el CÓDIGO DE ENCUESTA que aparece en la tarjeta que le fue entregada en el Metro de Sevilla

Indique código:

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



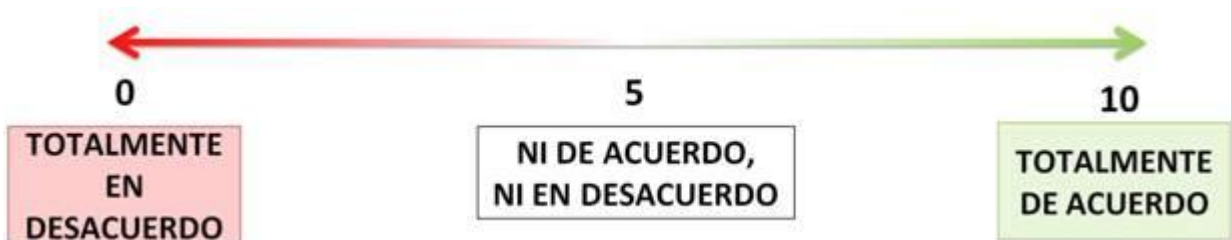
SATISFACCIÓN

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Me atrae viajar en metro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Me siento cómodo viajando en metro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El servicio de metro suministrado cumple con mis expectativas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



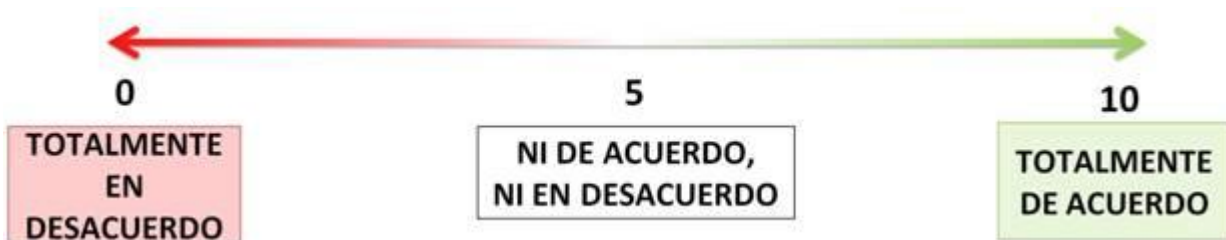
COSTOS PERCIBIDOS

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Considero que el tiempo de espera en los andenes es excesivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Las estaciones se encuentran alejadas de mi origen y/o destino.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El precio me parece alto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Creo que el precio del billete supera ampliamente los costes del metro (personal, electricidad, mantenimiento, etc.).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Considero altos los costes de viajar en metro (tiempo, dinero y confort).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



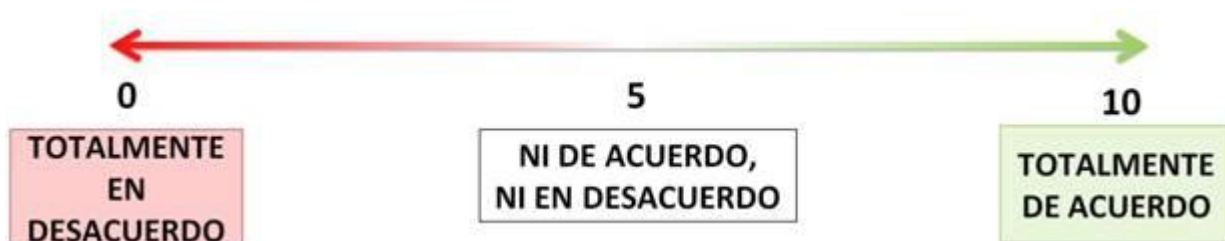
BENEFICIOS PERCIBIDOS

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Me gusta el metro porque me lleva rápido.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
La atención al cliente es buena.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El servicio es bueno.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Me parece adecuada la relación calidad-precio.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
El horario satisface mis necesidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



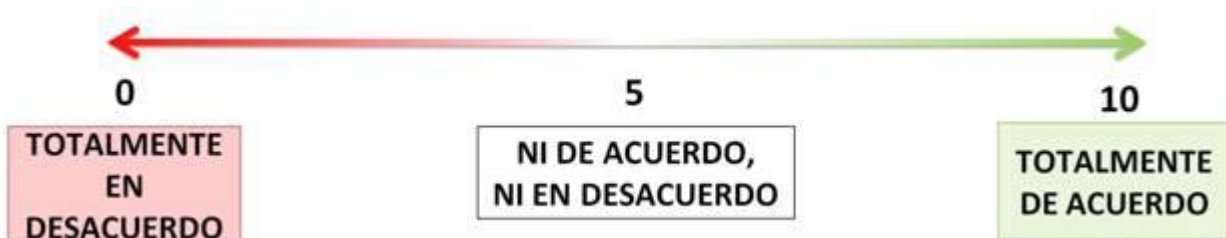
ALTERNATIVAS ATRACTIVAS

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Considero que existen buenas alternativas de transporte público al metro (como el autobús, taxi, etc.).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Prefiero el metro antes que otros transportes público (como el autobús, el taxi o préstamo de bicicletas).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pienso que otros modos de transporte (como el coche, bicicleta, autobús, taxi, etc.) ofrecen más ventajas que el metro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Da igual el modo de transporte empleado, mientras cumpla mis necesidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



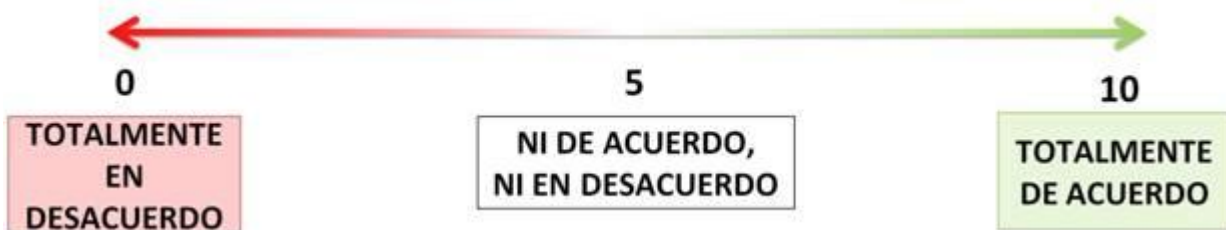
INTENCIÓN DE RECOMPRA

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Volveré a utilizar el metro con las mismas condiciones (dinero, tiempo y confort).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Suelo recomendar a otros que usen el metro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Seguro que volveré a utilizar el servicio de metro.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

VALORE CUÁL ES EL GRADO DE ACUERDO O DESACUERDO RESPECTO A LAS SIGUIENTES AFIRMACIONES (0 TOTALMENTE EN DESACUERDO Y 10 TOTALMENTE DE ACUERDO).



ACTITUD GENERAL HACIA EL TRANSPORTE PÚBLICO:

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Siento que usar el transporte público es acorde con mi estilo de vida.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Me gusta que la gente sepa que utilizo el transporte público.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Me gusta la gente que utiliza el transporte público.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
No importa cuál sea el motivo de viaje, que siempre prefiero coger el transporte público.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Siento que viajando en transporte público colaboro a proteger el medio ambiente.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Creo que usar el transporte público influye en la opinión de la gente sobre mí.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

PARTE A: ACTITUDES HACIA EL SERVICIO DE METRO

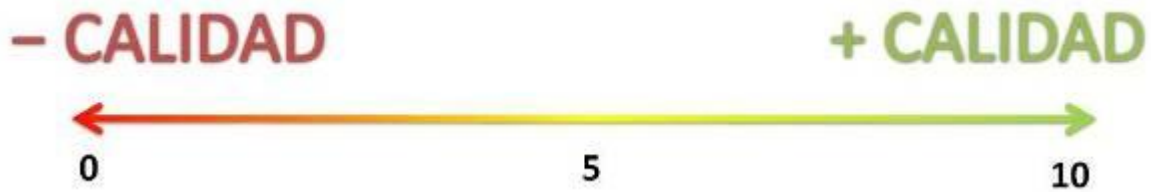
¿Podría indicarme, en una escala del 1 al 5 (siendo 1 la nota más baja y 5 la nota máxima), su grado de satisfacción con el funcionamiento del metro?

- 1 2 3 4 5 NS/NC

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



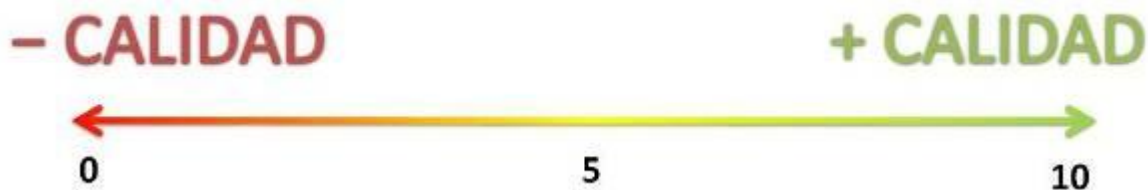
DISPONIBILIDAD

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Proximidad de las paradas a su origen y/o destino.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regularidad del servicio (Ausencia de interrupciones del servicio por averías o incidencias).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Horas de funcionamiento del servicio.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Número de servicios diarios (Frecuencia del servicio).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



INFORMACIÓN

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Información actualizada, precisa y fiable A BORDO (horarios, paradas, interrupciones en el servicio, etc.).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Información disponible a través de otros sistemas de comunicación (como internet, teléfono o aplicaciones móviles).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Información actualizada, precisa y fiable EN LAS ESTACIONES (tarifas, horarios, paradas, interrupciones en el servicio, etc.).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Paneles informativos y direccionales simples y claros en las estaciones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

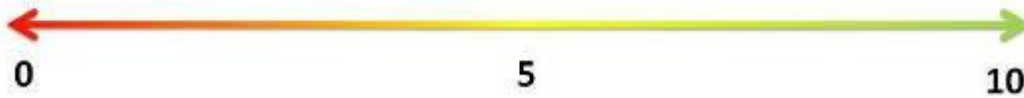
PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



- CALIDAD

+ CALIDAD



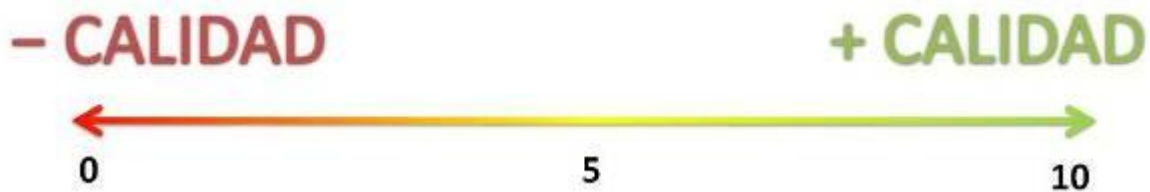
TIEMPO

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Puntualidad.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rapidez en la realización del viaje.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tiempo de espera en el andén.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



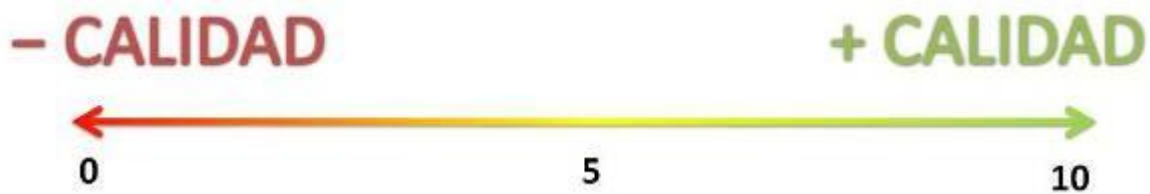
ATENCIÓN AL CLIENTE

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Amabilidad del personal.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Apariencia del personal.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionamiento del "Servicio de atención al cliente" (oficinas, web, reclamaciones, teléfonos, etc.).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eficacia y rapidez del personal para informar, atender y resolver problemas cotidianos de los usuarios.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



CONFORT Y COMODIDAD

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Disponibilidad de cobertura telefónica y 3G en estaciones y a bordo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disponibilidad de asientos en estaciones y andenes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Limpieza de los vehículos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conducción adecuada y no brusca.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nivel de comodidad a bordo de los vehículos (disponibilidad de asientos o viaje de pie sin apreturas).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Limpieza de las estaciones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Iluminación de las estaciones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Temperatura y ventilación de las estaciones y a bordo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Iluminación de los vehículos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

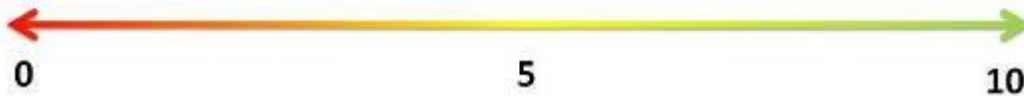
PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



- CALIDAD

+ CALIDAD



SEGURIDAD

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Sensación de seguridad frente a accidentes cuando viaja en el metro (impacto/descarrilamiento).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sensación de seguridad frente a robos y violencia en las estaciones y a bordo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sensación de seguridad frente a resbalones, caídas, atrapamiento por puertas y escaleras mecánicas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Señalización de las "Salidas de emergencia" y extintores.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

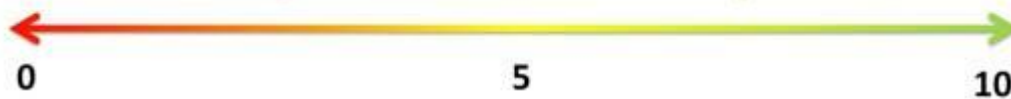
PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

UTILIZANDO UNA ESCALA DEL 0 AL 10 (0-VALOR MÁS BAJO Y 10-VALOR MÁS ALTO), PUNTÚE EL NIVEL DE CALIDAD PERCIBIDO DEL SERVICIO DEL METRO DE SEVILLA, RESPECTO A:



**MUY ALTOS
(- CALIDAD)**

**MUY BAJOS
(+ CALIDAD)**



CONTAMINACIÓN AMBIENTAL

	0	1	2	3	4	5	6	7	8	9	10	NS/NC
Niveles de ruido a bordo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Niveles de ruido en estaciones.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Niveles de vibración a bordo.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE B: SU OPINIÓN SOBRE LA CALIDAD PERCIBIDA EN EL USO DEL METRO DE ...

¿Cuál es su VALORACIÓN GLOBAL sobre el servicio de la Línea 1 del Metro de Sevilla (puntuándolo del 0 al 10, siendo 0 la nota más baja y 10 la nota más alta)?

- 0 1 2 3 4 5 6 7 8 9 10 NS/NC

PARTE C: SU VIAJE

CENTRANDO SU ATENCIÓN EN EL DÍA Y EN EL VIAJE QUE ESTABA REALIZANDO CUANDO RECIBIÓ LA TARJETA, RESPONDA A LAS SIGUIENTES CUESTIONES

¿Por qué utilizaba el metro en esa ocasión? (PUEDE MARCAR MÁS DE UNA OPCIÓN)

- | | | |
|-------------------------------------|--|--|
| <input type="checkbox"/> Precio | <input type="checkbox"/> Razones ecológicas | <input type="checkbox"/> Falta de aparcamiento |
| <input type="checkbox"/> Comodidad | <input type="checkbox"/> No dispongo de carnet de conducir | <input type="checkbox"/> Congestión del tráfico |
| <input type="checkbox"/> Rapidez | <input type="checkbox"/> No dispongo de vehículo | <input type="checkbox"/> Mi vehículo no lo puedo usar por algún motivo |
| <input type="checkbox"/> Frecuencia | <input type="checkbox"/> Es mi única alternativa | <input type="checkbox"/> Otra razón |

¿Cuál era el motivo principal de su desplazamiento cuando recibió la tarjeta?:

- Trabajo
- Estudios
- Ocio
- Otros

¿Cuál o cuáles de los siguientes aparcamientos que se ofertan en las estaciones de metro usa Vd.?. (PUEDE MARCAR MÁS DE UNA OPCIÓN)

- De coches
- De motos
- De bicicletas
- Ninguno

PARTE C: SU VIAJE

CENTRANDO SU ATENCIÓN EN EL DÍA Y EN EL VIAJE QUE ESTABA REALIZANDO CUANDO RECIBIÓ LA TARJETA, RESPONDA A LAS SIGUIENTES CUESTIONES

¿Cómo llegó desde su origen a la estación de metro?

- A pie/Andando
- Bicicleta
- Autobús urbano (Tussam)
- Autobús interurbano
- Coche
- Motocicleta/Ciclomotor
- Otro

¿Cuál fue la duración aproximada (en minutos) de su viaje desde el origen a la estación de metro?:

Indique la duración:

Desde la estación de metro en la que se baja, ¿cómo llegó a su destino final?

- A pie/Andando
- Bicicleta
- Autobús urbano (Tussam)
- Autobús interurbano
- Coche
- Motocicleta/Ciclomotor
- Otro

Desde la estación en la que se baja, aproximadamente, ¿cuánto tiempo tardó en llegar a su destino final?:

Indique la duración:

¿Cuál fue la DURACIÓN TOTAL (en minutos) aproximada de su viaje desde el origen a su destino? (Puerta a puerta): tiempo hasta estación + tiempo en metro + tiempo hasta destino final

Indique tiempo total:

Encuesta sobre la Línea 1 del Metro de Sevilla

PARTE C: SU VIAJE

CENTRANDO SU ATENCIÓN EN EL DÍA Y EN EL VIAJE QUE ESTABA REALIZANDO CUANDO RECIBIÓ LA TARJETA, RESPONDA A LAS SIGUIENTES CUESTIONES



Bonometro



Tarjeta del consorcio

Indique el tipo de billete que usted estaba utilizando cuando recibió la tarjeta:

- Billeto sencillo
- Billeto de ida y vuelta
- Bono de 1 día
- Bonometro
- Bono plus 45
- Tarjeta del consorcio
- Otros

¿Cuál es la frecuencia con la que usted utiliza el metro?

- Diariamente (más de 4 días a la semana)
- 3 o 4 veces por semana
- 1 o 2 veces por semana
- Esporádicamente

PARTE C: SU VIAJE

CENTRANDO SU ATENCIÓN EN EL DÍA Y EN EL VIAJE QUE ESTABA REALIZANDO CUANDO RECIBIÓ LA TARJETA, RESPONDA A LAS SIGUIENTES CUESTIONES

Indique cuál o cuáles de las siguientes medidas considera Vd. que podría mejorar la oferta de aparcamientos de bicicletas en las estaciones de metro:

- Colocar cubiertas para protegerlos de la lluvia
- Colocarlos en un lugar cerrado
- Poner servicios de préstamo como el Bus+Bici
- Poner cámaras de videovigilancia
- Ninguna me parece útil
- NS/NC

Si no pudiera utilizar el metro, ¿qué otro medio de transporte usaría?:

- A pie/Andando
- Bicicleta
- Autobús urbano (Tussam)
- Autobús interurbano
- Coche/Automóvil
- Motocicleta/Ciclomotor
- Tranvía
- Combinación de vehículos
- Otros

PARTE D: SOBRE USTED

PREGUNTAS RELACIONADAS CON SUS CARACTERÍSTICAS COMO USUARIO DEL METRO DE SEVILLA:

Género:

- Mujer
- Hombre

¿Posee algún tipo de minusvalía?

- No
- Temporalmente
- Permanente

Usted tiene...? (POSIBILIDAD DE RESPUESTA MÚLTIPLE)

- Carnet de conducir
- Acceso a coche propio
- Acceso a motocicleta o ciclomotor
- Acceso a bicicleta
- Ninguna de las anteriores

¿En qué franja de edad se encuentra?

- Menos de 18
- 18 a 25
- 26 a 40
- 41 a 65
- 66 o más
- NS/NC

Máximo nivel de estudios finalizado:

- Sin estudios
- Educación obligatoria (ESO, EGB, etc.)
- Bachiller o ciclo formativo de grado medio
- Título universitario (o superior)
- NS/NC

PARTE D: SOBRE USTED

PREGUNTAS RELACIONADAS CON SUS CARACTERÍSTICAS COMO USUARIO DEL METRO DE SEVILLA:

Situación laboral actual del encuestado:

- Empleado por cuenta propia
- Empleado por cuenta ajena
- Desempleado
- Jubilado
- Estudiante
- Ama/o de casa
- Otro

Número de personas que habitan en su hogar (incluido usted):

Nº de personas:

¿Cuál es su código postal?

PARTE D: SOBRE USTED

PREGUNTAS RELACIONADAS CON SUS CARACTERÍSTICAS COMO USUARIO DEL METRO DE SEVILLA:

Nivel de ingresos mensuales netos de la unidad familiar:

- Menos o igual a 1.200 €
- Entre 1.201 - 1.800 €
- Entre 1.801 - 2.400 €
- 2.401 € o más
- NS/NC

Encuesta sobre la Línea 1 del Metro de Sevilla

FIN DE ENCUESTA

Ha finalizado la encuesta. ¡¡Muchas gracias por su colaboración!!

Los resultados del sorteo aparecerán en la página web del grupo TRYSE de la Universidad de Granada (<http://tryse.ugr.es/>) a partir del día 28 de julio de 2014. El ganador del premio se anunciará con el **CÓDIGO DE ENCUESTA que introdujo al principio de la misma y que aparece en la tarjeta que le ha sido entregada.**

Para participar en el sorteo, y poder identificar correctamente al premiado, debe introducir en el siguiente recuadro las últimas cuatro cifras y letra de su NIF (Ejemplo: Si su NIF fuera 25344903-F, debería introducir en el siguiente recuadro 4903-F):

En el caso de ser premiado, para poder recoger dicho premio deberá presentar la tarjeta entregada con el **CÓDIGO DE ENCUESTA agraciado (cuyo número ya ha introducido al inicio de esta encuesta) y su DNI para la correspondiente identificación. !!!Conserve la tarjeta entregada porque puede valer un iPad!!!**

Si tiene alguna duda póngase en contacto con nosotros en la siguiente dirección de correo electrónico:

transportes@ugr.es

Si no desea estar atento a la página web y quiere que le informemos personalmente en caso de que resulte premiado, puede introducir a continuación su correo electrónico o un número de teléfono de contacto.

MUCHAS GRACIAS POR SU TIEMPO

Anexo 2.

RESULTADOS TEST CHI-CUADRADO

A continuación son mostrados los resultados de los test Chi-Cuadrado realizados para cada caso de estudio.

CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA)

		Disp		Info		Eq_Tan	
		Chi	Grados de libertad	Chi	Grados de libertad	Chi	Grados de libertad
Acces	Modelo con correlación	643,751	43	293,959	34	452,266	53
	Modelo sin correlación	1.979,245	44	1.834,216	35	2.074,43	54
	Diferencia	1.335,494	1	1.540,257	1	1.622,164	1
		<i>Significativo</i>		<i>Significativo</i>		<i>Significativo</i>	

		Info		Eq_Tan	
		Chi	Grados de libertad	Chi	Grados de libertad
Disp	Modelo con correlación	567,915	26	832,25	43
	Modelo sin correlación	1.841,782	27	2.127,115	44
	Diferencia	1.273,867	1	1.294,865	1
		<i>Significativo</i>		<i>Significativo</i>	

CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO

		Sat		CdS	
		Chi	Grados de libertad	Chi	Grados de libertad
Ben	Modelo con correlación	529,49	13	1.356,872	43
	Modelo sin correlación	2.173,152	14	3.358,214	44
	Diferencia	1.643,662	1	2.001,342	1
		<i>Significativo</i>		<i>Significativo</i>	

		Sat	
		Chi	Grados de libertad
CdS	Modelo con correlación	1742,243	53
	Modelo sin correlación	3520,249	54
	Diferencia	1778,006	1
		<i>Significativo</i>	

		Sat	
		Chi	Grados de libertad
Int_Rec	Modelo con correlación	583,879	13
	Modelo sin correlación	1688,816	14
	Diferencia	1104,937	1
		<i>Significativo</i>	

Anexo 3.

PUNTUACIÓN ARCOS DE LAS REDES BAYESIANAS

Se presentan las tablas con las puntuaciones obtenidas de acuerdo a las ocurrencias de los arcos en cada una de las Redes Bayesianas simuladas.

CASO DE ESTUDIO 1. APLICACIÓN DE LA METODOLOGÍA DE DOS ETAPAS PARA ESTUDIAR LA CALIDAD DEL SERVICIO EN EL METRO DE SEVILLA (ESPAÑA)

Padre	Hijo	Grow-Shrink	Increm. Asso.	Inte. Incre. Asso.	2-Phase Restrict. Max.	Hill Climbing (AIC)	Hill Climbing (BIC)	Hill Climbing (Bde)	Hill Climbing (mBde)	Hill Climbing (K2)	Hill Climbing (LogLik)	TABU (AIC)	TABU (BIC)	TABU (Bde)	TABU (mBde)	TABU (K2)	TABU (LogLik)	Max-Min Hill Climbing	Total
CS	Acces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Disp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	At_Cl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Med_Am	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Esp_Ind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Info	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Seg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CS	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acces	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Acces	Disp	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
Acces	At_Cl	0	0	0	0	1	0	1	1	1	1	1	0	1	1	1	1	0	10
Acces	Med_Am	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Acces	Esp_Ind	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Acces	Info	0.5	0.5	0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15.5
Acces	Seg	0	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	0	8
Acces	Eq_Tan	0	0	0	0	1	1	1	1	0	1	1	1	1	1	0	1	0	10
Disp	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
Disp	Acces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Disp	At_Cl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Disp	Med_Am	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	4
Disp	Esp_Ind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Disp	Info	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5

Disp	Seg	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Disp	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At_Cl	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
At_Cl	Acces	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	3
At_Cl	Disp	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
At_Cl	Med_Am	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At_Cl	Esp_Ind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At_Cl	Info	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
At_Cl	Seg	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	4
At_Cl	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Med_Am	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	17
Med_Am	Acces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Med_Am	Disp	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Med_Am	At_Cl	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Med_Am	Esp_Ind	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	3
Med_Am	Info	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Med_Am	Seg	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Med_Am	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Esp_Ind	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
Esp_Ind	Acces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Esp_Ind	Disp	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Esp_Ind	At_Cl	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Esp_Ind	Med_Am	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Esp_Ind	Info	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Esp_Ind	Seg	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Esp_Ind	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Info	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
Info	Acces	0.5	0.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.5

Info	Disp	0	0	0.5	0	1	1	1	1	1	1	1	1	1	1	1	1	0	12.5
Info	At_Cl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Info	Med_Am	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Info	Esp_Ind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Info	Seg	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Info	Eq_Tan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seg	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
Seg	Acces	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seg	Disp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seg	At_Cl	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seg	Med_Am	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0	0	10
Seg	Esp_Ind	0	0	0	0	1	0	1	1	1	0	1	0	1	1	1	0	0	8
Seg	Info	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seg	Eq_Tan	0.5	0.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.5
Eq_Tan	CS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
Eq_Tan	Acces	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	2
Eq_Tan	Disp	0	0	0	0	1	0	1	1	1	1	1	0	1	1	1	1	0	10
Eq_Tan	At_Cl	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Eq_Tan	Med_Am	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Eq_Tan	Esp_Ind	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	12
Eq_Tan	Info	0	0	0	0	1	0	1	1	1	1	1	0	1	1	1	1	0	10
Eq_Tan	Seg	0.5	0.5	0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15.5

CASO DE ESTUDIO 2. APLICABILIDAD DE LA METODOLOGÍA DE DOS ETAPAS A OTROS ÁMBITOS DE CONOCIMIENTO: ANÁLISIS DE LAS ACTITUDES DE LOS PASAJEROS HACIA EL SERVICIO DE METRO

Padre	Hijo	Grow-Shrink	Increm. Asso.	Inte. Incre. Asso.	2-Phase Restrict. Max.	Hill Climbing (AIC)	Hill Climbing (BIC)	Hill Climbing (Bde)	Hill Climbing (mBde)	Hill Climbing (K2)	Hill Climbing (LogLik)	TABU (AIC)	TABU (BIC)	TABU (Bde)	TABU (mBde)	TABU (K2)	TABU (LogLik)	Max-Min Hill Climbing	Total
CdS	Act_Tp	0	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1	0	11
CdS	Cost_Per	0.5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2.5
CdS	Int_rec	0.5	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	13.5
CdS	Ben	0	0.5	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	13.5
CdS	Sat	0	0	0	0	1	0	1	1	1	1	0	0	1	1	1	1	0	9
Act_Tp	CdS	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Act_Tp	Cost_Per	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Act_Tp	Int_rec	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Act_Tp	Ben	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Act_Tp	Sat	0	0.5	0.5	0	0	0	0	0	0	1	0	0	0	0	0	1	1	4
Cost_Per	CdS	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
Cost_Per	Act_Tp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cost_Per	Int_rec	0.5	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3
Cost_Per	Ben	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	3
Cost_Per	Sat	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Int_rec	CdS	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
Int_rec	Act_Tp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Int_rec	Cost_Per	0.5	0	0.5	1	1	1	1	1	1	1	1	1	1	1	1	1	0	14
Int_rec	Ben	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	3
Int_rec	Sat	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	2
Ben	CdS	0	0.5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1.5
Ben	Act_Tp	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Ben	Cost_Per	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	4
Ben	Int_rec	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Ben	Sat	0.5	0	0	1	1	1	1	1	1	0	1	1	0	0	1	0	0	9.5
Sat	CdS	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Sat	Act_Tp	0	0.5	0.5	0	1	0	0	0	1	0	1	0	0	0	1	0	0	5
Sat	Cost_Per	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sat	Int_rec	0	0	0	0	1	0	1	1	1	0	1	0	1	1	1	0	0	8
Sat	Ben	0.5	0	0	0	0	0	0	0	0	1	0	0	1	1	0	1	0	4.5

Anexo 4.

ARTÍCULO 1. DIEZ-MESA ET AL. (2018)

Se muestra el manuscrito del siguiente artículo publicado en revista científica de impacto:

Diez-Mesa, F., De Oña, R., De Oña, J. (2018). Bayesian Networks and Structural Equations Modelling to Develop Service Quality Models: Metro of Seville Case Study. Transportation Research Part A: Policy and Practice, 118, 1-13. <https://doi.org/10.1016/j.tra.2018.08.012>

Bayesian Networks and Structural Equation Modelling to develop service quality models: Metro of Seville case study

ABSTRACT

For many years it has been desirable to improve Service Quality (SQ) in Public Transport (PT) because of its strong influence on user satisfaction and its importance in attracting new passengers. Structural Equation Modelling (SEM) is one of the most widely used techniques for analysing SQ due to its ability to address different kinds of variables and to model a whole phenomenon occurring at one time. Nevertheless, its confirmative nature requires the presence of previous knowledge, a hurdle that can be overcome by applying Bayesian Networks (BN) as a technique that learns directly from data without any pre-assumptions. The aim of this paper is to apply a novel methodological approach in the field of SQ, based on a two-step process combining the techniques of BN and SEM, to model SQ in the Metropolitan Light Rail Transit (LRT) Service of Seville (Spain). The methodological approach proposed in this paper has been applied to discover and confirm the possible relationships between the LRT service characteristics and how they are related with passengers' overall perception of SQ directly from data and without the need to make assumptions. A BN was automatically learnt from the data and allowed to establish relationships between various SQ dimensions describing the service. SEM then checked the SQ model and the relationships between the dimensions extracted from the BN. The SEM model fit parameters and its consistency with the real life expected scenario supported and validated the proposed SQ model. Furthermore, the different relationships between the dimensions extracted from the BN were found to support the usefulness and potential of this methodological process for the development and confirmation of new theories and models in any field of knowledge based on data and expert supervision.

Keywords: Structural Equation Modelling, Bayesian Networks, Service Quality, Public Transport.

1. Introduction

Over recent years the study of Service Quality (SQ) and its improvement has become a relevant and crucial factor in many fields. Companies, operators and governments have focused on providing high levels of quality in their services in order to improve the current level of customer satisfaction and to attract new clients. It is important to highlight that this relationship between SQ and Satisfaction has been tested in a variety of fields, such as Marketing (Grönroos, 1984), Tourism (Shonk and Chelladurai, 2008), and On-line (Ho and Lee, 2007), etc.

In the specific field of Public Transport (PT), governments and operators have paid great attention to the study and analysis of SQ from the perspective of passengers (Andreassen, 1995; de Oña et al., 2015; Dell’Olio et al., 2011 a; Nathanail, 2008; Woods and Masthoff, 2017; etc.). Their main purpose, apart from addressing passenger satisfaction, has been to improve the ability of PT to compete with the private car by finding alternative modes of transport able to combat problems of mobility, traffic jams and pollution, etc. (Beirao and Cabral, 2007; de Oña and de Oña, 2015; Linda, 2003).

A wide range of techniques applied to the study of SQ in PT can be found in the literature (e.g., Celik et al., 2013; Islam et al., 2016; Kuo, 2011; etc.). Their main application has been to investigate the influence different service quality factors have on overall SQ. Such knowledge allows administrations and operators to concentrate their efforts and investment in specific areas in need of improvement or of greatest importance to the passengers.

Examples of such techniques are: Importance-Performance Analysis in Weinstein (2000); A Composite Index in de Oña et al. (2016a); Multinomial Logit models in Eboli and Mazzulla (2008); the VIKOR method in Kuo and Liang (2011); etc. Structural Equation Modelling (SEM) has, nevertheless, been one of the more widely used techniques and its application in this field has grown over recent years (e.g. Amin and Isa, 2008; Chen, 2008; Hapsari et al., 2017; Yang et al., 2012; Yilmaz and Ari, 2017; etc.). The main reason behind the increased interest in SEM is its ability to easily address large numbers of variables, both endogenous and exogenous, as well as latent variables (not observed variables) explained as a linear combination of observed variables (Golob, 2003). Indeed, SEM is generally considered to be one of the best integrated methods for measuring latent variables and assessing their structural relationships (Chiou and Chen, 2012; de Oña et al., 2015).

These characteristics are all crucial in the study of SQ as it has been defined as such a complex, fuzzy and abstract concept (Carman, 1990; Parasuraman et al., 1985). SQ also depends on a series of underlying observed and unobserved variables. These unobserved variables are commonly denominated as dimensions and are used to provide a better understanding of how customers perceive various service attributes (de Oña et al., 2013).

The suitability of using SEM to study SQ is, therefore, more than justified. Nonetheless, due to the confirmative nature of this technique (de Oña and de Oña., 2015; Golob, 2003), its use requires previous knowledge about how the different dimensions of the SQ models are related. This means that the bibliography or expert knowledge is usually required to develop the models (e.g., Bagozzi, 1994; Eboli and Mazzulla, 2012; Fillone et al., 2005; Kamaruddin et al., 2012; etc.) and, depending on the context, this could cause users to miss certain important relationships in the explanation of SQ.

The above has not usually been considered in the specific case of studying SQ in PT (de Oña et al., 2017), where the scientific community agrees that SQ is directly related to and influenced by all its dimensions. Nevertheless, there is evidence in the bibliography to suggest that SQ dimensions could be interrelated and influence each other, and that their relationship with overall SQ or user satisfaction is not always direct but sometimes indirect. Evidence of this can be found in different ways (i.e., directly or indirectly) of relating the dimensions in the PT SQ models found in the literature (e.g., Chou and Kim, 2009; de Oña et al., 2017; Eboli and Mazzulla, 2012; Rahman et al., 2016; etc.). For example, de Oña et al. (2017) grouped service attributes into two latent dimensions: primary attributes (transport service factors), and secondary attributes (comfort and convenience factors). They showed that the secondary factor dimension exerted an effect on the primary attribute dimension, and its relationship with user satisfaction was not direct, but rather indirect through the primary factors dimension.

To overcome this handicap, this paper proposes a two-step procedure based on the combined use of Bayesian Networks (BN) and SEM in a single methodology, as yet unrecorded in the field of SQ. The combined use of BN and SEM has already been applied in the health sector with outstanding results (e.g., Duarte et al., 2011; Scheines et al., 1999; Trentini et al., 2015). Moreover, Kenett and Salini (2011) have presented Bayesian networks and latent variable models as suitable methodologies for analysing customer surveys.

The first step in this approach is to use BN to find certain hypotheses about the relationships between the dimensions which define SQ, followed by a second step which applies SEM to validate these relationships and the model as a whole.

Both techniques can be considered as being complementary due to the different characteristics inherent to each one. On the one hand, BN is an exploratory technique which learns directly from data without the need for pre-assumptions (Heckerman, 1998). Thus, using BN solves the problem of requiring previous knowledge about dimensional relationships. On the other hand, SEM is a confirmatory technique which allows the modelling of a phenomenon in which a set of unidirectional effects or relationships between observed and unobserved variables are established by researchers (de Oña et al., 2015; Golob, 2003). Moreover, this technique examines more than one relationship and tests a set of hypotheses considering a large amount of information at the same time (de Oña et al., 2013; Hair et al., 2010). Therefore, SEM allows researchers to check and validate the relationships of the SQ framework extracted from the BN.

Thus, the proposed two-step methodological approach is appropriate for studying SQ. This is particularly relevant in the field of PT, due to the probable existence of several relationships between the SQ dimensions which might not be discovered and tested by other means.

The proposed research presented here is oriented towards the study of SQ in the Metropolitan Light Rail Transit Service (LRT) of Seville (Spain) by applying this two-step methodology. The main goal is to achieve a model which explains the SQ of this PT service from the point of view of passengers by using data from a Customer Satisfaction Survey (CSS) carried out on 2014.

This paper is structured as follows: Section 2 presents the PT case study, the data collection procedure and the sample characteristics. Section 3 describes the BN and SEM approaches and the two-step methodological process. The results obtained are explained and discussed in Section 4. Finally, the conclusions are reported in Section 5.

2. Methodology

This section describes all the processes that have been applied in this study. First, BN and SEM techniques are defined, followed by a detailed analysis of the two-step methodological process.

2.1. Bayesian Network (BN)

A BN is a framework for reasoning under uncertainty, and is widely used for representing uncertain knowledge (Pearl, 1988). This is a data mining technique with a wide range of advantages, one of the most important of which is that it makes complex problem analysis easy to understand as the interrelationships and dependencies of the model parameters become visible (Hänninen, 2008).

BN can be described according to two different terms. Firstly, in terms of a quantitative component consisting of a joint probability distribution that factorizes into a set of conditional probability distributions, governed by the structure of a Directed Acyclic Graph (DAG):

Let $U = \{x_1, \dots, x_n\}$, $n \geq 1$ be a set of variables. A BN over a set of variables U is a network structure, which is a DAG over U and a set of probability tables (1):

$$B_p = \{p(x_i | pa(x_i), x_i \in U)\} \quad (1)$$

where $pa(x_i)$ is the set of the antecedents of x_i in BN and $i = 1, 2, 3, \dots, n$.

A BN represents joint probability distributions (2):

$$P(U) = \prod_{x_i \in U} p(x_i | pa(x_i)) \quad (2)$$

Secondly, BN can also be described in terms of a qualitative component, consisting of a DAG. In other words, BN can be defined as graphic models of the interactions between a set of variables, where the variables are represented as the nodes of a graph and the interactions (direct dependences) as directed links (also known as arcs and edges) between the nodes.

In the model, the nodes can be connected, which shows direct dependence between them, and, in the opposite case, the nodes can be not connected, which shows independence among them. A particular nomenclature for these situations of dependence/independence between nodes is: nodes with arrows directed into them are called “child”, while the nodes from which the arrows depart are called “parent”.

Each node contains the states of the random variable and it represents a conditional probability table. The conditional probability table of a node contains the probability of the node being in a specific state, given the states of its parents.

For a detailed description of BN the reader is directed to Kjaerulff and Madsen (2008).

2.1.1. Bayesian Networks learning and the scoring metric

The first step in a BN analysis is to learn both the structure, which has to be validated, and the parameters. There are two main approaches to learning structures in BN:

- Automatic Learning Approach: The structure of the BN is provided by algorithms which learn the structure using only one database. There are three main approaches to structural learning in BN (De Oña et al., 2011):
 - a) Constraint based: Tests of conditional independence are performed on the data, and a search is made for a network that is consistent with the observed dependencies and independencies.
 - b) Scored based: A score that evaluates how well the dependencies or independencies in a structure match the data is defined and a search is made for a structure that maximizes the score.
 - c) Hybrid: Combines aspects of both constraint-based and score-based algorithms, as they use conditional independence tests and network scores at the same time.

Note that before learning the BN by applying any of the previous approaches, mandatory relationships or independence can be imposed between the considered variables.

- Manual Learning Approach: The structure of a BN is provided by human expertise. This can be a highly labour intensive task, requiring a great deal of skill and creativity as well as close communication with problem-domain experts (Kjaerulff and Madsen, 2008).

In this paper, the methodology of automatic learning is used due to its multiple advantages, an example being the possibility of combining previous knowledge with automatic learning from data or finding new and unknown relationships between the variables. In addition, the purpose of using BN in this research is to achieve a model (i.e., structure) directly from data.

To apply this technique and to get the most robust BN, we have relied on the methodology explained in Cugnata et al. (2016). It lies in the application of a wide range of different learning algorithms and in measuring the arc strength using re-sampling techniques. The most robust structure reappears with specific arcs in the most learned networks and can then be chosen.

2.2. Structural Equation Modelling (SEM)

SEM can be defined as a multivariate technique combining regression, factor analysis, and variance analysis to simultaneously estimate interrelated dependence relationships. If the latent variables are classified as endogenous (dependent) and/or exogenous (independent) variables, the two components of SEM can be defined. The first is called the structural model and describes the relationships between endogenous and exogenous variables, showing the direction and strengths of the relationships between the latent variables. The second is called a measurement model and assesses the relationships between latent and observed variables.

The basic SEM equation can be defined as in Bollen (1989).

$$\eta = B\eta + \Gamma\xi + \zeta \quad (3)$$

In which $\eta = m \times 1$ is a vector of the latent variables; $\xi = n \times 1$ is a vector of the latent exogenous variables; $B = m \times m$ is the matrix of the coefficients associated with the latent endogenous variables; $\Gamma = m \times n$ a matrix of the coefficients associated with the latent exogenous variables; $\zeta = m \times 1$ a vector of error terms associated with the endogenous variables.

The basic equations of the measurement model can be expressed as (5) and (6).

$$x = \Lambda_x \xi + \delta \quad (4)$$

$$y = \Lambda_y \eta + \varepsilon \quad (5)$$

In which $x =$ column q -vector related to the observed exogenous variables; $\delta =$ column q -vector related to the observed exogenous errors; $\Lambda_x = q \times n$ structural coefficient matrix for the effects of the latent exogenous variables on the observed variables; $y =$ column p -vector related to the observed endogenous variables; $\varepsilon =$ column p -vector related to the observed endogenous errors and $\Lambda_y = p \times m$ is a structural coefficient matrix for the effects of the latent endogenous variables on the observed ones.

Different methods can be used to estimate the parameters of the model (e.g., maximum likelihood, weighted and un-weighted least squares, generalized least squares, etc.). All of them have similar goals to minimize the differences between the predicted variance-covariance matrix of the variables in the model and the observed variable, while respecting the constraints of the model. Therefore, in order to select the most suitable, each case must be analysed by focusing on the probability distribution, the scale properties of the variables, the complexity of the SEM, and the sample size (Golob, 2003). Moreover, SEM is a confirmatory rather than an exploratory technique because the researcher constructs the model by defining unidirectional effects between variables (Golob, 2003), and with this purpose in mind it was used in this study.

2.3. Methodological process: step by step

The methodological process proposed in this paper is based on two crucial steps: 1) knowledge extraction from data; and 2) relationship validation.

Step 1: Knowledge extraction from data

The BN is calibrated to find the possible relationships between the dimensions explaining SQ.

This research applied the process described in Cugnata et al. (2016) to achieve the most robust BN structure. For this purpose, an adaptation of the R script provided in Cugnata et al. (2016) has been applied to our case of study.

This step begins when different network structures are learned. With this end in mind, the data is analysed using 17 algorithms implemented in the R package *bnlearn* (i.e., Constraint-based, Score-based and Hybrid learning algorithm) and encode in the R script of Cugnata et al. (2016). It is important to highlight that this number of algorithms was considered as being sufficient for the present case study in this research, however it could be very different depending on the criteria of the researcher.

The most robust network is then selected. The occurrence of an arc between two nodes at each learned BN is reported and scored following the process described in Cugnata et al. (2016). This process measures the occurrence of each arc in the learned BN which could be interpreted as the arc's strength. Therefore, the most robust BN has the highest number of arcs with the highest number of occurrences. A threshold of 11 was established for the occurrence score, which corresponds to almost 2/3 of occurrence in the total learned BN. An arc is considered to have a high level of occurrence if it has appeared at least 11 times in the algorithms that were implemented. Where two robust BN have the same number of arcs with high occurrence, the one with the lowest misclassification rate is selected.

Where different numbers of algorithms are considered, this threshold will be different depending on where it represents an almost 2/3 occurrence in the total learned BN.

Finally, a bootstrap re-sampling procedure is performed on the initial dataset to analyse the robustness of the chosen network and the extracted relationships. This consists of learning 1,000 BN using the same algorithm (which achieved the most robust BN in the previous step) on 1,000 randomly generated subsets with 1,000 random observations. The proportion of occurrence of each arc in the bootstrap replicates is obtained indicating the robustness of the dimension relationships extracted from the BN.

Step 2: Relationship validation

In this step the relationships between the dimensions identified with the BN are modelled and validated by SEM (*Amos Graphics v.22* software was used). For this purpose, the goodness-of-fit parameters of the structural model were analysed following Hooper et al. (2008). They showed different kinds of indices to determine the model fit (i.e., absolute, incremental and parsimony), and suggested acceptable threshold levels for each one.

The following goodness-of-fit indices were used for this research:

- Absolute fit indices: the chi-squared test, the goodness of fit index (GFI) and the adjusted goodness of fit index (AGFI), the root mean square error of approximation (RMSEA) and the root mean square residual (RMR).
- Incremental fit indices: The normed fit index (NFI) and the comparative fit index (CFI).
- Parsimony fit indices: The parsimony goodness-of-fit index (PGFI) and the parsimonious normed fit index (PNFI).

Moreover, all the relationships in the model are tested in order to be significant at a 0.05 level of confidence. If not, they are removed.

Table 1 shows a summary of the process.

Table 1. Summary of Methodology.

Steps	Content
Step 1: Knowledge extraction from data	- Following Cugnata et al. (2016): a) Learning of different network structures. b) Selection of the most robust network. c) Analysis of the robustness of the chosen network and extraction of the relationships between dimensions.
Step 2: Relationships validation	-Test Structural Model: ➤ Fit indices: - Chi-squared test, GFI, AGFI, RMR, RMSEA. - NFI, CFI. - PGFI, PNFI. ➤ Significance of structural coefficients.

3. Data Collection

The data of this study was gathered from a CSS asked about the LRT Service of Seville (Spain). This service consists of one 18 Km long line and 22 stations which are distributed throughout the city of Seville (690,566 inhabitants in 2016). The survey implementation and data collection were carried out online, via a web-based platform for conducting surveys. Passengers had three weeks (May-June 2014) to complete the online survey providing 3,365 registered responses, of which 3,198 were found to be valid for subsequent analysis.

Different kinds of subjects were covered in the survey (e.g., attitude of the passengers towards the LRT service, travel habits, socioeconomic characteristics, etc.). Nevertheless, this study was mainly focused on the perceptions of passengers about the service characteristics (i.e., the availability of the service, accessibility, information, timeliness, customer service, comfort, safety and environmental pollution). Moreover, overall passenger satisfaction with the service and their global score for SQ were also considered.

The main socioeconomic characteristics and travel habits of the sample showed a distribution between females and males of 53.30% and 46.70%, respectively. Around half of them were aged between 18-25 years (41.70%), followed by 26-40 (28.90%) and 41-65 years (25.60%). Note that in this sample the age groups younger than 18 and older than 65 were under-represented (2.80% and 1.00%, respectively). The main reasons for travelling were studies (38.80%) and work (35.50%), followed by leisure and other reasons (15.30% and 10.30%, respectively). Their frequency of use of this service is daily (52.10%) and, generally, they have a high-school diploma (41.90%) or are university graduates (48.50%). A smaller group is also represented in the sample, those who only have secondary compulsory education (8.40%). Most of the respondents have a low household monthly family income (lower than 1,800 Euros) and, there is not much difference between the percentage of passengers who had a private vehicle available to make the trip and those who did not have one (54.78% and 45.22%, respectively). The answers given for the question "Overall Service Quality" show that the passengers in this sample perceived a suitable level of SQ (average rate of 7.6 and standard deviation of 1.5).

4. Results and discussion

Many variables were considered for the case of the Seville Metro which could have influenced the results of the BN. Therefore, the data base needed to be pre-processed before taking the steps described in the proposed methodology.

4.1. Data pre-processing

The following questions were used as variables in the analysis: Thirty seven questions/SQ attributes which are related to various aspects of the LRT service, such as Availability of the service, Accessibility, Safety, etc.; one question about the overall perceived level of quality with the LRT (i.e., "Overall Service Quality"); and, finally, another question about the overall satisfaction of the passengers (i.e., "Overall Service Satisfaction") (see Table 2). The perceived level of quality of each of the 37 attributes and the "Overall Service Quality" were asked with an 11-point Likert scale (0-lowest quality and 10-highest quality). The "Overall Service Satisfaction" was scored on a 5-point Likert scale (1-lowest satisfaction and 5-highest satisfaction).

The data pre-processing began with a re-categorization of the attributes (i.e., 37 SQ attributes and "Overall Service Quality"). The accuracy of the BN model, its complexity and the probability of each class are influenced by the number of categories of different variables (Kashani and Mohaymany, 2011). Therefore, because of the wide range of categories (11 categories) and the number of variables, the authors decided to reduce them by a re-categorization of the perceived scores, moving from an 11-point scale to a 5-point scale. The "Overall Service Satisfaction" was performed on a 5-point scale and no re-categorization was necessary.

This was followed by performing a Principal Component Analysis (PCA) on the 37 SQ attributes in order to determine the SQ latent dimensions. Another PCA was performed on the "Overall Service Quality" and "Overall Service Satisfaction" to determine the Quality of Service (QoS) latent dimension. The use of a PCA meant the number of dimensions was reduced to 9 (Table 2 shows the attributes and the dimensions obtained with the PCA). The resulting dimensions represent the variables of this study and are as follows: Accessibility (ACCESS), Availability of the Service (AVAIL), Customer Service (CUST_SER), Tangible Service Equipment (TANG_E), Security (SEC), Individual Space (IND_S), Information (INF), Environmental Pollution (ENV_POL), Quality of Service (QoS).

It is important to highlight that, based on the PCA results (Table 2), the attributes that showed a factor loading of 0.4 or higher in the same dimension were grouped together (Brons et al., 2009). Finally, the attributes B30, B1 and B3 were removed and the dimensions were extracted according to 36 attributes.

Table 2. Principal Component Analysis (PCA) results.

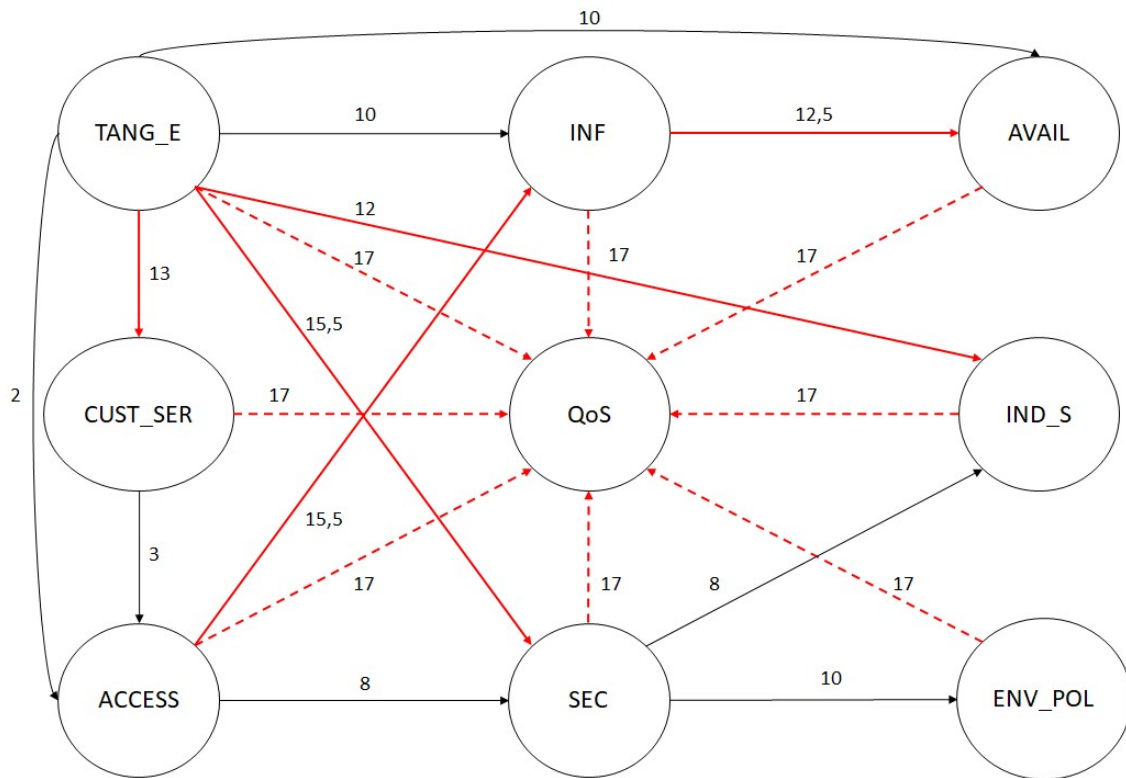
	PCA Factor loadings	PCA Factor weight scores
Accessibility (ACCESS)		
B8 Easy access of persons with reduced mobility	0.696	0.348
B6 Easy access to stations and platforms from the street	0.691	0.327
B7 Operation of elevators, escalators, etc.	0.671	0.326
B9 Operation of ticket validators at the entrance and exit of stations	0.635	0.311
B10 Easy use of ticket vending machines	0.613	0.301
B5 Easy connection with other transportation modes such as bike rental, taxis, buses, etc.	0.548	0.242
Availability of the Service (AVAIL)		
B2 Number of trains per day (frequency of the service)	0.769	0.434
B17 Waiting time on the platform	0.728	0.414
B16 Speed of the trip	0.632	0.315
B4 Regularity of the service (absence of interruptions caused by breakdown or incidents)	0.565	0.296
B15 Punctuality	0.554	0.243
Customer Service (CUST_SER)		
B20 Effectiveness and speed of employees to respond, give information and deal with users' daily problems	0.806	0.412
B19 Courtesy of the employees	0.794	0.400
B21 Performance of the Customer Service (offices, web site, contact by phone, dealing with complaints, etc.)	0.733	0.359
B18 Appearance of employees	0.693	0.323
Environmental Pollution (ENV_POL)		
B36 Noise level on the vehicle	0.868	0.425
B37 Vibration level on the vehicle	0.847	0.407
B35 Noise level in stations	0.819	0.401
Individual Space (IND_S)		
B26 Seat availability in stations and on platforms	0.739	0.514
B27 Level of comfort on vehicle (seat availability or enough room while standing up)	0.728	0.486
Information (INF)		
B12 Updated, precise and reliable information in stations (price, operating hours, stops, service interruptions, etc.)	0.733	0.451
B11 Updated, precise and reliable information on vehicles (operating hours, stops, service interruptions, etc.)	0.733	0.442
B14 Clear and simple notice boards with information and directions in stations	0.648	0.368
B13 Information available through other communication technologies (internet, phone, mobile applications, etc.)	0.596	0.356

Security (SEC)			
B32	Sense of security against theft and aggression in stations and on vehicles	0.740	0.458
B31	Sense of security against accidents while traveling (crash/vehicle derailment)	0.735	0.431
B33	Sense of security against slipping, falling and accidents at vehicle doors and escalators.	0.715	0.416
B34	Signage of emergency exits and extinguishers	0.592	0.315
Tangible service equipment (TANG_E)			
B22	Cleanliness of the stations	0.732	0.370
B24	Lighting in stations	0.707	0.341
B25	Lighting on vehicle	0.689	0.325
B23	Cleanliness of the vehicle	0.643	0.298
B28	Temperature and ventilation system on vehicle and in stations	0.418	0.160
B29	Appropriate driving	0.340	0.085
Quality of Service (QoS)			
SQ1	Overall Service Quality	0.921	0.543
SQ2	Overall Service Satisfaction	0.921	0.543

4.2. Bayesian Networks results

The 9 dimensions extracted from the PCA and re-categorized into a 5 point – Likert scale were used for learning the BN. Before calibration, the condition that QoS was directly linked to all the SQ dimensions was imposed as mandatory for the learning structures. Although the literature provides examples where the dimensions indirectly influence SQ (de Oña et al., 2017), most researches show a direct influence (e.g., de Oña et al., 2013; Eboli and Mazzulla, 2008; Redman et al., 2013; etc.).

Figure 1 shows the most robust network. This is learned using the TABU algorithm with mBDeu score. The total score of all 17 algorithms is reported for each arc. The red arcs have a score equal to or higher than 11. Note that the arcs between QoS and the rest of the variables have a score of 17 because these relationships were set as mandatory in all the algorithms. Other red arcs showing strong relationships between SQ dimensions are TANG_E with CUST_SER, SEC and IND_S; ACCESS with INF and INF with AVAIL. There are also other relationships with scores close to the thresholds, for example, TANG_E with INF and AVAIL or SEC with ENV_POL.

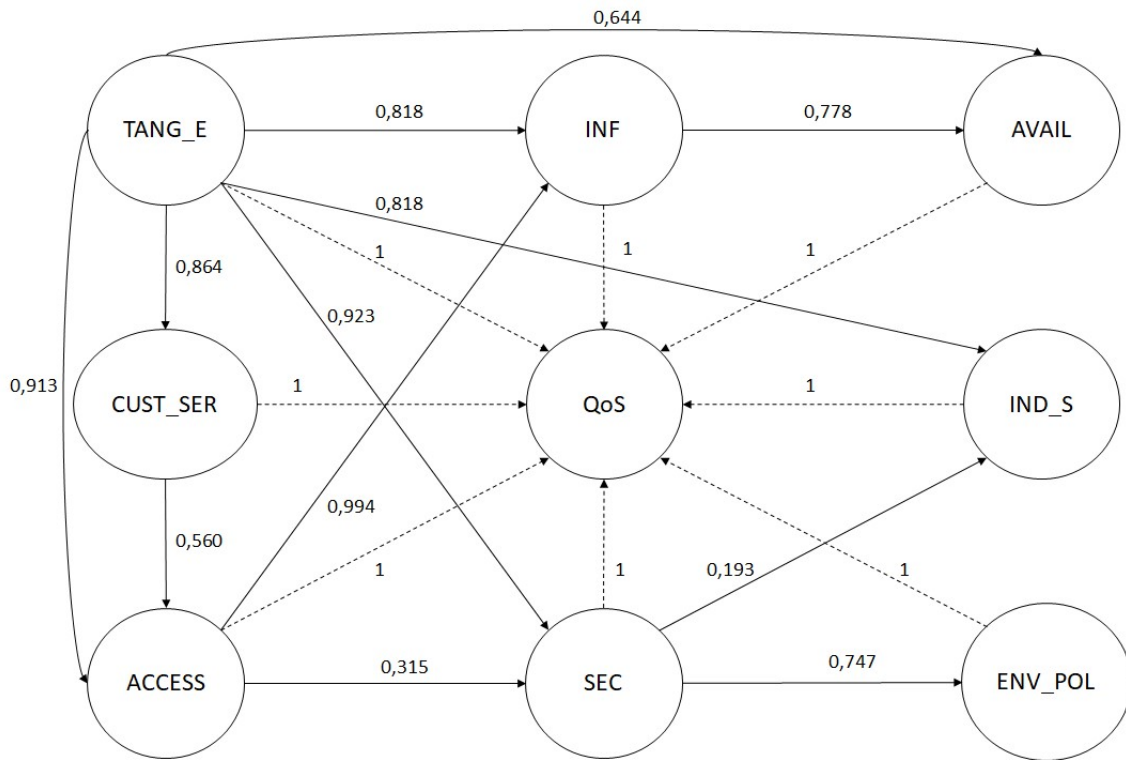


Note: the dotted arcs are mandatory relationships.

Figure 1. Learned Bayesian Network.

Figure 2 shows the occurrence proportion of each arc in the bootstrap replicates. The arcs with a proportion close to 1 are considered to be robust while those close to 0 need further analysis.

Note that almost all the red arcs have a value close to 1, which means that the relationships in the present BN are robust. Only the relationships between ACCESS and SEC, and SEC and IND_S present values under 0.5, so, their significance and relevance were analysed using SEM in order to discover whether or not they should be removed. The relationships between the 8 SQ dimensions and QoS show a value of 1 because these arcs were imposed as mandatory in accordance with the literature.



Note: the dotted arcs are mandatory relationships.

Figure 2. BN with the proportion of occurrence of each arc in bootstrap replicates.

4.3. Structural Equation Model results

The SEM was built using the dimension relationships obtained from the previous BN model. The latent variables representing the 8 SQ dimensions were explained using the 36 SQ attributes (from B2 to B37, without B1, B3 and B30) based on the PCA (Table 2) whereas the latent variable QoS, was explained with the two observed variables: “Overall Service Quality” (SQ1) and “Overall Service Satisfaction” (SQ2)

Table 3. Regression weights of Measurement Relationships.

Measurement Relationships			Unst	S.E.	St	p-value
B5	<---	ACCESS	1	-	0.625	
B6	<---	ACCESS	1.021	0.030	0.760	***
B7	<---	ACCESS	0.986	0.030	0.712	***
B8	<---	ACCESS	1.088	0.033	0.717	***
B9	<---	ACCESS	1.054	0.034	0.659	***
B10	<---	ACCESS	1.076	0.035	0.640	***
B2	<---	AVAIL	1	-	0.639	
B4	<---	AVAIL	0.835	0.031	0.561	***
B15	<---	AVAIL	0.827	0.025	0.733	***
B16	<---	AVAIL	0.966	0.029	0.739	***
B17	<---	AVAIL	1.101	0.033	0.715	***
B18	<---	CUST_SER	1	-	0.795	
B19	<---	CUST_SER	1.200	0.023	0.853	***
B20	<---	CUST_SER	1.293	0.024	0.852	***
B21	<---	CUST_SER	1.207	0.025	0.797	***
B35	<---	ENV_POL	1	-	0.775	
B36	<---	ENV_POL	1.095	0.022	0.866	***
B37	<---	ENV_POL	1.069	0.022	0.853	***
B26	<---	IND_S	1	-	0.673	
B27	<---	IND_S	1.087	0.036	0.791	***
B11	<---	INF	1	-	0.782	
B12	<---	INF	1.019	0.022	0.797	***
B13	<---	INF	0.914	0.028	0.588	***
B14	<---	INF	0.906	0.021	0.757	***
B31	<---	SEC	1	-	0.764	
B32	<---	SEC	0.925	0.024	0.711	***
B33	<---	SEC	1.013	0.025	0.741	***
B34	<---	SEC	0.858	0.021	0.733	***
B22	<---	TANG_E	1	-	0.740	
B23	<---	TANG_E	1.171	0.028	0.741	***
B24	<---	TANG_E	1.060	0.024	0.782	***
B25	<---	TANG_E	1.082	0.025	0.782	***
B28	<---	TANG_E	1.118	0.037	0.558	***
B29	<---	TANG_E	1.192	0.036	0.605	***
SQ1	<---	QoS	1	-	0.891	
SQ2	<---	QoS	0.872	0.020	0.773	***

Note: Unst., unstandardized; St., Standardized; S.E., estimate of the standard error of the covariance. *** ($p < 0.001$).

Table 4.Regression weights of Structural Relationships.

Structural Relationships			Unst	S.E.	St	p-value
ACCESS	<---	CUST_SER	0.278	0.022	0.306	***
ACCESS	<---	TANG_E	0.608	0.031	0.534	***
AVAIL	<---	INF	0.381	0.024	0.418	***
AVAIL	<---	TANG_E	0.525	0.033	0.427	***
CUST_SER	<---	TANG_E	0.862	0.027	0.690	***
ENV_POL	<---	SEC	0.581	0.024	0.518	***
IND_S	<---	SEC	0.437	0.036	0.414	***
IND_S	<---	TANG_E	0.537	0.052	0.340	***
INF	<---	ACCESS	0.661	0.036	0.557	***
INF	<---	TANG_E	0.371	0.037	0.275	***
SEC	<---	ACCESS	0.402	0.037	0.305	***
SEC	<---	TANG_E	0.757	0.044	0.506	***
QoS	<---	AVAIL	0.530	0.030	0.507	***
QoS	<---	CUST_SER	0.139	0.023	0.135	***
QoS	<---	ENV_POL	0.058	0.014	0.075	***
QoS	<---	IND_S	0.284	0.024	0.349	***
QoS	<---	SEC	0.078	0.028	0.090	0.006
QoS	<---	TANG_E	-0.213	0.050	-0.165	***

Note: Unst., unstandardized; St., Standardized; S.E., estimate of the standard error of the covariance. *** (p<0.001).

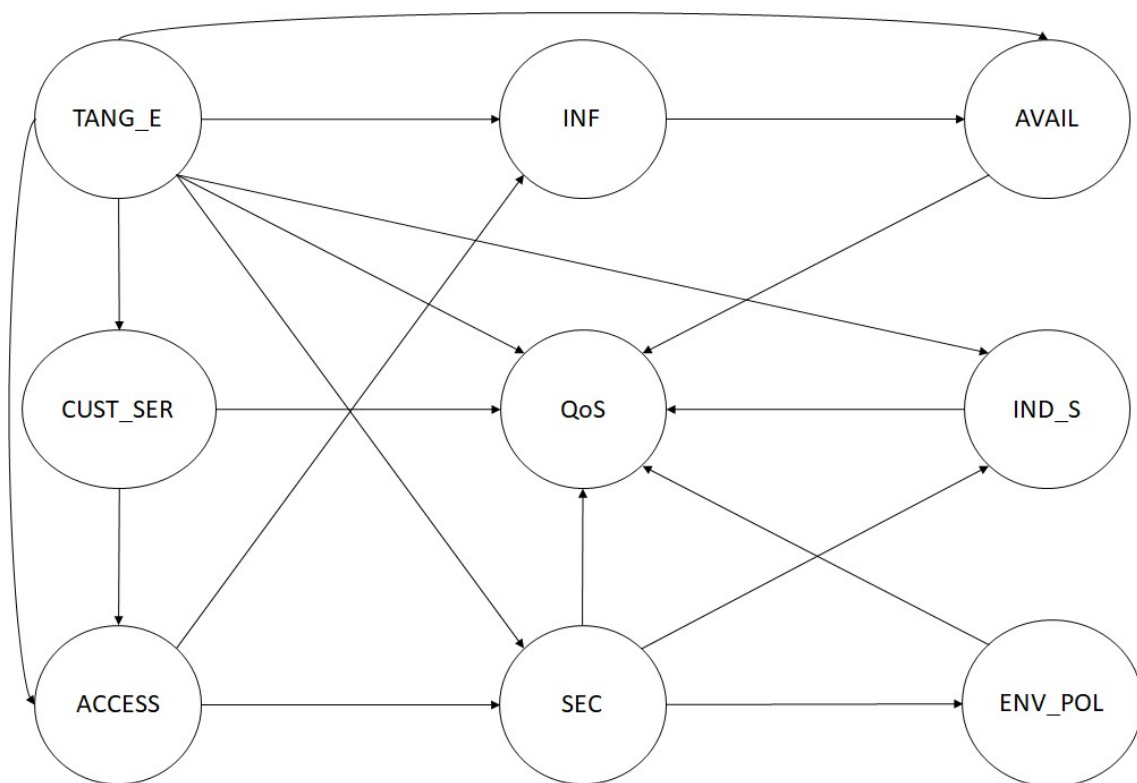


Figure 3. Structural model of SEM with significant relationships.

The results of the SEM are displayed in Table 3 and Table 4. The measurement model shows the relationships among the latent and observed variables (i.e., the 37 SQ attributes, SQ1 and SQ2 (Table 3)), while the structural model is formed by the 9 dimensions and their relationships (Table 4). Concerning the measurement model, all relationships were significant (at a significance level of 0.001), and all the standard regression weights were considerably high (St. > 0.5). SQ1 and SQ2 explained the QoS dimension very well with a standard regression weight of 0.891 and 0.773, respectively.

For the case of the structural model, most of the relationships were significant at a significance level of 0.001, and all of them at a level of 0.01. Even the two relationships where the proportion of occurrence in the final BN structure (Figure 2) was under 0.5 (i.e., INF with AVAIL and SEC with ENV_POL) were significant in the SEM model. Similarly, all the attributes show significant and positive relationships with their respective dimensions at a significance level of 0.001. Almost all the dimensions show significant and positive relationships with QoS. Only TANG_E shows a significant but negative relationship with QoS. However, this fact must be analysed through the overall effect the variables have on QoS.

Therefore, Table 5 shows that all the dimensions had a direct effect and ACCESS, INF, SEC and TANG_E had an indirect effect on QoS. The case of TANG_E can be seen to have had a slightly negative direct effect which can be considered to be not as important as its indirect effect.

Any analysis or study of the SQ model needs to take into account the overall effect, because if the research only considered the direct effects then TANG_E would have a negative influence on QoS, meaning the pre-assumption made in this study would not have been logical and validated and, therefore, the model would have been incorrect. However, the existence of relationships between the SQ dimensions allows the indirect effects on QoS to be included and these effects have an overall positive effect which supports our hypothesis about the different relationships between SQ and its dimensions (i.e., direct or indirect).

Considering the Standardized total effects shown in Table 5, the TANG_E and AVAIL dimensions are shown to be the most relevant aspects influencing passenger perception of QoS, whereas ENV_POL has the lowest total influence on QoS. These results agree with other analyses which used the same data base (i.e., De Oña et al., 2016b; De Oña et al., 2015; Machado-Leon et al., 2016; etc.).

Table 3. Direct, indirect and total effect on the "Quality of Service" latent variable.

		ACCESS	AVAIL	CUST_SER	ENV_POL	IND_S	INF	SEC	TANG_E
Direct Effects	St.	0.000	0.507	0.135	0.075	0.349	0.000	0.090	-0.165
Indirect Effects	St.	0.202	0.000	0.062	0.000	0.000	0.212	0.184	0.776
Total Effects	St.	0.202	0.507	0.197	0.075	0.349	0.212	0.274	0.611

The model's fit parameters can be seen in Table 6. The fit parameters of a basic model in which the SQ dimensions have direct relationships with QoS but without any interrelationships between them are also shown (Third column). A comparison between the fit parameters of both models clearly highlights that the proposed two-step process model described in this study achieves better results than the alternative used and thereby supports the existence of relationships between the various dimensions forming SQ.

The two-step process model proposed in this study achieves the following fit parameters: Model chi-square indicates that the magnitude of discrepancy between the sample and fitted covariance matrix is insignificant at a level of 0.05, which is the threshold value suggested by several authors (e.g., Golob, 2003; Hooper et al., 2008; Mulaik et al., 1989, etc.). Moreover, the same authors indicate that the sample size should be greater than 200 for an acceptable model, as is the case here. Absolute fit indices like GFI and AGFI have values similar to the recommended value, which is 0.90 (GFI=0.921, AGFI=0.909); RMSEA (0.045) is also under 0.08 for a good fit. Incremental fit indices have comparable values (NFI=0.929 and CFI=0.938); in this case a value closer to 1 indicates a good fit. The parsimony fit indices, PGFI and PNFI, have values of around 0.796 and 0.850, respectively, consistent with the statement expressed by Mulaik et al. (1989). Moreover, RMR (0.044) is below 0.05 which indicates well-fitting models. Therefore, the level of goodness-of-fit of the model was considered to be fine, even better than in other studies in the literature which have applied SEM in the field of PT and others (e.g., De Oña et al., 2013; Dimitrov, 2006; Philip et al., 2003, etc.).

Table 4. Goodness of fit measures of SEM.

	Model with Relationships between SQ dimensions	Model without Relationships between SQ dimensions
Fit Indices		
Chi-square	4,273.87	15,112.00
Degrees of freedom	576	586
Probability level	0.000	0.000
Number of distinct parameters to be estimated:	90	80
Absolute fit indices		
GFI	0.921	0.687
AGFI	0.909	0.643
RMSEA	0.045	0.088
RMR	0.044	0.281
Incremental fit indices		
NFI	0.929	0.750
CFI	0.938	0.757
Parsimony fit indices		
PGFI	0.796	0.604
PNFI	0.850	0.698

4.4. Discussion about the SQ Model

Figure 1 shows the relationships that are learnt using BN, and Figure 3 shows the definitive relationships of the SQ model which have been validated by SEM. Note that the relationships between INF and ACCESS with QoS have been removed. However, their influence on QoS is made indirectly through other dimensions. Therefore, it can be stated that all the different SQ dimensions have a direct or indirect influence on overall SQ.

TANG_E was shown to be the most influential dimension on SQ (total effect=0.611). Note that the important influence it has is made in an indirect way which means that the influence of TANG_E comes through other dimensions (e.g., AVAIL, ACCESS, CUST_SER, IND_S, INF or SEC). This fact can be explained by analysing the attributes grouped into TANG_E (i.e., temperature, cleanliness, lighting, appropriate driving). They might be the “base” of a good performance for these dimensions. For example, ACCESS will be perceived in a better way if the cleanliness and the lighting work perfectly, which in turn, will have a positive influence on passenger perception of QoS.

AVAIL was shown to be the second most influential dimension on SQ (total effect=0.507). This dimension has been shown in the literature to represent a basic and fundamental aspect of any public transport service (Eboli and Mazzulla, 2008; UNE-EN 13186, Tripp and Drea, 2002, etc.).

Note the cases of ACCESS and INF which did not have a direct influence on QoS. The passengers considered that the different attributes covered by ACCESS and INF did not have a direct influence on their global perception of QoS, nevertheless, they can still have a bearing on other dimensions (e.g., AVAIL or SEC) and therefore have an indirect influence on QoS. For example, if elevators and escalators operate appropriately, the sense of security against falls will improve and have a positive direct influence on passenger perception of QoS.

Finally, the ENV_POL was shown to be the dimension that had the least influence on QoS. This can be explained by the fact that the LRT Seville service is a modern environmentally friendly electric vehicle and the only pollution being produced is acoustic (noise and vibration). The passengers are fully aware of this.

Certain relevant relationships between SQ dimensions can be highlighted and could not have been identified if the proposed two-step process had not been applied.

A relevant case is the TANG_E dimension which influences the remaining dimensions apart from ENV_POL. It covers all the aspects relating to lighting, cleanliness (both vehicles and stations), temperature, comfort, ventilation, or even appropriate driving. A positive or negative perception of this dimension by the passengers is going to influence their perception of the remaining dimensions:

- a) Suitable temperature, cleanliness and ventilation allow passengers to perceive a good level of on-vehicle comfort (Relationships with IND_S).
- b) Good illumination of elevators, escalators, etc., allow passengers to easily access the stations and platforms from the street. Moreover, a good level of cleanliness for these facilities means they achieve a good score. Thus, the relationship with Access is more than justified (Relationship with ACCESS).
- c) In the case of security, lighting has a direct influence on the numbers of robberies or assaults and the number of physical accidents due to bad lighting (i.e. falling, crash, slips, etc.), proving the relationship (Relationship with SEC).
- d) Bad driving is going to influence the passenger perceptions about the performance of this service, from the point of view of frequency, punctuality, etc. These aspects are covered by AVAIL and support their relationships.
- e) Adequate lighting in stations also helps passengers to identify the different information boards in order to guide them around the stations or in the train. (Relationships with INF).

- f) Finally, passengers usually demand an adequate performance from the LRT, no excessive speeds, a good level of comfort, cleanliness, etc. In other words, these are all daily requirements or problems, the resolution of which will impact the perception passengers have about the effectiveness and speed of employees in dealing with them. Therefore, TANG_E has an influence on the CUST_SER dimension. (Relationships with CUST_SER)

The ACCESS dimension is connected to the INF dimension because well operating ticket validating or vending machines allow passengers to be reliably informed which improves their satisfaction with aspects associated to information.

The INF dimension is related to the AVAIL dimension. Updated, precise and reliable information means passengers can plan and organize their trips in advance. This fact reduces the waiting time on the platform, improves the sense of punctuality and, if this information is about the same service in other cities, it can improve their current level of satisfaction over the operating hours of the service, its frequency and regularity, etc.

Finally, the SEC dimension is linked up with the ENV_POL dimension. When passengers feel safe it entails a lower level of vibration and noise perception.

To sum up, these examples show that the relationships extracted from the BN are suitable for modelling SQ in this LRT and agree with real life expected scenarios. Bayesian Networks have also been shown to be a suitable appropriate technique for finding potential relationships which may exist between the variables used in the study. But, when combined with SEM it is possible to determine if these relationships are significant or not and, in consequence, to check and validate the SQ model framework obtained from the BN. Therefore, this study supports the use of the proposed two-step methodology and demonstrates its usefulness and relevance in developing new theories to identify new relationships between variables and to confirm their validity in this field and others.

5. Conclusions

This research demonstrates that the combined use of BN and SEM in a two-step methodological process represents a powerful tool which can be used to develop new theories or frameworks in any field of knowledge. In this study, the authors have applied this methodological approach for analysing SQ in the LRT Service of Seville (Spain). The research has demonstrated the usefulness of this methodology in identifying and validating hidden relationships between the SQ dimensions of the LRT service.

The authors have used the BN as an exploratory technique due to its ease of application, it learns directly from the data without the requirement of any pre-assumptions and its results can be easily interpreted. Furthermore, its characteristics overcome the limitations of SEM as a confirmatory technique. BN results are represented as a DAG from which the relationships between the model's variables can be extracted and analysed. This technique allows information to be introduced to condition the links between the variables (e.g., compulsion of link, prohibition of link, etc.) where researchers have partial expert knowledge about the phenomena being studied.

Similarly, SEM has been used as a confirmatory technique. It has been used to analyse the achieved SQ model due to its suitability as a technique for describing complex phenomena. SEM can be considered as a similar but more advanced technique to regression modelling; in fact, it allows researchers to introduce latent constructs which really appear in such phenomenon where latent dimensions are present. Therefore, the limitations associated with learning directly from data or the requirement for pre-existing knowledge are overcome by the characteristics of the two-step methodological procedure applied in this research.

In this study, the authors have followed the learning BN methodology shown in Cugnata et al., (2016). It consists of applying a wide range of different learning algorithms and measuring the arc strength using re-sampling techniques. In this paper, the most robust BN is learned by Hill Climbing with mBDeu score from the data of a CSS of the LRT Service of Seville (Spain). The robustness of this BN was checked by bootstraps. The final model was formed with 9 latent variables/dimensions (i.e., ACCESS, AVAIL, INF, TANG_E, CUST_SER, SEC, IND_S, ENV_POLCUST_SERENV_POL and QoS) and various observed variables (i.e. 37 SQ attributes, "Overall service quality" and "Overall Service Satisfaction"). The goodness-of-fit parameters of the SEM model provided excellent results, better than those from a model without relationships between the SQ dimensions. Moreover, combining both techniques in a two-step methodological process represents an interesting approach to model SQ in PT, as a more complex phenomenon SQ dimensions not only influence service quality, but also other SQ dimensions.

Furthermore, for this specific case, the resulting model structure provides valuable information for understanding what aspects of the service have the greater influence on passengers when they are deciding to use the service. This information can help transport managers to prepare new strategies and investment plans in order to continually improve the quality perceived by their passengers, and consequently the use of the system. Transit operators can also use these findings to attract new and retain existing passengers.

6. Acknowledgments

Support from the Spanish Ministry of the Economy and Competitiveness (Research Project TRA2015-66235-R) is gratefully acknowledged. The authors would also like to thank FEDER of the European Union for financial support via project “Mejora de la calidad del TP para fomentar la movilidad sostenible: Metro de Sevilla” of the “Programa Operativo FEDER de Andalucía 2007-2013”.

7. References

Amin, M. and Isa, Z., 2008. An examination of the relationship between service quality perception and customer satisfaction: A SEM approach towards Malaysian Islamic banking. *International Journal of Islamic and Middle Eastern Finance and Management*, 1(3), pp. 191-209.

Andreassen, T.W., 1995. (Dis)satisfaction with public services: the case of public transportation. *Journal of Services Marketing*, 9, pp. 30-41.

Bagozzi, R.P., 1994. *Structural equations models in marketing: Basic principles*. Principles of Marketing Research. Ed. Cambridge: Blackwell, pp. 317-385.

Beirao, G. and Cabral, J.A.S., 2007. Understanding attitudes towards public transport and private car: A qualitative study. *Transport Policy*, 14(6), pp. 478-489.

Bollen, K.A., 1989. *Structural Equations with Latent Variables*. Wiley, New York.

Brons, M., Givoni, M. and Rietveld, P., 2009. Access to railway stations and its potential in increasing rail use. *Transportation Research Part A: Policy and Practice*, 43(2), pp. 136-149.

Carman, J., 1990. Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions. *Journal of Retailing*, 66, pp. 33-55.

Celik, E., Bilisik, O.N., Erdogan, M., Gumus, A.T., Baracli, H. (2013). An Integrated Novel Interval Type-2 Fuzzy MCDM Method to Improve Customer Satisfaction in Public Transportation for Istanbul. *Transportation Research Part E*, 58, 28-51.

Chen, C.F., 2008. Investigating Structural Relationship Between Service Quality, Perceived Value, Satisfaction, and Behavioral Intentions for Air Passengers: Evidences from Taiwan. *Transportation Research Part A: Policy and Practice*, 42(4), pp. 709-717.

Chiou, Y.C., and Chen, Y.H., 2012. Service quality effects on air passenger intentions: A service chain perspective. *Transportmetrica*, 8(6), pp. 406-426.

Chou, J. and Kim, C., 2009. A structural equation analysis of the QSL relationship with passenger riding experience on high speed rail: An empirical study of Taiwan and Korea. *Expert Systems with Applications*, 36, pp. 6945-6955.

Cugnata, F., Kenett, R.S. and Salini, S., 2016. Bayesian Networks in Survey Data: robustness and sensitivity issues. *Journal of Quality Technology*, 48(3), pp. 253-264.

De Oña, J. and De Oña, R., 2015. Quality of Service in Public Transport Based on customer satisfaction surveys: A review and assessment of methodological approaches. *Transportation Science*, 49(3), pp. 605-622

De Oña, J., De Oña, R., Díez-Mesa, F., Eboli, L. and Mazzulla, G., 2016a. A composite index for evaluating transit Service quality across different user profiles. *Journal of Public Transportation*, 19(2), pp. 128-153.

De Oña, J., De Oña, R., Eboli, L. and Mazzulla, G., 2013. Perceived Service quality in bus transit service: A structural equation approach. *Transport Policy*, 29, pp. 219-226.

De Oña, J., De Oña, R., Eboli, L. and Mazzulla, G., 2016b. Transit passengers' behavioural intentions: The influence of service quality and customer satisfaction. *Transportmetrica A: Transport Science*, 12(5), 385-412.

De Oña, J., Mujallli, R.O. and Calvo, F.J., 2011. Analysis of traffic accident injury severity on spanish rural highways using bayesian networks. *Accident Analysis & Prevention*, 43(1), pp. 402-411.

De Oña, R., de Abreu, J., Muñoz-Monge, C. and de Oña, J., 2017. Users' satisfaction evolution of a metropolitan transit system in a context of economic downturn. *International Journal of Sustainable Transportation* (Just-Accepted)

De Oña, R., Machado, J.L. and De Oña, J., 2015. Perceived Service Quality, Customer Satisfaction and Behavioral Intentions Structural Equation Model for the Metro of Seville, Spain. *Transportation Research Record*, 2538, pp. 76-85.

Dell'Olio, L., Ibeas, A. and Cecin, P., 2011. The quality of service desired by public transport users. *Transport Policy*, 18(1), pp. 217-227.

Dimitrov, D.M., 2006. Comparing groups on latent variables: A structural equation modeling approach. *Speaking of Research*, 26, pp. 429-436.

Duarte, C. W., Klimentidis, Y. C., Harris, J. J., Cardel, M. and Fernandez, J. R., 2011. A hybrid Bayesian Network/Structural Equation Modelling (BN/SEM) approach for detecting physiological networks for obesity-related genetic variants. *Proceedings IEE International Conference Bioinformatics and Biomedicine*, pp. 696-702.

Eboli, L. and Mazzulla, G., 2008. A stated preference experiment for measuring service quality in public transport. *Transportation Planning and Technology*, 31(5), pp. 509-523.

Eboli L. and Mazzulla, G., 2012. Structural equation modelling for analyzing passengers' perceptions about railway services. *Procedia-Social and Behavioural Science*, 54, pp. 96-106.

Fillone, A.M., Montalbo, C.M. and Tiglaio., N.C., 2005. Assessing urban travel: A structural equations modelling (SEM) approach. *Proceedings of the Eastern Asia Society for Transportation Studies*, 5, pp. 1050-1064.

Golob, T.F., 2003. Structural equation modeling for travel behavior research. *Transportation Research Part B: Methodological*, 37(1), pp. 1-25.

Grönroos, C., 1984. A service quality model and its marketing implications. *European Journal of marketing*, 18(4), pp. 36-44.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E., 2010. *Multivariate Data Analysis: A Global Perspective*. Prentice Hall, New Jersey.

Hänninen, M., 2008. Analysis of human and organizational factors in marine traffic risk modeling. Literature Review. Series AM. Espoo, Helsinki University of Technology. Faculty of engineering and architecture. Department of applied mechanics.

Hapsari, R., Clemes, M.D. and Dean, D., 2017. The impact of service quality, customer engagement and selected marketing constructs on airline passenger loyalty. *International Journal of Quality and Services*, 9(1), pp. 21-40

Heckerman, D., 1998. A tutorial on learning with Bayesian networks. *Learning in graphical models*. Springer Netherlands.

Ho, C. and Lee, Y., 2007. The development of an e-travel service quality scale. *Tourism Management*, 28(6), pp. 1434-1449.

Hooper, D., Coughlan, J. and Mullen, M.R., 2008. Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), pp. 53-60.

Islam, M.R., Hadiuzzaman, M. Banik, R., Hasnat, M.M., Musabbir, S.R., Hossain, S. (2016). Bus Service Quality Prediction and Attribute Ranking: A Neural Network Approach. *Public Transport*, 8(2), 295-313.

Kamaruddin, R., Osman, I. and Pei, C.A.C., 2012. Public Transport Services in Klang Valley: Customer expectations and its relationship using SEM. *Procedia-Social and Behavioral Sciences*, 36, pp. 431-438.

Kashani, A.T. and Mohaymany, A.S., 2011. Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. *Safety Science*, 49(10), pp. 1314-1320.

Kenett R.S. and Salini, S., 2011. Modern analysis of customer surveys: with applications using R (Vol. 117). John Wiley & Sons.

Kjaerulff, U.B. and Madsen, A.L., 2008. Bayesian networks and influence diagrams: a guide to construction and analysis. Information Science and Statistics, Jordan, M., Kleinberg, J. and Schölkopf, B. (eds). Springer.

Kuo, M. and Liang, G., 2011. Combining VIKOR with GRA techniques to evaluate service quality of airports under fuzzy environment. *Expert System Appl.* 38(3), pp. 1304-1312.

Linda, S., 2003. Can public transport compete with the private car?. *IATSS Research*, 27(2), pp. 27-35.

Machado-Leon, J.L., de Oña, R. and de Oña, J., 2016. The role of involvement in regards to public transit riders' perceptions of the service. *Transport Policy*, 48, 34-44.

Mulaik, S.A., James, L.R., Van Alstine, J., Bennett, N., Lind, S. and Stilwell, C.D., 1989. Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105, pp. 403-445.

Nathanail, E. (2008). Measuring the Quality of Service for Passengers on the Hellenic Railways. *Transportation Research Part A*, 42(1), 48-66.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L., 1985. A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49, pp. 41-50.

Pearl, J., 1988. Probabilistic reasoning in intelligent systems. San Francisco. CA: Morgan Kaufman.

Philip, K., Gus, M., Rodney, A., John, A., 2003. Customer repurchase intention: A general structural equation model. *European Journal of Marketing*. 37(11, 12), pp. 1762-1800.

Rahman, F., Das, T., Hadiuzzaman, M.D. and Hossain, S., 2016. Perceived service quality of paratransit in developing countries: A structural equation approach. *Transportation Research Part A: Policy and Practice*, 93, pp. 23-38.

Redman, L., Friman, M., Gärling, T. and Hartig, T., 2013. Quality attributes of public transport that attract car users: A research review. *Transport Policy*, 25, pp. 119-127.

Scheines, R., Hoijtink, H. and Boomsma, A., 1999. Bayesian estimation and testing of structural equation models. *Psychometrika*, 64(1), 37-52.

Shonk, D., and Chelladurai, P., 2008. Service Quality, Satisfaction, and Intent to Return in Event Sport Tourism. *Journal of Sport Management*, 22, pp. 587-602.

Trentini, F., Malgaroli, M., Camerini, A.L., Di Serio, C. and Schulz, P., 2015. Multivariate determinants of self-management in Health Care: assessing Health Empowerment

Model by comparison between structural equation and graphical models approaches. *Epidemiology Biostatistics and Public Health*, 12 (1), 1-13.

Weinstein, A., 2000. Customer satisfaction among transit riders. How customer rank the relative importance of various service attributes. *Transportation Research Record*, 1735, pp. 123-132.

Woods, R., Masthoff, J., 2017. A Comparison of Car Driving, Public Transport and Cycling Experiences in Three European Cities. *Transportation Research Part A*, 103, pp. 211-222.

Yang, K., Hsieh, T., Li, H. and Yang, C., 2012. Assessing how service quality, airline image and customer value affect the intentions of passengers regarding low cost carriers. *Journal Air Transport Management*, 20, pp.52-53.

Yilmaz, V. and Ari, E., 2017. The effects of service quality, image, and customer satisfaction on customer complaints and loyalty in high-speed rail service in Turkey: a proposal of the structural equation model.

Anexo 5.

ARTÍCULO 2. DE OÑA ET AL. (UNDER REVIEW)

Se muestra el manuscrito del siguiente artículo que se encuentra en revisión en una revista científica de impacto:

De Oña, R., Diez-Mesa, F., De Oña, J. (under review) Bayesian Network and Structural Equation Model for Analysing Passenger Behavioural Intentions, Perceptions, Satisfaction and Involvement. Transportation Research Board 99th Annual Meeting/Transportation Research Record.

Bayesian Network and Structural Equation Model for Analysing Passenger Behavioural Intentions, Perceptions, Satisfaction and Involvement

ABSTRACT

The study of passenger behavioural intentions towards Public Transport (PT) is an important tool for transport operating companies and service managers. Various latent constructs (e.g. service quality, satisfaction, involvement, perceived benefits, etc) can be found in the bibliography to explain passenger behavioural intentions in terms of once again using this service or recommending it to other people. Knowledge about the reality of this phenomenon is a conditioning factor for marketing campaigns encouraging service uptake or the focus for making service improvements. Given that there is no consensus about how said constructs interact and, in some cases some important relationships could be missed due to a lack of knowledge or because such relationships are specific to the service and context being analysed, this study will apply a two stage methodology. This methodology will use Bayesian Networks (BN) and Structural Equation Models (SEM) together to extract knowledge directly from the data and validate it. The context will be passenger behaviour on the Light Rail Transit of Seville (Spain) using a data base of 3,198 valid surveys. The results confirm certain previously established relationships, such as the Service Quality-Satisfaction-Behavioural Intentions paradigm, as well as identifying new relationships. This paper demonstrates the utility of this methodology to generate models directly from the data with satisfactory fits.

Keywords: Bayesian Networks, Structural Equation Model, Service Quality, Satisfaction, Behavioural Intentions, Public Transport.

1. Introduction

Understanding passenger behavioural intentions (BI) after experiencing a certain type of public transport (PT) is a requirement for any transport manager. This knowledge allows them to develop efficient strategies to deliver passenger needs, thereby retaining existing and attracting new passengers from other modes of transport (1). The study of these behavioural intentions is based on a series of latent constructs (2). The most analysed constructs found in the literature are: service quality (SQ), satisfaction (SA), attractive alternatives (AA), involvement (INV), perceived costs (PC) and perceived benefits (PB), (1-4), although sometimes perceived value construct has been analysed, as it is considered a trade-off between perceived costs and perceived benefits (1)

If the concepts of satisfaction and service quality are analysed in detail, they are found to be of a similar nature as both derive from disconfirmation theory (5). Satisfaction is defined as an overall affective response to a perceived discrepancy between prior expectations and perceived performance after consumption, while service quality is more specific and related to cognitive judgments (2). This fact has meant that on more than one occasion, not only their meanings have over crossed in the literature (6), but also the way in which they interrelate. Put simply, this explains why we can state that most authors accept that service quality is a determinant factor in satisfaction (an antecedent of it), although it does not completely explain it (7-8). Nevertheless, some work considers service quality to be a consequence of satisfaction (9).

If this point is accepted, we can state that the conflict begins or widens with the consideration of the behavioural intentions (BI) construct, defined by Oliver (10) as the strong compromise to once again buy or use a determined preferred product or service in the future, or recommend it to other people. In fact, it was created the so-called "Service quality–Satisfaction–Loyalty" paradigm, which suggests that satisfaction is a link between service quality and loyalty (11). Most researchers tend to employ the BI measure to represent customer loyalty, since action loyalty is difficult to observe and is often equally difficult to measure, and it is accepted that favourable BI lead to customer loyalty (1). Then, in this paper we will use the BI construct. Furthermore, some evidence can be found in the literature showing that service quality could also have a direct influence on BI (1; 4;12-14). Authors (1-3) have tried to explain the behavioural intentions of passengers not only through service quality and satisfaction, but also by using other constructs. Each of these constructs covers different fields influencing passenger behaviour. The perceived costs could be defined as what the passenger renounces or sacrifices in order to use the service (15), which covers things that have monetary value or not (16-17). The perceived benefits are the main benefits obtained by using the service (15). The attractive alternatives cover the preferences for using alternative transport modes to those being studied (2). Finally, involvement indicates the pre-disposition and opinion of passengers towards PT in general and the people that use it (1).

There is currently no agreement about how these constructs are interrelated to explain BI. For example, (1) proved that satisfaction has a direct effect over involvement, while (2) validated that the relationship between both constructs was in the opposite direction (from involvement to satisfaction). In general, the work of different authors (7;18-19) uses structural equation models (SEM) to confirm the relationships between the constructs, the hypotheses of which were already established in the literature and used as guidelines. However, certain important relationships between these constructs may not have been considered, with the consequent repercussions on the results and the models obtained, as they would not be in complete agreement with reality. It should be highlighted that this circumstance would be completely outside the responsibility of the researcher, as their origin would be ignorance of the presence of such relationships, or simply addressing individual characteristics relevant to a particular context or case study.

Therefore, in order to avoid having to propose literature based hypotheses which may not represent reality, this paper proposes a two-stage methodology to explore and validate the existing relationships between different constructs. This methodology is made up of an exploratory stage and a confirmatory stage using the application of Bayesian Networks (BN) and SEM. The BN are able to extract hypotheses about the existing relationships between constructs directly from the data, without requiring expert knowledge about the phenomenon being analysed, while SEM validates these hypotheses, thereby avoiding any identified relationships at the BN stage being due to purely statistical properties. This methodology has been used in other fields PT research (20), specifically for the analysis of service quality. The aim of this work is to use the stated methodology to extract and confirm new knowledge about the behavioural intentions, passenger perceptions, satisfaction and involvement in the Light Rail Transit (LRT) system of Seville (Spain). This research will validate the applicability of this stated methodology to another field of travel behaviour.

The paper is organised as follows: the methodology section presents the BN and SEM models, as well as the steps followed for the proposed process; the data section briefly describes the survey, the data collection process and some descriptive statistics; the section results and discussion summarise and highlight the main outcomes obtained with the analysis; and finally, some conclusions and policy implications are offered in the last section.

2. Methodology

The techniques BN and SEM are briefly described below, followed by an introduction to the steps followed in developing the two-stage methodology.

2.1. Bayesian Networks (BN)

A BN (B) is defined as a Directed Acyclic Graph (DAG) representing a probability distribution together with a group of random variables V . Two basic parts of the BN can be differentiated ($B = (G, \Theta)$). A qualitative component (G), which would be the DAG, formed of nodes X_1, \dots, X_n , representing the random variables being considered in the study, and by links, representing the direct probable dependency between the variables. And a quantitative component (Θ), consisting of a joint probability distribution that factorises into a set of conditional probability distributions, governed by the structure of the DAG (21).

The qualitative part guarantees that none of the nodes will be a predecessor or descendent of itself, which is a basic condition of BN. Considering this condition, we can state that graph G codifies the independence hypothesis, so that each variable X_i is independent of its non descendents given its predecessors in G , generically termed π_i .

The quantitative part of this network, group Θ contains the parameter $\Theta_{x_i|\pi_i} = P_B(x_i|\pi_i)$ for each category x_i of X_i conditioned π_i , where X_i is the group of predecessors in G . Therefore, B defines a unique grouped probability distribution about V , as:

$$P_B(X_1, \dots, X_n) = \prod_{i=1}^n P_B(X_i|\pi_i) = \prod_{i=1}^n \Theta_{x_i|\pi_i} \quad (1)$$

The work of (22) is recommended for a detailed description of BN.

2.2. Structural Equation Modelling (SEM)

SEM can be defined as a multivariate technique combining regression, factor analysis, and variance analysis to simultaneously estimate interrelated dependence relationships. A series of equations, similar to those used in multiple regressions, describes the relationships between the analysed constructs. SEM consists of two components, a measurement model that assesses unobserved latent constructs as linear functions of observed variables (items), and a structural model that shows the direction and strengths of the relationships of the latent constructs. The relationships construct-item and construct-construct are associated to a standardised factor loading (or standardised regression weight, SRW, using AMOS software terminology) that indicates the strength and significance of the relationship.

The basic SEM equation can be defined as in (23)

$$\eta = B\eta + \Gamma\xi + \zeta \tag{2}$$

In which η is a $m \times 1$ vector of the latent constructs, ξ is a $n \times 1$ vector of the latent exogenous constructs, B is an $m \times m$ matrix of the coefficients associated with the latent endogenous constructs, Γ is an $m \times n$ matrix of the coefficients associated with the latent exogenous constructs and ζ is an $m \times 1$ vector of error terms associated with the endogenous constructs.

The basic equations of the measurement model can be expressed as:

$$x = \Lambda_x\xi + \delta \tag{3}$$

$$y = \Lambda_y\eta + \varepsilon \tag{4}$$

In which x and δ are column q -vectors related to the observed exogenous constructs and errors, respectively; Λ_x is a $q \times n$ structural coefficient matrix for the effects of the latent exogenous constructs on the observed variables, y and ε are column p -vectors related to the observed endogenous variables and errors respectively, and Λ_y is a $p \times m$ structural coefficient matrix for the effects of the latent endogenous constructs on the observed variables.

Different methods can be used to estimate the parameters of the model (e.g. ML, GLS, etc.), depending on the probability distribution, the scale properties of the variables, the complexity of the SEM, and the sample size (24). Moreover, SEM is a confirmatory technique because the researcher constructs the model by defining unidirectional effects between variables (24).

2.3. Methodological process: Step by step

The methodology followed in this article is based on (20), the reader is therefore referred to this work for a more detailed description of the methodology, however a schematic summary is provided in Table 1.

Table 1. Methodological process step by step.

Steps	Tasks	Description
Step 0: Data-preprocessing	EFA	<ul style="list-style-type: none"> ✓ Sample size > 5 obs./variable; MSA > 0.8; Bartlett's test (p-value < 0.001); Factor loadings > 0.4 ✓ Extraction of latent constructs
	CFA	<ul style="list-style-type: none"> ✓ Assess construct validity of measurement model: Factor loadings > 0.7; AVE > 0.5; Cronbach's Alpha and CR > 0.7; AVE > squared intercorrelation between constructs

Step 1: Exploratory Analysis (BN)	BNs learned	<ul style="list-style-type: none"> ✓ Sample split into training and test sub-samples: 70/30 ✓ Algorithms used: 7 different algorithms and 6 metrics
	Selection of the most robust BN	<ul style="list-style-type: none"> ✓ Relationships score: 0, 0.5 and 1 (no arc, no directed arc, directed arc, respectively) ✓ BN most robust: the one with most relationships scored over a threshold >11. If various BNs are tied, the one with lowest misclassification rate based on test sub-sample
	Relationships' strength	<ul style="list-style-type: none"> ✓ Bootstrap resampling: BNs calibrated with 1000 random subsets with a proportion of data 1/10 ✓ Influence ranking based on Distance-Weighted Influence (DWI)
Step 2: Confirmatory Analysis (SEM)	Model structure	<ul style="list-style-type: none"> ✓ The relationships for the SEM are based on the DAG of the most robust BN
	Model's reformulation	<ul style="list-style-type: none"> ✓ Remove non-significant relationships (p-value > 0.05)
	Model fit	<ul style="list-style-type: none"> ✓ Check goodness-of-fit indexes based on recommended values

Note: EFA: Exploratory Factor Analysis; CFA: Confirmatory factor Analysis; MSA: Measure of Sampling Adequacy; CR: Construct Reliability; AVE: Average Variance Extracted; DWI: Distance-Weighted Influence; DAG: Directed Acyclic Graph

The phases linked to the application of the BN are based on the work in (26) and a modification of the code in R. It needs to be highlighted that to create the BNs the researchers had to work with the constructs extracted during data pre-processing (from Exploratory Factor Analysis, EFA, and Confirmatory Factor Analysis, CFA). With regards to the threshold established to choose the most robust BN (the one with the most relationships scoring above the fixed threshold) the value of 11 represents 70% of all the learnt networks, as recommended by Cugnata et al. (26). This work has incorporated the Distance-Weighted Influence (DWI) metric, which can be used to calculate a ranking for the influential variables around an objective node based on the network structure (27), in the present case it will be used to determine the most influential constructs. The DWI between X and Y is defined in the following way:

$$DWI(X, Y; w) = \sum_{s \in S(X, Y)} w^{|s|} \quad (5)$$

Where $S(X, Y)$ is the group of routes in the BN that connect nodes X and Y, $|s|$ is the length of route s, and w is the weighting of the route, defined as the product of the force of the links forming the route.

3. Data

The study concentrates on the LRT system of Seville (Spain). The length of the line is 18 kilometres with 21 stations. This LRT line connects four municipalities in the metropolitan area of Seville with an overall population of 850,000 inhabitants.

Passengers' perceptions in regard to the LRT system were collected with an on-line survey. The survey was structured into 4 sections:

- Part A: Score the different items related to passenger behavioural intentions, satisfaction and perceptions of the LRT service as well as towards PT in general.
- Part B: Score the perceived quality with respect to different aspects of the service, as well as provide an overall score for the service.
- Part C: Travel habits (i.e. frequency of use, ticket used, etc.)
- Part D: Socioeconomic characteristics (i.e. gender, age, income, etc.).

The questions in Parts A and B were scored with an 11-point Likert scale. Only the question related to the overall satisfaction of the passenger with the service was measured with a 5-point Likert scale (1-lowest level of satisfaction, 5-highest level of satisfaction).

19,863 cards were administered to users by trained interviewers during a period of two weeks (May-June 2014), on weekdays, Saturdays and Sundays. A raffle was conducted as a reward for participants who completed the survey. 3,365 responses were registered (response rate value of 17.09%), from which 3,198 were valid. Table 2 summarises the socioeconomic characteristics of the complete sample.

Table 2. Sample socio-economic characteristics

<i>Characteristics</i>	<i>Percentage</i>
Gender	Woman (53.35%) Man (46.65%)
Disability	Permanent (0.97%) Temporary (0.47%) None (98.56%)
Availability of...	Driver License (74.95%) Access to private car (54.69%) Access to bicycle (43.18%) Access to motorcycle (6.72%) None (11.98%)

Age	<18 (2.81%) 18-25 (41.59%) 26-40 (28.89%) 41-65 (25.58%) >65 (1.03%) No response (0.1%)
Level of studies completed	None or Secondary school (9.01%) Bachelors or higher (48.47%) High school or Professional education (41.96%) No response (0.56%)
Employment status	Employed (43.68%) Student (41.53%) Retired (2.56%) Other (12.23%)
Household size	<3 (24.04%) 3-4 (60.44%) >4(15.52%)
Household monthly income	<1,201€ (28.80%) 1,201 -1,800€ (21.04%) 1,801-2,400€ (16.48%) >2,401€ (15.98%) No response (17.70%)

4. Results and discussion

Step 0. Pre-processing of the data

This study used 26 variables for part A (A1-A26), 8 service quality dimensions (part B), 1 question about overall perceived service quality (part B: SQ1), and 1 question about overall satisfaction (part A: SQ2). A total of 36 variables were considered and re-categorised from an 11 point scale to a 5 point scale (scores 0-2 were recoded as 1, 3-4 as 2, 5-6 as 3, 7-8 as 4, and 9-10 as 5). SQ2 was not re-categorised because it had already been asked on a 5 point scale. The 8 service quality dimensions are the same as those found by Diez-Mesa et al. (20) through an EFA (Accessibility, Customer service, Availability of service, Tangible equipment, Individual space, Information, Environmental pollution and Security) and they were therefore treated as variables in this work. The reader can find more details about the service quality dimensions extraction in Diez-Mesa et al. (20).

The results obtained during the data pre-processing stage are described below:

Exploratory Factor Analysis

An initial EFA was performed with 23 items from part A of the survey (from A4 to A26), with the hope of identifying different latent constructs. Sample size, sampling adequacy, and Bartlett's test sphericity were satisfactory, with an observation ratio per attribute of 139.04 (>5), the Bartlett test had a value of 253 significant to 0.001, and the MSA had a value of 0.864. Nevertheless, because the factorial weight of various items was lower than the minimum recommended threshold (0.4), it was necessary to sequentially remove the variables A10 (0.318), A6 (0.314) and A7 (0.315). Finally, the 20 remaining items were grouped into 6 constructs explaining a variance of 58.68%, which was considered adequate (25). The identified constructs were named as involvement (INV), perceived benefits (PB), perceived costs (PC), attractive alternatives (AA) and behavioural intentions (BI).

A second EFA was performed on the items which were directly linked to overall satisfaction (from A1 to A3 and SQ2). Their observation ratio per attribute was 799.5 (>5), the Bartlett test was 3,824, significant to 0.001, and the MSA was 0.746. The attributes were grouped into a unique latent construct, identified as satisfaction (SA) and explaining a variance of 61.43%.

Finally, the third EFA was performed with the 8 quality dimensions and overall service quality (SQ1), obtaining satisfactory results (observations ratio per attribute equal to 399.75, the Bartlett test had a value of 10,566, significant to 0.001, and the MSA was 0.926). Only one construct was identified and was named as service quality (SQ), explaining a variance of 50.10%. Table 3 shows factor loadings and the factor score weights of the different items in each construct.

Table 3. EFA and CFA final results.

Attributes		Factor Loadings	Factor Score Weights
Involvement (INV)			
<i>(CR: 0.810; AVE: 0.462; Cronbach's Alfa: 0.803)</i>			
21	I feel that taking public transport is consistent with my lifestyle	0.786	0.293
26	I like people who take public transport	0.786	0.292
25	I like others to know the fact that I take public transport	0.751	0.283
22	Independently of trip purpose, I always prefer to travel by PT	0.709	0.270
23	I feel that by taking PT I help to protect the environment	0.669	0.236

Perceived Benefits (PB)			
<i>(CR: 0.735; AVE: 0.484; Cronbach's Alfa: 0,713)</i>			
12	The attention to the costumer is good	0.824	0.564
9	The service of LRT is good	0.805	0.490
13	I like the LRT because of the speed of the trip	0.621	0.341
Service Quality (SQ)			
<i>(CR: 0.874; AVE: 0.465; Cronbach's Alfa: 0.866)</i>			
Tan_Eq	Tangible Equipment (Tan_Eq)	0.765	0.179
Acces	Accessibility	0.764	0.179
Info	Information	0.749	0.176
SQ1	Overall SQ	0.739	0.173
Sec	Security	0.729	0.171
Cust_Serv	Customer Service	0.726	0.170
Avail	Availability of service	0.724	0.170
Ind_Space	Individual space	0.635	0.149
Perceived Cost (PC)			
<i>(CR: 0.837; AVE: 0.631; Cronbach's Alfa: 0.833)</i>			
4	I believe that the price is high.	0.880	0.421
5	I believe that the ticket price exceeds the costs of the LRT (staff, electricity, maintenance, etc.)	0.873	0.415
8	I consider the costs of travelling by LRT to be high (time, money and comfort).	0.797	0.357
Behavioural Intention (BI)			
<i>(CR: 0.782; AVE: 0.545; Cronbach's Alfa: 0.760)</i>			
20	I will travel by LRT again under the same conditions (money, time and comfort)	0.857	0.529
18	Surely. I will use the LRT service again	0.816	0.476
19	I usually recommend the LRT service to others	0.645	0.318
Satisfaction (SA)			
<i>(CR: 0.788; AVE: 0.484; Cronbach's Alfa: 0.781)</i>			
2	I feel comfortable travelling by LRT	0.818	0.333
SQ2	Overall Satisfaction with the service of the LRT	0.792	0.322
3	The service of LRT meets my expectations	0.79	0.321
1	Travelling by LRT attracts me	0.733	0.298
<i>Items removed with EFA or CFA: A6. The stations are far away from my origin and/or destination, A7. I consider the waiting time on the platforms to be excessive, A10. The quality-price relationship is acceptable, A11. The timetable meets my needs, A14. I prefer the metro to other types of public transport (bus, taxi or hired bicycle), A15. I believe there are good public transport alternatives to the metro (bus, taxi, etc.), A16. I don't mind which transport mode I use so long as it meets my needs. A17. I think other modes of transport (car, bicycle, bus, taxi, etc.) offer more advantages than the metro, A24. I think that using public transport affects other peoples' opinion about me; and Enviromental Pollution.</i>			

Goodness-of-fit	
X^2	4,293.842
Degree of freedom	284
GFI	0.896
AGFI	0.871
RMSEA	0.066
RMR	0.047
NFI	0.885
CFI	0.892
PGFI	0.725

Note: CR: Construct Reliability; AVE: Average Variance

Confirmatory factor analysis

The first CFA removed the AA construct, made up of items A15 to A17 because its Cronbach's Alpha was very low (0.482) and it was not possible to prove its discriminant or convergent validity. It was also necessary to sequentially remove the item environmental pollution from the SQ construct, A11 from the PB construct, A14 from BI and A24 from APT, as they did not fulfil the convergent validity.

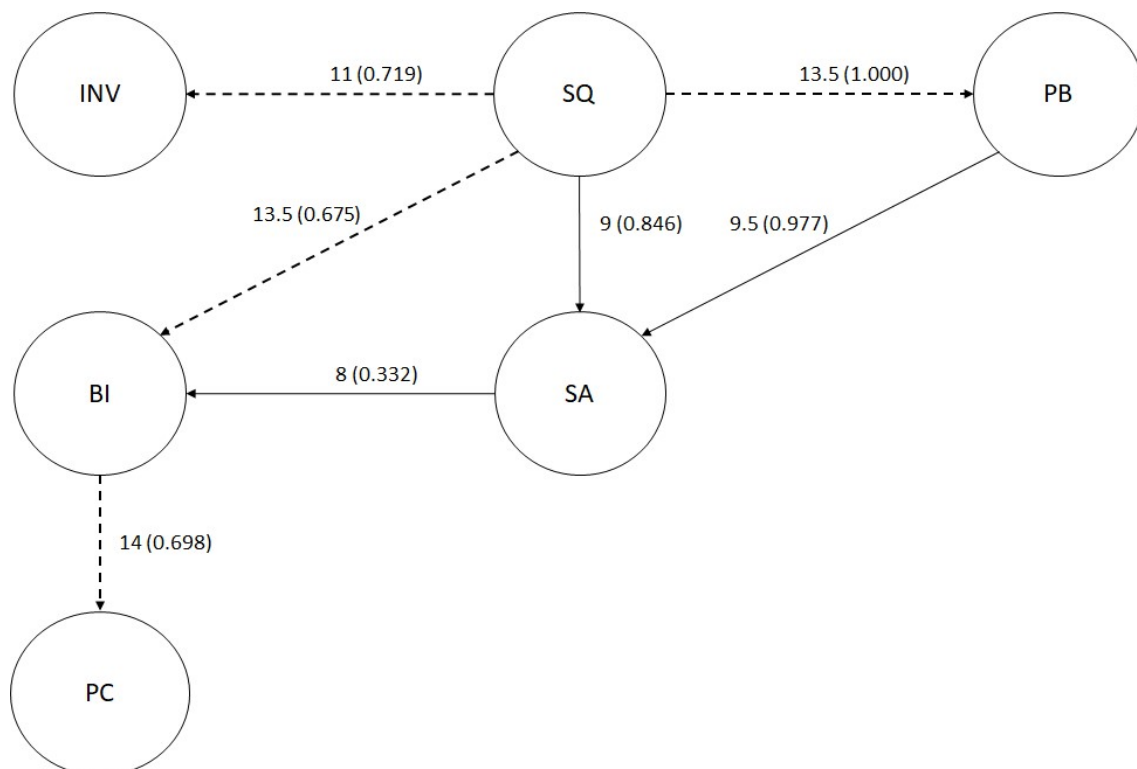
Finally, the 6 identified constructs were used to define a measuring model: INV, PB, PC, BI, SA and SQ. This model had fit indexes values close to the recommended thresholds (CFI of 0.892, very close to 0.9; RMSEA of 0.066, lower than 0.08). The calibrated measuring model was therefore considered to be suitable.

The convergent validity was confirmed when most of the factor loadings were found to be above the ideal threshold of 0.7. Only the PC and BI constructs had a value over 0.5 (0.631 and 0.545, respectively) for the AVE. Nevertheless, the remaining constructs showed values closer to the threshold. Their Cronbach's Alpha and CR values were also above 0.7. Table 3 presents the EFA and CFA results, which are considered to be acceptable.

Step 1. Bayesian network results

Following the process shown in Table 1, the most robust BN was identified and is presented in Figure 1 (Hill Climbing algorithm and mBDe scoring). The relationships that were repeated on more than 70% of the networks learned with the different algorithms and metrics (above the threshold of 11) are represented by a dotted line in the figure. These are the relationships of SQ with INV, SQ with PB, SQ with BI, and BI with PC. Nevertheless, other relationships have values which are close to the threshold, indicating that the model is fairly robust, as the structure usually repeats itself in almost all the learned networks.

The strength of the links was analysed using iterative resampling of the initial database (bootstrap resampling). Figure 1 also shows how often each link occurs in this iterative replication process. Almost all the links had high proportions and in some cases they were close to 1 (e.g. SQ with PB and PB with SA), showing the robustness of the model and its relationships. The strength of the links was then used to calculate the DWI considering the BI as the objective node. Only three latent constructs had any direct or indirect influence on the BI, which allowed us to rank the importance of these three constructs over BI: SQ was first, with a DWI=0.79, SAT was second with a DWI=0.33 and PB was in third place, with a DWI=0.11.



Note: INV: Involvement; SQ: Service quality; PB: Perceived benefits; BI: Behavioural intentions; SA: Satisfaction; PC: Perceived Costs. The value of the links indicates the score obtained by each relationship as the sum of the partial scores of each of the learnt models (17 in total). The value in brackets indicates the proportion each link occurs in the bootstrap resampling.

Figure 1. Most robust BN from the 17 learnt models (Hill Climbing algorithm and mBDe scoring).

Step 2. Validation of the relationships

In order to validate the model described in the previous section, a SEM model was calibrated (using AMOS v.22), using the 6 constructs (INV, PB, SQ, PC, BI and SA) as latent variables and all the items and dimensions that were not removed during the EFA and CFA (19 items and 7 dimensions) as observed variables as shown in Table 3.

Table 4. Regression Weights, Standard Errors and Goodness-of-Fit Indexes from SEM model.

			ERW	SRW	S.E.	Sig.
Measurement model relationships						
A21	<---	INV	1.000	0.736		
A22	<---	INV	0.976	0.611	0.032	***
A23	<---	INV	0.754	0.625	0.024	***
A25	<---	INV	0.933	0.672	0.028	***
A26	<---	INV	0.924	0.743	0.025	***
A9	<---	PB	1.000	0.811		
A12	<---	PB	0.940	0.653	0.026	***
A13	<---	PB	0.751	0.606	0.023	***
Acces	<---	SQ	1.000	0.706		
Cust_Serv	<---	SQ	1.149	0.679	0.032	***
Avail	<---	SQ	1.045	0.698	0.028	***
Ind_Space	<---	SQ	1.207	0.573	0.040	***
Tan_Eq	<---	SQ	0.906	0.694	0.025	***
Info	<---	SQ	1.073	0.678	0.030	***
SQ1	<---	SQ	1.101	0.748	0.028	***
Sec	<---	SQ	1.124	0.660	0.032	***
A4	<---	PC	1.000	0.819		
A5	<---	PC	0.997	0.816	0.023	***
A8	<---	PC	1.000	0.747	0.024	***
A18	<---	BI	1.000	0.775		
A19	<---	BI	1.023	0.702	0.028	***
A20	<---	BI	0.709	0.721	0.019	***
A1	<---	SA	1.000	0.592		
A2	<---	SA	1.010	0.699	0.033	***
A3	<---	SA	1.285	0.704	0.042	***
SQ2	<---	SA	1.045	0.775	0.032	***
Structural model relationships						
SA	<---	PB	0.443	0.511	0.040	***
INV	<---	SQ	0.715	0.451	0.035	***
PB	<---	SQ	1.068	0.845	0.029	***
BI	<---	SQ	0.384	0.282	0.050	***
SA	<---	SQ	0.405	0.371	0.046	***
PC	<---	BI	-0.612	-0.443	0.030	***
BI	<---	SA	0.622	0.499	0.050	***

Goodness-of-fit	Proposed Model	(28)	(4)	(7)	(19)
Sample Size	3,198	3,198	3,211	337	412
Chi-Squared	4,404.44	16,685.00	3,084.30	-	543.191
Degrees of freedom	292	1415	147	-	258
GFI	0.893	0.820	0.892	-	0.902
AGFI	0.872	-	0.861	-	0.880
RMSEA	0.066	0.060	0.079	0.045	0.047
RMR	0.051	-	0.173	0.060	-
NFI	0.882	-	0.902	0.904	-
CFI	0.889	0.850	0.906	0.911	-
PGFI	0.743	-	0.690	0.751	-
PNFI	0.792	-	-	0.750	-

Note: ERW. Estimate (non-standardized) Regression Weights; SRW. Standardized Regression Weights; S.E., Standard Error; Sig. *** significant at level $p < 0.001$.

Table 4 shows the results of this model, which were satisfactory: all the relationships were significant to 0.001 and all the standardised load weights were high ($SRW > 0.5$); in the structural model all the relationships are significant to 0.001. Note the negative sense of the relationship BI with PC. This was to be expected as the PC construct has negative connotations for the passengers.

Table 4 also shows the main fit indexes for this model. This study has used indexes from other works that used the same database and the same constructs (4; 28); and from other comparable works (similar items and constructs) such as Irtema et al. (19) and Li et al. (7).

As with the other models (24; 29-30), the model used in this paper rejects the X^2 test. However, the absolute fit indexes (GFI and AGFI) show values which are very close to the recommended threshold of 0.9 (0.893 and 0.872, respectively), which allowed them to be accepted. The RMSEA (0.066) is found to be lower than 0.08, and the RMR is equal to the recommended value of 0.05. The incremental fit indexes show similar values (NFI=0.882 and CFI=0.889) which are close to the recommended value of 0.9. Finally, the values of the parsimony indexes PGFI (0.743) and PNFI (0.792) are consistent with the recommendations set by (30). The overall fit of the model can therefore be considered as acceptable.

A comparison of these indexes with those of other works clearly shows the validity of this model. Similar values to those provided by the work of Irtema et al. (19) and Li et al. (7) have been obtained, with a slight improvement on the fits of Machado-Leon et al. (4, 28).

Finally, Table 5 shows the different kinds of effects produced between the constructs.

Table 5. Standardized Direct, Indirect and Total effects of SEM model.

		INV	PB	SQ	PC	BI	SA
Direct Effects	INV	0	0	0.451	0	0	0
	PB	0	0	0.845	0	0	0
	PC	0	0	0	0	-0.443	0
	BI	0	0	0.282	0	0	0.499
	SA	0	0.511	0.371	0	0	0
Indirect Effects	INV	0	0	0	0	0	0
	PB	0	0	0	0	0	0
	PC	0	-0.113	-0.302	0	0	-0.221
	BI	0	0.255	0.401	0	0	0
	SA	0	0	0.432	0	0	0
Total Effects	INV	0	0	0.451	0	0	0
	PB	0	0	0.845	0	0	0
	PC	0	-0.113	-0.302	0	-0.443	-0.221
	BI	0	0.255	0.683	0	0	0.499
	SA	0	0.511	0.803	0	0	0

Table 5 shows that all the effects whether direct, indirect or total, of the constructs PB, SQ, BI and SA on PC are negative, given that these have a positive attitude towards the service whereas PC has negative connotations. Therefore, the type of relationship between these constructs is expected to be negative. However, the construct that has the strongest effect on BI is SQ, followed by SA and finally PB. This ranking in importance SQ, SA and PB on BI coincides with the results of the DWI metric obtained by using the bootstrap resampling and the DAG of the BN. Note that involvement (INV) has no influence on BI. In this case, the fact that passengers have a pro-active attitude towards using PT does not imply that they feel tempted to return to using the LRT, it could be it has no utility for them or they are using another mode of PT. Only the SQ construct has a positive effect on INV, because if the passengers perceive the LRT as having good quality their general INV will be more positive. SQ has a direct influence on BI and, in turn, influences it indirectly through SA, which agrees with the results from previous works (1;4;12-13). The work of Machado-Leon et al. (4) calibrated various SEM models to model the role of INV in the relationships between SQ, SA and BI in different ways: as mediator, antecedent and moderator. Within the mediator role they proposed 3 different ways of mediating, one of them without any relationship between INV and BI, rather SQ and SA influenced INV, and SQ and SA influenced BI.

Therefore, this model has shown the validity of the proposed relationships, in other words: the lack of a relationship between INV and BI, and the relationship between SQ over INV and SQ over BI (directly and indirectly through SA). Finally, note that SQ is the only construct that has any influence over all the other constructs (directly and/or indirectly), making it a key factor in this behavioural model.

5. Discussion about Model

The model created with the BN and validated with the SEM has resulted in satisfactory adjustment parameters, in line with other research in this field (4, 7, 19, 28). A series of relationships have been found which have also been verified in other work (such as SQ-SA, SQ-BI, SA-BI, PB-SA, SQ-INV) and from which we have extracted specific relationships for this case study (SQ-PB and BI-PC) (see Figure 1). The design of the working models used as references was supported by the literature, whereas the model used in this research has been drawn directly from the data without any kind of imposition. Therefore, any improvement or similarity in the parameters, as well as certain equality in the relationships of the models, will support the suitability of this applied methodology for extracting and validating knowledge directly from the data.

Service quality shows a direct effect on the remaining constructs except perceived costs where the effect is indirect, showing the importance of this construct in passenger perceptions towards the service, and towards PT in general:

- An acceptable level of overall quality provided by the LRT service makes passengers have a better opinion about PT in general (e.g. environmental or social), because many of them are people with a predisposition to use the car. Therefore, if they perceive this service as a clear competitor to the car, in terms of quality and facilities, their involvement may improve (relationship SQ with INV). This relationship has also been identified in previous work (1, 4, 19).
- If the passengers' perceived quality of the LRT service is high they consider they have many benefited from it (relationship of SQ with PB). Similarly, this relationship has also been identified in other work (15, 31), although those cases referred to perceived value rather than perceived benefits.
- If the quality provided by the service contributes to fulfilling passenger expectations, it is verified its relationship with satisfaction. This relationship (SQ with SA) has been frequently identified in the literature (4, 7, 19, 28).
- Finally, the quality of service is going to establish certain conditions referring to money, time, comfort, etc., which if they are perceived as good will have a bearing on passenger loyalty towards the service, in other words, their intention to use the service again or recommend it to others (relationship SQ with BI). This relationship has also been identified previously (4, 7, 14, 19, 28).

The perceived benefits construct is found to be connected with satisfaction. This is because as the passengers feel they have obtained certain benefits having used the service (related to their perception of comfort and the fulfilment of their expectations), this will generate a positive feeling towards the service and, therefore, will have a bearing on their satisfaction. Various studies have identified this relationship between PB and SA (2; 28; 31).

Satisfaction is, in turn, related to behavioural intentions. The attraction of the LRT, with levels of comfort and expectations fulfilled, are aspects covered by the satisfaction construct and which directly influence a passenger when they decide to use the service again under the same conditions or recommend it to other people. This relationship (SA with BI) forms part of the well known service quality-satisfaction-behavioural intentions paradigm, which has been validated by several authors (11-13).

Behavioural intentions are found to be connected with perceived costs. If there is a low intention to repurchase, this influences the feeling that the costs of using the service are perceived to be high. Note that in the literature this relationship has normally been found in the opposite direction (PC influences BI) (1), therefore further analysis needs to be carried out in the present case to uncover whether this is a specific characteristic of this case study, or could it be generalised to other cases.

Finally, two statements can be made based on the results obtained in this work. The first is the suitability of applying this methodology to the study and analysis of models, allowing previous unidentified relationships to be discovered which could be generalised or which may be specific to a case in question. And, secondly, the possibility of applying the proposed methodology to other different fields from the perception of PT service quality (20), to explore new fields of knowledge, as well as confirm and/or improve existing models.

6. Conclusions

The application of the proposed two stage methodology has allowed us to obtain and validate a theoretical model underpinning behavioural intentions, passenger perceptions, satisfaction and involvement directly from the data obtained in a survey on the LRT system of Seville (Spain). Diez-Mesa et al. (20) applied the same methodology to explain service quality, but with the inclusion of certain relationships that had to be fulfilled based on their literature review, so expert knowledge was linked to automatic learning. However, in the present study the relationships between constructs have been extracted without any kind of constraint, obtaining satisfactory results which confirm accepted theories: the Service quality-Satisfaction-Behavioural intentions paradigm; the direct effect of service quality on behavioural intentions; or the influence of perceived benefits on satisfaction. Involvement and perceived benefits have also been found to be directly and positively influenced by service quality, therefore, any improvements made in this latter construct can help to change the mentality of passengers in favour of PT. As involvement has been found not to influence behavioural intentions, it opens up a new line of research to determine whether or not the absence of this relationship is specific to the case in question, the LRT service in Seville, or it can be extrapolated to other PT services. The results found in this research are quite rational as the LRT service is made up of only one line and does not provide a service to different zones in the city and therefore, although there is a high level of involvement with PT in general, there is no reason to have high levels of behavioural intentions, simply because this particular service does not suit the passenger.

The results obtained in this work show that the relationships automatically extracted by the BN and later validated by SEM are valid as they show a great similarity with the relationships established in other work based on previous knowledge and bibliographic review. Therefore, this methodology can be useful and relevant in the development of new theories, the identification of new relationships between variables and the confirmation of existing model validity. In fact, this methodology can be considered as a robust and suitable way to model a phenomenon straightforwardly from a database. In the age of big data in the transport field, the combined use of exploratory methods, based on data mining, together with confirmatory theories, such as structural equations modelling, could represent a revolution in the identification of new travel behaviour theories, as well as in other fields of transport.

In this specific case of the behavioural intentions towards a particular mode of public transport, the results obtained could be a great help for transport managers and planners, as they allow them to acquire a better understanding of passengers' opinions in order to develop more appropriate marketing strategies and tailor their services to meet passengers' needs, attracting new passengers from other modes and retaining existing ones.

7. Acknowledgements

Support from the Spanish Ministry of Economy and Competition (Research Project TRA2015-66235-R) is gratefully acknowledged. The authors would like to thank the FEDER of the European Union for financial support via the project “Mejora de la calidad del TP para fomentar la movilidad sostenible: Metro de Sevilla” of the “Programa Operativo FEDER de Andalucía 2007-2013”.

8. Author contributions

The authors confirm contribution as follows: study conception and design, Diez-Mesa, de Oña and de Oña; data collection, Diez-Mesa and de Oña; analysis and interpretation of results, Diez-Mesa, de Oña and de Oña; draft manuscript preparation, Diez-Mesa, de Oña, and de Oña. All authors reviewed the results and approved the final version of the manuscript.

9. References

- [1] Lai, W., and C., Chen. Behavioral Intentions of Public Transit Passengers – The Roles of Service Quality, Perceived Value, Satisfaction and Involvement. *Transport Policy*, 18, 2011, pp. 318-325.
- [2] De Oña, J., R., De Oña, L., Eboli, C., Forciniti, and G., Mazzulla, G. Transit Passengers’ behavioural Intentions: The Influence of Service Quality and Customer Satisfaction. *Transportmetrica A: Transport Science*, 12, 2016, pp. 385-412.
- [3] Jen, W., and K.C., Hu. Application of Perceived Value Model to Identify Factors Affecting Passenger’s Repurchase Intentions on City Bus: A Case Study of the Taipei Metropolitan Area. *Transportation*, 30, 2013, pp. 307-327.
- [4] Machado-Leon, J.L., R., de Oña, and J., de Oña. The Role of Involvement in Regards to Public Transit Riders’ Perceptions of the Service. *Transport Policy*, 48, 2016, pp. 34-44.
- [5] Parasuraman, A., V.A., Zeithaml, and L.L., Berry. SERVQUAL: A Multiple Item Scale for Measuring Customer Perceptions of Service Quality. *Journal of Retailing*, 64, 1988, pp. 12-40.
- [6] Oliver, R.L. *Satisfaction: A Behavioral Perspective on the Consumer*. M.E. Sharpe, Armonk, N.Y., 2010.
- [7] Li, L., Y., Bai, Z., Song, A., Chen, and B., Wu. Public Transportation Competitiveness Analysis Based on Current Passenger Loyalty. *Transportation Research Part A: Policy and Practice*, 113, 2018, pp. 213-226.

- [8] Park, J., R., Robertson, and C.L., Wu. The Effect of Airline Service Quality on Passengers' Behavioural Intentions: A Korean Case Study. *Journal of Air Transport Management*, 10, 2004, pp. 435-439.
- [9] Bolton, R.N., and J.H., Drew. A Multistage Model of Customers' Assessment of Service Quality and Value. *Journal of Consumer Research*, 17, 1991, pp. 375-384.
- [10] Oliver, R.L. Whence Customer Loyalty?. *Journal of Marketing*, 63, 1999, pp. 33-44.
- [11] Jen, W., R., Tu, and T., Lu. Managing Passenger Behavioral Intention: An Integrated Framework for Service Quality, Satisfaction, Perceived Value, and Switching Barriers. *Transportation*, 38, 2011, pp. 321-342.
- [12] Chou, P.F., C.S., Lu, and Y.H., Chang. Effects of Service Quality and Customer Satisfaction on Customer Loyalty in High-Speed Rail Services in Taiwan. *Transportmetrica A: Transport Science*, 10, 2014, pp. 917-945.
- [13] De Oña, R., J.L., Machado, and J., de Oña. Perceived Service Quality, Customer Satisfaction, and Behavioral Intentions: Structural Equation Model for the Metro of Seville, Spain. *Transportation Research Record: Journal of the Transportation Research Board*, 2538, 2015, pp. 76-85.
- [14] Minser, J., and V., Webb. Quantifying the Benefits: Applications of Customer Loyalty Modeling in Public Transportation Context. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2144. Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 111-120.
- [15] Zeithaml, V.A. Consumer Perceptions of Price, Quality and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, 52, 1988, pp. 2-22.
- [16] Choi, K.S., W.H., Cho, S., Lee, H., Lee, and C., Kim. The Relationships among Quality, Value, Satisfaction and Behavioral Intention in Health Care Provider Choice: A South Korean Study. *Journal of Business Research*, 57, 2004, pp. 913-921.
- [17] Wang, Y., H.P., Lo, R., Chi, and Y., Yang. An Integrated Framework for Customer Value and Customer Relationship-Management Performance: a Customer-Based Perspective from China. *Managing Service Quality: An International Journal*, 14, 2004, pp. 169-182.
- [18] Fu, X., J., Zhang, and F., Chan. Determinants of Loyalty to Public Transit: A Model Integrating Satisfaction-Loyalty Theory and Expectation-Confirmation Theory. *Transportation Research Part A: Policy and Practice*, 113, 2018, pp. 476-490.
- [19] Irtema, H., A., Ismail, M., Borhan, A., Das, and A., Alshetwi. Case Study of the Behavioural Intentions of Public Transportation Passengers in Kuala Lumpur. *Cases Studies on Transport Policy*, 6, 2018, pp. 462-474.

- [20] Diez-Mesa, F., R., De Oña, and J., De Oña. Bayesian Networks and Structural Equations Modelling to Develop Service Quality Models: Metro of Seville Case Study. *Transportation Research Part A: Policy and Practice*, 118, 2018, pp. 1-13.
- [21] Pearl, J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2000.
- [22] Kjaerulff, U.B., and A.L., Madsen. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis (2nd Ed.)*. Springer Science+Business Media, New York, 2013.
- [23] Bollen, K.A. *Structural Equations with Latent Variables*. Wiley, New York, 1989.
- [24] Golob, T.F. *Structural Equations Modeling for Travel Behavior Research*. *Transportation Research Part B*, 37, 2003, pp. 1-25.
- [25] Hair, J.F., W.C., Black, B.J., Babin, and R.E., Anderson. *Multivariate Data Analysis*. In P. Education Ed., 7th Edition. Prentice Hall, New Jersey, 2010.
- [26] Cugnata, F., R.S., Kenett, and S., Salini. Bayesian Networks in Survey Data: Robustness and Sensitivity Issues. *Journal of Quality Technology*, 48, 2016, pp. 253-264.
- [27] Albrecht, D., A.E., Nicholson, and C., Whittle, C. *Structural Sensitivity for the Knowledge Engineering of Bayesian Networks*. *Probabilistic Graphical Models*, pp. 1-16. Basel, Switzerland: Springer International Publishing, 2014.
- [28] Machado-Leon, J.L., R., de Oña, F., Diez-Mesa, and J., de Oña. Finding Service Quality Improvement Opportunities across Different Typologies of Public Transit Customers. *Transportmetrica A: Transport Science*, 14, 2018, pp. 761-783.
- [29] Hooper, D., J., Coughlan, and M.R., Mullen. *Structural Equation Modelling: Guidelines for Determining Model Fit*. *The Electronic Journal of Business Research Methods*, 6, 2008, pp. 53-60.
- [30] Mulaik, S.A., L.R., James, J., Van Alstine, N., Bennet, S., Lind, and C.D., Stilwell. Evaluation of Goodness-of-fit Indices for Structural Equation Models. *Psychological Bulletin*, 105, 1989, pp. 430-445.
- [31] Sumaedi, S., G., Bakti, and M., Yarmen. The Empirical Study of Public Transport Passengers' Behavioral Intentions: The Roles of Service Quality, Perceived Sacrifice, Perceived Value and Satisfaction (Case Study: Paratransit Passengers in Jakarta, Indonesia). *International Journal for Traffic and Transport Engineering*, 2, 2012, pp. 83-97.