

**UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO**

**“MEDICIÓN DE LA PERCEPCIÓN DE CALIDAD
EN LOS SERVICIOS PÚBLICOS”**

**Tesis sometida a consideración de la Comisión del
Programa de Estudios de Posgrado en Estadística para
optar por el grado y título de Maestría Académica en
Estadística**

JUAN PABLO SOLÍS ALVARADO

Ciudad Universitaria Rodrigo Facio, Costa Rica

2017

Dedicatoria y agradecimientos

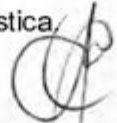
Dedico la confección de esta Tesis, con mucho cariño, a mi familia quienes siempre me han apoyado en todos los aspectos de mi vida y quienes me han impulsado a alcanzar cada una de mis metas y me han dado el valor necesario para lograr el éxito.

A mis hermanos y a mis padres por el apoyo brindado a lo largo de los años, especialmente a mi madre por su gran esfuerzo para darme la oportunidad de estudiar, y con todo mi cariño a mi esposa y a mi hijo, quienes han sido el principal soporte de mis logros alcanzados y con quienes comparto todos mis éxitos.

Un agradecimiento muy especial a los profesores miembros del comité asesor: Fernando Ramírez, Edgar Gutiérrez, Olman Ramírez, Guaner Rojas y Andrey Zamora por acceder a ser parte de este proyecto, por la guía y por los consejos que me brindaron a lo largo del proceso y que culminaron con el éxito alcanzado.

A la Escuela de Estadística y a todo el personal miembro de la misma, le agradezco el haberme aceptado en el Programa de Estudios de Posgrado y darme la oportunidad de realizar mis estudios superiores. Finalmente, agradezco a la Universidad de Costa Rica por la oportunidad y las facilidades brindadas para lograr este objetivo.

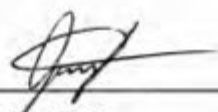
Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica, como requisito parcial para optar por el grado y título de Maestría Académica en Estadística.



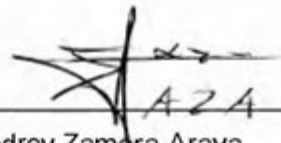
Ph.D. Gilbert Brenes Camacho
Representante del Decano
Sistema de Estudios de Posgrado



M.Sc. Fernando Ramírez Hernández
Director de Tesis



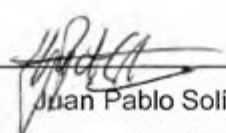
Ph.D. Guaner Rojas Rojas
Asesor



M.Sc. Andrey Zamora Araya
Asesor



M.Sc. Jenny Madrigal Pana
Director
Programa de Posgrado en Estadística



Juan Pablo Solís Alvarado
Candidato

Tabla de contenido

Dedicatoria y agradecimientos	ii
Tabla de contenido	iv
RESUMEN	viii
Lista de cuadros y tablas.....	ix
Lista de figuras	x
Lista de abreviaturas	xi
Glosario.....	xiii
I. Introducción.....	1
1.1 Problema de investigación.....	2
1.2 Objetivos	2
1.3 Justificación	3
1.4 Aplicabilidad	4
II. Marco Teórico	6
2.1 Servicios públicos	6
2.2 Calidad percibida en los servicios	8
2.2.1 Conceptualización de calidad percibida.....	9
2.2.2 Dimensiones y modelos de calidad en servicios	11
2.2.3 Modelo Gaps	15
2.2.4 SERVQUAL: Service Quality (Calidad del Servicio).....	18
2.3 Indicadores de calidad.....	20
2.3.1 Indicadores en función del formato	21
2.3.2 Indicadores en función del contenido.....	22
III Antecedentes	24
3.1 Interés en la calidad en los servicios públicos nacionales.....	24
3.2 Monitoreo de calidad en los servicios públicos nacionales	26
3.2.1 Diagnóstico de la Secretaría Técnica del SNCS	26

3.2.2 Índice de Gestión Institucional (IGI).....	28
3.2.3 Índice GpRD.....	29
3.4 Estudio: "Percepción en la Calidad de los Servicios Públicos".....	30
IV Marco Metodológico.....	34
4.1 Recolección de datos.....	34
4.2 Instrumentos de recolección.....	34
4.3 Población.....	35
4.4 Muestra.....	35
4.5 Conceptualización del índice de calidad percibida.....	36
4.6 Variables de estudio.....	37
4.6.1 Selección preliminar de variables.....	37
4.6.2 Análisis de datos faltantes (no respuesta).....	38
4.6.3 Tratamiento de datos faltantes (no respuesta).....	38
4.6.4 Estandarización de variables.....	40
4.7 Modelos ajustados.....	40
4.7.1 Regresión Lineal Múltiple (RLM).....	41
4.7.2 Análisis de Factores Confirmatorio (AFC).....	44
4.8 Evaluación del modelo.....	47
4.8.1 Regresión Lineal Múltiple (RLM).....	48
4.8.2 Análisis de Factores Confirmatorio (AFC).....	49
4.9 Índice calculado.....	55
4.9.1 Índice de calificación de calidad.....	55
4.9.2 Índice de calidad percibida.....	56
4.10 Fiabilidad.....	56
4.11 Validez.....	59
4.11.1 Evidencias de contenido.....	59
4.11.2 Evidencias de criterio.....	61

4.11.3 Evidencias de constructo	62
4.12 Invarianza factorial.....	64
4.13 Análisis de usuarios.....	67
4.13.1 Segmentación de usuarios	68
4.13.2 Perfiles de usuarios	69
V Resultados	70
5.1 Selección de variables	70
5.2 Análisis de datos faltantes (no respuesta).....	70
5.3 Tratamiento de datos faltantes.....	71
5.4 Modelo RLM	72
5.4.1 Ecuación de regresión	72
5.4.2 Ajuste del modelo	72
5.4.3 Índice de Calificación Calidad (ICC)	76
5.4.4 Fiabilidad.....	77
5.4.5 Validez.....	78
5.5 Modelo AFC.....	82
5.5.1 Ecuaciones estructurales.....	82
5.5.2 Ajuste del modelo AFC	84
5.5.3 Índice de Calidad Percibida (ICP).....	90
5.5.4 Fiabilidad	91
5.5.5 Validez.....	91
5.6 Invarianza factorial.....	96
5.7 Análisis de usuarios.....	98
5.7.1 Segmentación de usuarios	99
5.7.2 Perfiles de usuarios	100
VI Conclusiones y recomendaciones	102
6.1 Alcances y limitaciones del estudio.....	102

6.2 Discusión	104
6.3 Conclusiones	106
6.4 Recomendaciones para futuras investigaciones	108
VII. Referencias bibliográficas.....	110
VIII. Anexos	119
Anexo A. Variables de estudio	119
Anexo B. Simbología Modelos de Ecuaciones Estructurales.....	124
Anexo C. Análisis de no respuesta.....	125
Anexo D. Distribución de la no respuesta.....	126
Anexo E. Análisis de regresión lineal múltiple	128
Anexo F. Análisis de factores confirmatorio.....	132
Anexo G. Análisis de usuarios.....	144
Anexo H. Propuesta aspectos a evaluar.....	149

RESUMEN

El objetivo de la presente investigación es construir un índice de Calidad Percibida en Servicios Públicos, con evidencias suficientes de fiabilidad y validez, como complemento a las mediciones actuales sobre calidad en servicios públicos y que ayude a mejorar la toma de decisiones.

Se realizó una revisión bibliográfica sobre los fundamentos teóricos del concepto calidad percibida, así como de los modelos establecidos para medirlo. Se utilizaron el modelo Gaps y la escala ServQual como esquema conceptual para definir el concepto y proponer el modelo de medición. La calidad percibida en servicios se definió como la diferencia entre las expectativas y las percepciones de los usuarios, al momento de hacer uso de los servicios.

De acuerdo al esquema conceptual, el concepto de calidad percibida en servicios se compone de cinco dimensiones: confiabilidad, responsabilidad, aseguramiento, empatía y elementos tangibles. Se pusieron a prueba dos modelos para calcular un índice de calidad percibida en servicios públicos: un modelo de Regresión Lineal Múltiple (RLM) y un modelo de Análisis de Factores Confirmatorio (AFC) de segundo orden.

Los modelos se estimaron en una muestra de 1.041 individuos a quienes se les aplicó un cuestionario para evaluar la calidad de cinco servicios públicos brindados por instituciones estatales. Los modelos se sometieron a pruebas de validez y fiabilidad en una segunda muestra de 1.035 personas evaluando los mismos cinco servicios. Los datos se extrajeron del estudio "Percepción en la Calidad de los Servicios Públicos" realizado por la Escuela de Estadística de la Universidad de Costa Rica, entre 2009 y 2011.

El modelo AFC de segundo orden presentó la mejor bondad de ajuste y logró cumplir con los diagnósticos respectivos. El índice de calidad percibida en servicios públicos se estimó a partir de las cargas factoriales del factor de segundo orden. Adicionalmente se estimaron cinco subíndices para medir cada una de las dimensiones del modelo, a partir de las cargas factoriales de primer orden.

El índice calculado presentó evidencias favorables de fiabilidad, medidas por los coeficientes alfa de Cronbach, alfa ordinal, Spearman-Brown y Guttman-Flanagan. Se recolectaron evidencias de contenido, a partir del esquema conceptual; evidencias de criterio, al obtenerse resultados congruentes con las mediciones externas disponibles y evidencias de constructo por medio del análisis de la estructura interna, diversos indicadores de validez, pruebas de invarianza factorial y análisis de validación cruzada.

De acuerdo con el índice calculado, los servicios mejor calificados son los brindados por el Registro Nacional y los de menor calidad percibida son los brindados por la SETENA. Finalmente, se obtuvo que la edad y el nivel educativo de los usuarios están relacionados con los puntajes del índice, de forma positiva e inversa, respectivamente. Se identificaron tres perfiles de usuarios: exigentes, neutrales y flexibles.

Lista de cuadros y tablas

3.1	Tamaños de muestra finales. Estudio Percepción de la Calidad en Servicios Públicos. 2008-2010.....	32
4.1	Ecuaciones estructurales. Modelo AFC de segundo orden.....	47
4.2	Restricciones impuestas: variables con cargas fijadas en 1.....	51
4.3	Índices de ajuste. Modelo AFC.....	53
5.1	Estadísticos de ajuste. Modelo RLM. Muestra de calibración.....	73
5.2	Significancia de coeficientes de regresión. Muestra de calibración	73
5.3	Coefficientes de fiabilidad. Modelo RLM. Ambas aplicaciones.....	77
5.4	Análisis de Factores Exploratorio. Solución rotada. Modelo RLM. Muestra de validación.....	78
5.5	Puntaje del ICC y del IGI por servicio y coeficiente de validez, según año de aplicación. Modelo RLM. 2008 a 2011.....	80
5.6	Inconformidades externas e ICC estimado por dimensión. Modelo RLM. 2008 a 2010.....	80
5.7	Ecuaciones estructurales estimadas. Modelo AFC	82
5.8	Índices de ajuste estimados. Modelo AFC. Muestra de calibración.....	87
5.9	Coefficientes de fiabilidad. Modelo AFC. Ambas muestras.....	91
5.10	Análisis de Factores Exploratorio. Solución rotada. Modelo AFC. Muestra de validación	93
5.11	Puntaje del ICP y del IGI por servicio y coeficiente de validez, según año de aplicación. Modelo AFC. 2008 a 2011.....	94
5.12	Inconformidades externas e ICP estimado por dimensión. Modelo AFC. 2008 a 2010.....	95
5.13	Coefficientes de validez por dimensión. Modelo AFC. Muestra de calibración...	96
5.14	Análisis de invarianza factorial. Modelo AFC.....	97
5.15	Análisis de validación cruzada. Modelo AFC.....	98
5.16	Perfiles de usuarios. Muestra de calibración.....	101

Lista de figuras

2.1	Modelo de la Calidad Total Percibida.....	13
2.2	Modelo Gaps.....	16
4.1	Esquema conceptual de calidad percibida.....	36
4.2	Dimensiones y temas. Modelo de Regresión Lineal Múltiple.....	42
4.3	Modelo AFC de segundo orden. Diagrama de trayectoria.....	46
5.1	Modelo AFC de segundo orden. Solución completamente estandarizada.....	83

Lista de abreviaturas

- AFC: Análisis de Factores Confirmatorio (CFA, por sus siglas en inglés)
- AFE: Análisis de Factores Exploratorio (EFA, por sus siglas en inglés)
- AGFI: Adjusted Goodness of Fit Index (Índice Ajustado de Bondad de Ajuste)
- ANDEVA: Análisis de varianza (ANOVA, por sus siglas en inglés)
- BID: Banco Interamericano de Desarrollo
- CEA: Comisión de Eficiencia Administrativa
- CEARE: Comisión de Eficiencia Administrativa y Reforma del Estado
- CFI: Comparative Fit Index (Índice de Ajuste Comparativo)
- CGR: Contraloría General de la República
- CI: Centrality Index (Índice de Centralidad)
- CN: Critical N (N Crítico)
- DFOE: División de Fiscalización Operativa y Evaluativa
- DGME: Dirección General de Migración y Extranjería
- DGSC: Dirección General de Servicio Civil
- ECVI: Expected Cross Validation Index (Índice Esperado de Validación Cruzada)
- FIV: Factor de Incremento de la Varianza
- FEP: Factores de Ejes Principales
- GFI: Goodness of Fit Index (Índice de Bondad de Ajuste)
- GpRD: Gestión para Resultados en el Desarrollo
- ICP: Índice de Calidad Percibida
- ICC: Índice de Calificación de Calidad
- IFI: Incremental Fit Index (Índice de Ajuste Incremental)
- IGI: Índice de Gestión Institucional
- IGM: Índice de Gestión Municipal
- IT: Índice de Tolerancia
- MAPI: Matriz de Planificación Institucional
- MCP: Mínimos Cuadrados Ponderados
- MEE: Modelos de Ecuaciones Estructurales (SEM, por sus siglas en inglés)
- MH: Ministerio de Hacienda
- MIDEPLAN: Ministerio de Planificación Nacional y Política Económica
- MOPT: Ministerio de Obras Públicas y Transportes

- NFI: Normed Fit Index (Índice de Ajuste Normado)
- NNFI: Non Normed Fit Index (Índice de Ajuste No Normado)
- PGFI: Parsimony Goodness of Fit Index (Índice Parsimonioso de Bondad de Ajuste)
- PME: Programa de Modernización del Estado
- PND: Plan Nacional de Desarrollo
- PNFI: Parsimony Normed Fit Index (Índice Parsimonioso de Ajuste Normado)
- PNUD: Programa de las Naciones Unidas para el Desarrollo
- POI: Plan Operativo Institucional
- PRODEV: Programa de Implementación del Pilar Externo del Plan de Acción a Mediano Plazo para la Efectividad en el Desarrollo
- RFI: Relative Fit Index (Índice de Ajuste Relativo)
- RLM: Regresión Lineal Múltiple
- RMR: Root Mean Square Residual (Raíz del Residuo Cuadrático Medio)
- RMSEA: Root Mean Square Error of Approximation (Raíz del Error Cuadrático Medio de Aproximación)
- RN: Registro Nacional
- SEP: Sistema de Evaluación PRODEV
- SERVQUAL: Service Quality (Calidad de Servicio)
- SETENA: Secretaría Técnica Nacional Ambiental
- SIDES: Sistema de Indicadores de Desarrollo
- SINE: Sistema Nacional de Evaluación
- SNCS: Sistema Nacional de Contralorías de Servicios
- SNP: Sistema Nacional de Planificación
- SRMR: Standardized Root Mean Squared Residual (Raíz del Residuo Cuadrático Medio Estandarizado)
- TCT: Teoría Clásica de los Tests
- TRI: Teoría de Respuesta al Ítem
- UCR: Universidad de Costa Rica
- USES: Unidad de Servicios Estadísticos

Glosario

- Administración Pública: “el Estado y los demás entes públicos, a los cuales el ordenamiento jurídico les reconoce personalidad jurídica y capacidad de Derecho público y privado los cuales realizan actividad o función administrativa” (MIDEPLAN, 2007a pág.11). “Actividad administrativa de los órganos del Estado en todas sus escalas o jerarquías” (Ossorio, 2007, pág.45).
- Bien: “todos los objetos que, por útiles y apropiables, sirvan para satisfacer las necesidades humanas” (Cabanellas de Torres, 2011, pág.53). “Destinados a satisfacer necesidades o deseos humanos” (Ossorio, 2007, pág.109).
- Calidad en servicios: “medida del grado en que un servicio cumple con las especificaciones necesarias para satisfacer las necesidades y expectativas de los ciudadanos. Generalmente está determinada por un estándar o norma” (MIDEPLAN & DGSC, 2010, pág. 6).
- Ciudadanos: “personas que viven dentro de los límites territoriales de un Estado, poseen derechos y están sujetos a obligaciones públicas” (Sancho, 1999, pág.44). “El residente en alguna ciudad o Estado libre, cuando sus leyes y Constitución le dan ciertos derechos” (Cabanellas, 2011, pág.74).

El término ciudadano incluye tanto a los usuarios como a la colectividad social restante que no hace uso de los servicios públicos, pero que podría hacerlo (Sancho, 1999).

- Cliente: “designa a los ciudadanos, empresas y grupos organizados con los que está en contacto la Administración” y que demandan bienes y servicios (Sancho, 1999, pág.45). En el sector público, el término ‘cliente’ incorpora en sí diferentes conceptos tales como: usuario, administrado, contribuyente y beneficiario (Sancho, 1999).
- Contraloría de servicios: “instancia de participación ciudadana que facilita la comunicación entre los (as) usuarios (as) y la institución responsable de prestar los servicios públicos, con el fin de velar por la calidad de los servicios, la satisfacción del usuario (a) y el uso racional de los recursos públicos” (MIDEPLAN, 2009, pág.5).

- Eficiencia: "logro de los objetivos previamente establecidos, utilizando un mínimo de recursos" (MIDEPLAN & DGSC, 2010, pág.8). "Grado de optimización del resultado obtenido de la actuación pública en relación con los recursos que se han empleado para proveer los servicios públicos" (Sancho, 1999, pág.163).
- Entidad: "colectividad, institución, establecimiento, agrupación o empresa" (Ossorio, 2007, pág.370).
- Estado: "grupo de individuos establecidos sobre un territorio determinado y sujetos a la autoridad de un mismo gobierno" (Ossorio, 2007, pág.380). "Conjunto de personas con comunes caracteres culturales, históricos y sociales regidos por las mismas leyes y un solo gobierno / Conjunto de los poderes públicos" (Cabanellas, 2011, pág.168).
- Gobierno: "dirección o administración del Estado" (Cabanellas de Torres, 2011, pág.196). "Conjunto de los ministros superiores de un Estado" (Ossorio, 2007, pág.437). "Sistema de gestión y control realizados en la administración de la organización" (MIDEPLAN & DGSC, 2010, pág.10).
- Indicador: "es un valor identificable en una variable, que sirve para medir su comportamiento en función del nivel de logro de una meta planificada" (MIDEPLAN & DGSC, 2010, pág.10). "Número que expresa la relación entre cantidad y la frecuencia de un fenómeno" (Cabanellas, 2011, pág.220).

"Estadístico que proporciona información relevante sobre alguna característica relativa a la calidad de la prestación de los servicios públicos" (López y Gadea, 1995, citados por Sancho, 1999, pág.160).

- Institución pública (organización pública, entidad pública, institución estatal): "organismo que desempeña una función de interés público" (RAE, 2011¹). "Organismo que cumple con una función de interés público y que es parte del Gobierno" (Observatorio Ciudadano, 2011)².

¹ Tomado de http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=institución

² Tomado de http://cgrobservatorio.cqr.go.cr/blog/?page_id=112

- Plan Nacional de Desarrollo: “marco orientador del Gobierno de la República que define las políticas que normarán la acción del gobierno para promover el desarrollo del país”. (MIDEPLAN & MH, 2013, pág.10).
- Poderes de la República (poderes del Estado, poderes públicos): “cada uno de los órganos fundamentales de que el Estado se sirve o en los que se concreta y personifica al ejercer su soberanía territorial y personal” (Cabanellas, 2011, pág.331). “Órganos fundamentales del Estado. Son tres: el Legislativo, el Ejecutivo y el Judicial” (MIDEPLAN, 2007a, pág.13).
- Sector público: “conjunto de organismos públicos que incluye: El Poder Legislativo, Poder Ejecutivo, Poder Judicial y organismos públicos autónomos, instituciones, empresas y personas que realizan alguna actividad económica en nombre del Estado” (Observatorio Ciudadano, 2011)³

“Es el conjunto total de las organizaciones públicas. Lo integran los Poderes de la República, las instituciones autónomas, las municipalidades, los Bancos del Estado, las empresas públicas instituciones públicas no estatales” (MIDEPLAN, 2007a pág.13).

- Servicio: “toda prestación de carácter intangible que contribuye a la satisfacción de una necesidad, sea individual o colectivamente” (MIDEPLAN & DGSC, 2010, pág.14).
- Servicio público: servicio “de carácter técnico, prestado al público de manera regular y continua, para satisfacer una necesidad pública y por una organización pública” (Ossorio, 2007, pág.887). Servicio “que ha de satisfacer una necesidad colectiva por medio de una organización administrativa o regida por la Administración pública” (Cabanellas, 2011, pág.390).
- Usuario: “es toda persona física o jurídica que tiene derecho a demandar servicios ante las entidades públicas” (MIDEPLAN, 2009, pág. 5). Se refiere al sector social que “efectivamente utiliza las prestaciones ofrecidas por los poderes públicos” (Sancho, 1999, pág.44).

³ Tomado de http://cgobservatorio.cgr.go.cr/blog/?page_id=112

I. Introducción

Las Administraciones del Gobierno de Costa Rica han puesto de manifiesto la preocupación por incrementar el rendimiento y calidad de las instituciones estatales, en el cumplimiento de su función: dar respuesta a las demandas de la ciudadanía para satisfacción de las necesidades sociales, mediante la prestación de servicios públicos.

En el Plan Nacional de Desarrollo (PND), el Programa de Modernización del Estado (PME), el Sistema Nacional de Planificación (SNP), la Comisión de Eficiencia Administrativa y Reforma del Estado (CEARE) y el Sistema Nacional de Evaluación (SINE), se proponen objetivos y metas dirigidos a monitorear e incrementar el rendimiento de las instituciones públicas en la prestación de servicios.

Los instrumentos disponibles (diagnósticos del Sistema Nacional de Contralorías de Servicios -SNCS- y el Índice de Gestión Institucional -IGI-) están dirigidos a evaluar la gestión global de las instituciones estatales, incluyendo el tema de prestación de servicios públicos. Estos instrumentos son completados por el personal de las instituciones induciendo a una evaluación independiente, mediante indicadores propios de cada área, impidiendo realizar comparaciones del resultado de la evaluación en diferentes entidades.

El desarrollo de este trabajo surge con el fin de realizar una propuesta para abordar la problemática de medir la calidad percibida por los usuarios de servicios brindados en el sector público, a nivel nacional, en una base común que permita realizar comparaciones y fomentar la toma de decisiones. Esta propuesta consiste en el desarrollo de un indicador compuesto válido y confiable, que permita medir calidad percibida.

Este documento incluye la descripción del contexto de servicios públicos, la conceptualización del concepto "calidad percibida" y su aplicación en los servicios, el análisis de modelos utilizados para su medición y su composición.

El estudio describe el cálculo de un índice para medir calidad percibida en servicios públicos, la metodología aplicada, el modelo utilizado, la estrategia analítica (técnicas estadísticas) aplicada para dar respuesta al problema de investigación, además de una evaluación de cinco servicios públicos y un análisis de sus usuarios.

1.1 Problema de investigación

Existe un interés permanente en la Administración del Gobierno para medir, evaluar y monitorear el rendimiento y desempeño de las instituciones públicas, debido a que constantemente se cuestiona la capacidad de dichas instituciones para cumplir con su cometido de brindar servicios de calidad a la ciudadanía (CEARE, 2008).

El problema que motiva esta investigación es la ausencia de una medición de calidad percibida, con suficiente fundamento teórico y con criterios de validez y confiabilidad que englobe en un sólo valor numérico la evaluación deseada.

El desarrollo del presente trabajo conlleva a dar solución al siguiente problema de investigación: ¿Cómo medir la percepción de los usuarios sobre la calidad de los servicios públicos en Costa Rica, por medio de un indicador que permita evaluar la prestación de servicios por parte de las instituciones del Estado, realizar comparaciones y mejorar la toma de decisiones?

1.2 Objetivos

Con el fin de dar respuesta al problema de investigación se proponen los siguientes objetivos:

Objetivo general:

- Construir un índice compuesto que permita medir y evaluar la calidad, percibida por los usuarios, de los servicios públicos ofrecidos por cinco instituciones estatales de Costa Rica, que contribuya a mejorar la toma de decisiones.

Objetivos específicos:

- Proponer un modelo teórico-conceptual, a partir de investigaciones previas, que brinde las bases para construir un índice de calidad percibida en los servicios públicos.

- Identificar los factores asociados a la percepción de la calidad en los servicios públicos y medir sus relaciones y efectos, según el modelo conceptual propuesto.
- Medir el nivel de calidad, percibida por los usuarios, de los servicios de cinco instituciones públicas por medio de un índice, construido a partir de la información recolectada por las encuestas sobre percepción de los servicios públicos.
- Medir el nivel percibido por los usuarios de los servicios brindados por cinco instituciones públicas, de las dimensiones del índice de calidad percibida, por medio de indicadores individuales.
- Realizar una caracterización, por medio de análisis multivariantes, de los usuarios de los servicios de cinco instituciones públicas de acuerdo a los niveles en el índice de calidad percibida.

1.3 Justificación

Los sistemas de planificación y evaluación nacionales no disponen de una medición numérica estandarizada, estadísticamente válida y confiable, para evaluar la calidad percibida en los servicios brindados por las instituciones del sector público, desde la perspectiva de los usuarios.

El análisis de la calidad de los servicios públicos nacionales se realiza, principalmente, por medio del informe anual que las contralorías de servicios institucionales entregan a la Secretaría Técnica del SNCS. Mediante este informe se evalúan los resultados de la gestión de las contralorías de servicios, a su vez, son utilizados para identificar las principales deficiencias que poseen las instituciones al momento de brindar servicios y que requieren atención (MIDEPLAN, 2009).

En el informe anual, las contralorías institucionales, realizan un análisis cualitativo de la institución a partir de las inconformidades presentadas por los usuarios (MIDEPLAN, 2011c). Dicho análisis resulta incompleto por cuanto no se toma en cuenta la percepción de los usuarios de los servicios públicos que no presentan inconformidades.

El análisis se realiza desde la perspectiva interna de la organización, por esta razón, se hace necesaria la aplicación de otros instrumentos complementarios que permitan conocer las percepciones de todos los usuarios (Madrigal, 2009).

Por medio del informe anual se evalúan aspectos tales como: accesibilidad y calidad de la información, trato a los usuarios, calidad del servicio y estado de instalaciones; pero no posibilita una medición cuantitativa, normalizada y objetiva, que permita detectar diferencias entre instituciones y realizar comparaciones.

El desarrollo de un índice de calidad percibida en los servicios públicos permitirá cuantificar y evaluar objetivamente, en una escala numérica, la calidad percibida por los usuarios de los servicios brindados por entidades públicas. Adicionalmente, comparar la gestión de las distintas instituciones públicas en la prestación de servicios e identificar las áreas de mayor debilidad a ser reforzadas.

De acuerdo con la CEARE (2008) la importancia de impulsar una mejora continua en la prestación de servicios públicos se debe a que, cuando se brindan servicios de calidad, éstos contribuyen de manera significativa al bienestar y al mejoramiento de las condiciones de vida de todos los ciudadanos.

1.4 Aplicabilidad

Con un índice de calidad percibida en servicios públicos y el informe anual de labores de las contralorías de servicios institucionales, la Secretaría Técnica del SNCS estará en capacidad de realizar un análisis de la calidad de los servicios, integrando la perspectiva interna de la institución y la perspectiva externa de los usuarios, en aras de mejorar la prestación de servicios públicos.

Se plantea utilizar el índice de calidad percibida en servicios públicos como un mecanismo adicional de control de calidad del SINE, para evaluar si las acciones propuestas en los instrumentos de planificación, tienen efecto en la evaluación y mejoramiento en la prestación de servicios públicos.

Se visualiza al índice de calidad percibida en servicios públicos como un componente del sistema de monitoreo de los servicios públicos, planteado por el Ministerio de Planificación Nacional (MIDEPLAN), a utilizarse para la búsqueda de la excelencia en la gestión pública. Cumpliendo así con lo establecido en la Carta Iberoamericana de la Calidad en la Gestión Pública (Madrigal, 2009).

El índice de calidad percibida en servicios públicos se propone para ser utilizado como un parámetro adicional en el otorgamiento del "Premio Nacional a la Calidad en la Gestión Pública" y el "Reconocimiento a Prácticas Promisorias en la Gestión Pública", además se plantea como integrante del Sistema de Indicadores de Desarrollo (SIDES).

II. Marco Teórico

La delimitación del ámbito de acción de la presente investigación, requiere exponer el esquema teórico del cual se desprende el modelo conceptual de calidad percibida.

En este capítulo se realiza una descripción de las características de los servicios públicos y su proceso de diseño, se define el concepto de calidad percibida en los servicios y se analizan los modelos propuestos para su medición, además se incluye una breve descripción de lo qué son los indicadores de calidad.

2.1 Servicios públicos

Los servicios públicos corresponden a la respuesta del Gobierno ante las demandas realizadas por distintos grupos sociales, que buscan auxilio para la satisfacción de sus necesidades, individuales y colectivas. Por medio de los servicios públicos se responden las necesidades sociales con el fin de incrementar el bienestar colectivo (Sancho, 1999).

A diferencia de los servicios brindados en el entorno privado, donde la finalidad es obtener un beneficio económico, el objetivo de la prestación de servicios públicos es el bienestar social (Sancho, 1999). Según Esping-Andersen (1993, citado por Sancho, 1999) la prestación de servicios públicos pretende realizar una distribución equitativa de las riquezas entre los diversos grupos sociales.

El Estado puede brindar servicios públicos de dos formas: directamente, cuando los bienes y servicios son producidos o brindados por entidades públicas e indirectamente, con la producción o prestación privada de bienes y servicios pero, con financiación y regulación pública (Sancho, 1999).

Debido a que la cantidad de demandas presentadas por los grupos sociales al Gobierno es grande y que la actuación de éste último se ve limitada en recursos, no se puede hacer frente a todas las peticiones ciudadanas. No todas las necesidades sociales son consideradas problemas públicos a ser resueltos por el Gobierno (Subirats, 1989, citado por Sancho, 1999).

De acuerdo con Sancho (1999) el diseño de la oferta de servicios públicos debe considerar el nivel de receptividad de la Administración; es decir, su capacidad de dar respuesta a las necesidades sociales. Por ello, el Gobierno realiza un proceso de sistematización, interpretación y priorización de las necesidades ciudadanas.

Al ser, la oferta de servicios públicos limitada, la misma debe definirse tomando en cuenta condiciones tales como: igualdad de trato, distribución espacial del beneficio y equidad, procurando que la atención de necesidades beneficie a la mayor cantidad posible de ciudadanos.

Tomando en consideración lo anterior, el diseño de los servicios públicos se puede estructurar en los siguientes pasos (Sancho, 1999):

- 1) Identificar y estudiar las demandas de la ciudadanía: para ello se pueden utilizar diversos métodos de recolección de datos, cuantitativos y cualitativos, como la realización de: encuestas por muestreo, entrevistas a profundidad, grupos focales u observación directa. Aplicando en cada caso su respectiva metodología.

El estudio de las necesidades sociales se realiza con el fin de identificar los elementos a incluir dentro de los servicios para dar respuesta al requerimiento de los ciudadanos, mientras que el estudio de las características del ciudadano, se realiza con el objetivo de ajustar el servicio a su realidad concreta (Sancho, 1999).

- 2) Realizar un análisis del entorno: en este análisis se estudian tanto las características del entorno actual como del futuro, con el fin de prever posibles problemas. En el análisis es importante considerar los factores políticos, sociales, culturales, económicos, de producción, tecnológicos y ecológicos que intervienen en la prestación de los servicios públicos (Godet, 1993, citado por Sancho, 1999).

Entre las técnicas utilizadas para analizar el entorno se encuentran: el análisis FODA (Fortalezas, Oportunidades, Debilidades y Amenazas), el método de los escenarios (futuros alternativos posibles) o el método Delphi (técnica del panel de expertos).

- 3) Diseñar la estructura del servicio: se refiere a la estructuración de los elementos de la oferta. Eiglier y Langeard (1989, citados por Sancho, 1999) utilizan el término "servucción" para referirse al diseño de la producción de un servicio. Este diseño debe garantizar la disponibilidad de recursos materiales, humanos y económicos necesarios, así como los procesos que se requieran para su producción y prestación.
- 4) Prever las estrategias de comunicación: las políticas de comunicación deben girar en torno a un sistema de información y participación del ciudadano, al trato con los medios de comunicación y a la relación con agentes sociales (asociaciones y grupos empresariales y gubernamentales).

Entre los instrumentos disponibles para efectuar las relaciones públicas se encuentran: la definición de hechos noticiables (crear noticias favorables sobre la organización, el servicio o el personal), la organización de conferencias (difusión de la imagen de la organización) y las publicaciones (difusión de material informativo).

- 5) Establecer sistemas de control de calidad: se refiere a un proceso de monitoreo constante y dinámico de la prestación del servicio, procurando una mejora continua de los procesos de la organización con el fin de garantizar su calidad y adaptación a las expectativas de sus usuarios (Gröroos, 1994, citado por Sancho, 1999).

En resumen, la prestación de servicios públicos es un proceso complejo que incluye: la identificación de necesidades sociales demandadas por la ciudadanía, el análisis de los usuarios o beneficiarios y su entorno, el diseño y prestación del servicio, su publicidad y el establecimiento de un proceso de evaluación que procure su mejora continua.

2.2 Calidad percibida en los servicios

En el siguiente apartado se examinan diferentes esquemas conceptuales sobre calidad percibida aplicada a los servicios, se analizan algunas definiciones de este concepto y su relación con otros conceptos y se estudian dos corrientes de pensamiento que determinan los principios conceptuales propuestos para su definición y medición.

La medición de este concepto requiere tomar en cuenta todas las perspectivas y aspectos que intervienen en su conceptualización, a fin de poder establecer una definición clara del concepto a medir y de los parámetros bajo los cuales será medido.

2.2.1 Conceptualización de calidad percibida

La calidad en los servicios es distinta de la calidad en los bienes. La calidad en los bienes materiales, al tratarse de elementos tangibles, puede ser definida por medio de variables como color, olor, dureza o estilo y medida objetivamente utilizando diversos indicadores como durabilidad o cantidad de defectos. (Crosby, 1979; Garvin, 1983; Parasuraman, Zeithaml & Berry, 1988).

La calidad en los servicios, por el contrario, es un concepto confuso y difícil de definir, debido a tres características únicas de los servicios: intangibilidad, heterogeneidad e inseparabilidad de la producción y el consumo (Parasuraman, Zeithaml & Berry, 1985). Estas características diferencian los servicios de los bienes.

Según Casino (1999, citado por Díaz & Pons, 2009) la definición de calidad en servicios ha evolucionado. Inicialmente se asoció a la calidad en los bienes, después se le agregó la perspectiva de los consumidores, vinculándola con la satisfacción, pasando de una calidad objetiva a una calidad subjetiva (Morin & Jallais, 1991, citados por Díaz & Pons, 2009, pág.2).

La calidad subjetiva se refiere a un concepto humanístico, mientras la calidad objetiva a un concepto mecánico. La calidad mecánica, por un lado, es objetiva al considerar las características de un producto. La calidad humanística, por el otro, se refiere a la respuesta subjetiva de las personas a los productos (Holbrook & Corfman, 1985).

Para Olshavsky (1985, citado por Parasuraman et al., 1988) la calidad es una forma de evaluación total de un producto y puede ser visto como una actitud. Díaz & Pons (2009) apoyan la noción de que calidad es similar a una actitud. La calidad en los servicios también está asociada a la satisfacción. De acuerdo con Díaz & Pons (2009) hay una relación recíproca entre ambos conceptos.

La relación entre calidad y satisfacción surge con la confirmación de las expectativas del usuario (Smith & Houston, 1982, citados por Parasuraman et al., 1985). Para Bolton & Drew, (1991, citados por Díaz & Pons, 2009) tanto la calidad del servicio como la satisfacción nacen de la diferencia entre las expectativas y la prestación del servicio.

Aunque existe una relación entre ambos conceptos (calidad y satisfacción), para Díaz & Pons (2009) la calidad se refiere a una valoración global mientras que la satisfacción a una evaluación específica. Esta interpretación también es asumida por Parasuraman et al. (1988), estos autores asocian la calidad percibida con la actitud y la diferencian de la satisfacción.

Parasuraman et al. (1988) también hacen una diferenciación entre la calidad objetiva (actitud que contrasta las percepciones con las expectativas) y la calidad percibida (valoración de la excelencia del proveedor del servicio). Parasuraman et al. (1988) mencionan otros autores que definen calidad del servicio como la diferencia entre las expectativas y las percepciones, tales como: Sasser, Olsen & Wyckoff (1978), Grönroos (1982), Lehtinen & Lehtinen (1982) y Lewis y Booms (1983).

Según Parasuraman et al. (1988), el término expectativas puede obtener connotaciones distintas de acuerdo con el ámbito de estudio. En la calidad del servicio, las expectativas se refieren a lo que debería ser ofrecido por el proveedor (servicio deseado), pero en términos de satisfacción, las expectativas se pueden ver como pronósticos sobre lo que podría ser ofrecido (servicio posible).

Para Chías (1995, citado por Sancho, 1999) la calidad en los servicios debe ser variante en tanto las expectativas y las necesidades de los usuarios también cambian en el tiempo.

En cuanto a las percepciones, Parasuraman et al. (1988) las definen como opiniones sobre el servicio recibido, las cuales resultan de comparar las expectativas con el resultado de la prestación del servicio.

Dado que la calidad en los servicios se relaciona con otros aspectos de la prestación como la actitud, la satisfacción, las expectativas y las percepciones, es razonable indicar que la calidad en los servicios corresponde a un concepto de más de una dimensión que integra elementos tangibles e intangibles (Díaz & Pons, 2009).

Para Díaz & Pons (2009) “la calidad percibida por los clientes es una actitud, diferente a la satisfacción específica y equivalente a la satisfacción global, es considerada una variable multidimensional que debe tomar en cuenta las expectativas deseadas y/o las esperadas y las percepciones de los clientes” (pág.8).

Parasuraman et al. (1988) consideran que la “calidad del servicio percibida es por lo tanto vista como el grado y dirección de la discrepancia entre las percepciones y expectativas de los consumidores” (pág.17).

Para efectos de la presente investigación, siguiendo la definición planteada por Parasuraman et al. (1988), se considera calidad percibida en los servicios como la valoración realizada por los usuarios sobre el servicio, producto de la comparación entre la percepción del servicio recibido y las expectativas de lo que esperaban recibir.

2.2.2 Dimensiones y modelos de calidad en servicios

A continuación se discuten varias propuestas de operacionalización del concepto calidad percibida en servicios, las cuales han culminado con la postulación de diversos modelos para su medición y análisis, compuestos por distintas dimensiones.

En 1978 Sasser, Olsen y Wyckoff propusieron medir calidad en servicios a partir de tres dimensiones: materiales, facilidades y personal (Parasuraman et al., 1985). Lehtinen y Lehtinen (1982, citados por Díaz & Pons, 2009), propusieron tres dimensiones de calidad en servicios: calidad física (aspectos físicos), calidad corporativa (imagen del proveedor) y calidad interactiva (interrelaciones proveedor-usuarios).

La propuesta de Garvin (1988, citado por Díaz & Pons, 2009) es una de las más extensas con ocho dimensiones: desempeño, características, confiabilidad, apego, durabilidad, aspectos del servicio, estética y calidad percibida

El planteamiento realizado por Eiglier y Langeard en 1989, establece tres dimensiones: calidad en el proceso de fabricación del servicio (facilidades), calidad en el desarrollo del servicio (interacción con los clientes) y calidad del servicio prestado como resultado final (Díaz & Pons, 2009).

Senlle (1993, citado por Díaz & Pons, 2009) propone dos tipos de dimensiones: cuantitativa (tiempos de espera y entrega, facturación, cumplimiento) y cualitativa (estética, cortesía, amabilidad, simpatía, confianza, seguridad).

Díaz & Pons (2009) mencionan diferentes propuestas conceptuales que integran distintas dimensiones en un modelo utilizado para guiar la medición de la calidad percibida en servicios. Las propuestas mencionadas son las de: Grönroos (1988), Bitner (1990), Nguyen (1991), Bolton & Drew (1991) y Parasuraman et al. (1993).

El modelo de Bitner (1990) contrasta las expectativas hacia el servicio con el resultado percibido del servicio, produciendo un efecto sobre la experiencia de cliente, la cual se encuentra directamente relacionada con la calidad percibida en el servicio. Según la satisfacción o insatisfacción del usuario, se recomendará el servicio.

El modelo de Nguyen (1991) expone que la calidad percibida en el servicio es el resultado de la interacción de 5 componentes: la imagen de la empresa, el personal en contacto con el cliente, la organización de la servucción, el soporte físico de la servucción y la satisfacción del cliente. En este modelo todos los componentes se relacionan entre sí.

El modelo propuesto por Bolton y Drew (1991) propone que la percepción de calidad en el servicio proviene de la satisfacción del cliente, la cual responde a una diferencia entre las expectativas (afectadas por las características del cliente) y las percepciones (afectadas por las características del servicio) del usuario.

Es importante destacar que en los tres modelos anteriores, la calidad percibida en servicios se analiza a partir de las expectativas y las percepciones de los usuarios. Este parece ser el enfoque más utilizado para medir calidad en servicios.

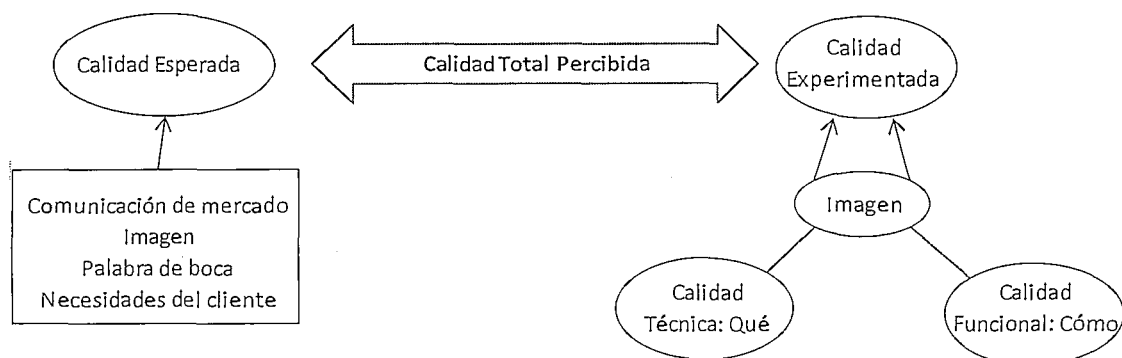
De la literatura revisada, se identifica que las propuestas más conocidas y discutidas en el tema de calidad en los servicios han sido clasificadas en dos grandes corrientes o enfoques: la escuela nórdica y la escuela americana (Duque, 2005; Serrano & López, 2007; Serrano, López & García, 2007).

La escuela nórdica propone que la calidad en servicios se debe analizar a partir de dos dimensiones: calidad técnica y calidad funcional (Serrano & López, 2007). El modelo de Grönroos es el más representativo de esta escuela. Dicho modelo analiza el efecto de la imagen del proveedor sobre la percepción de calidad del servicio (Grönroos, 1990).

Según el modelo de Grönroos el usuario experimenta dos tipos de calidad, la calidad técnica que hace referencia a lo que el usuario recibe y la calidad funcional que se refiere a cómo se recibe el servicio, es decir, el modo en que el servicio es entregado (Parasuraman, et. al, 1985; Duque, 2005; Díaz & Pons, 2009).

La figura 2.1 muestra el modelo de la Calidad Total Percibida de Grönroos, el cual indica que si las percepciones (calidad experimentada, procedente de la calidad técnica y la calidad funcional) satisfacen las expectativas (calidad esperada, afectada por la comunicación externa y las necesidades del usuario) el servicio es percibido por el usuario como de buena calidad (Grönroos, 1990). A este contraste se le denomina paradigma de la desconfirmación.

Figura 2.1
Modelo de la Calidad Total Percibida



Fuente: Grönroos, 1990, pág.41

Los trabajos de Parasuraman et al. (1985, 1988) son la base de la escuela norteamericana. Dichos trabajos se concentraron en obtener una definición del concepto calidad del servicio, en proponer un instrumento para medirlo (ServQual) y en plantear un modelo conceptual (Modelo Gaps) para explicar la diferencia entre las percepciones y las expectativas (Serrano & López, 2007). Para realizar su propuesta, Parasuraman, et. al, también partieron del paradigma de la desconfirmación.

Parasuraman et al. propusieron cinco dimensiones para medir calidad en los servicios: elementos tangibles, confiabilidad, responsabilidad, aseguramiento y empatía. Para Díaz & Pons (2009) el planteamiento de Parasuraman et al. es el más discutido, el más completo y el más sencillo de comprender. Según Duque (2005) es el planteamiento más utilizado por los académicos y el más mencionado en artículos sobre el tema.

En la literatura es posible identificar otras propuestas más recientes para definir y medir calidad en los servicios como los modelos propuestos por Brogowicz, Delene y Li, Corin y Taylor, Teas, Rust y Oliver y Candido y Morris (Duque, 2005).

Estos modelos proceden de los planteamientos realizados por Grönroos (escuela nórdica) y por Parasuraman et al. (escuela americana). En ellos, cada autor hace un esfuerzo por sintetizar los principales aspectos de cada escuela o corriente de pensamiento, corrigiendo sus debilidades y explotando sus fortalezas.

A criterio de Díaz & Pons (2009) "sólo el modelo de Parasuraman y otros (1993) aborda la calidad de servicio percibida con un enfoque total, explicado y materializado en sus dos partes en la producción de un servicio: los clientes y la empresa" (pág.13).

El planteamiento de Parasuraman et al. (1985, 1988) ha sido el más discutido y debatido en torno a la calidad en servicios. Este planteamiento se ha logrado mantener en el tiempo, siendo base de las propuestas más recientes. La escuela norteamericana, en relación a la escuela nórdica, ha recibido más adeptos, los cuales siguen la misma corriente de pensamiento expuesta por Parasuraman et al. en sus investigaciones.

Debido a lo anterior, se seleccionó el modelo de 5 dimensiones planteado por Parasuraman et al. (1985, 1988) como guía conceptual de la presente investigación.

2.2.3 Modelo Gaps

El modelo GAPS planteado por Parasuraman et al. (1985) se deriva de una investigación exploratoria desarrollada con el fin de definir el concepto de calidad aplicada a los servicios. El estudio consistió en aplicar entrevistas a ejecutivos de empresas prestadoras de servicios y realizar grupos focales con usuarios de distintos servicios.

En la investigación se indagó sobre el significado de calidad del servicio, las características de los servicios de calidad y los criterios utilizados para evaluar la calidad en los servicios (Parasuraman et al., 1988).

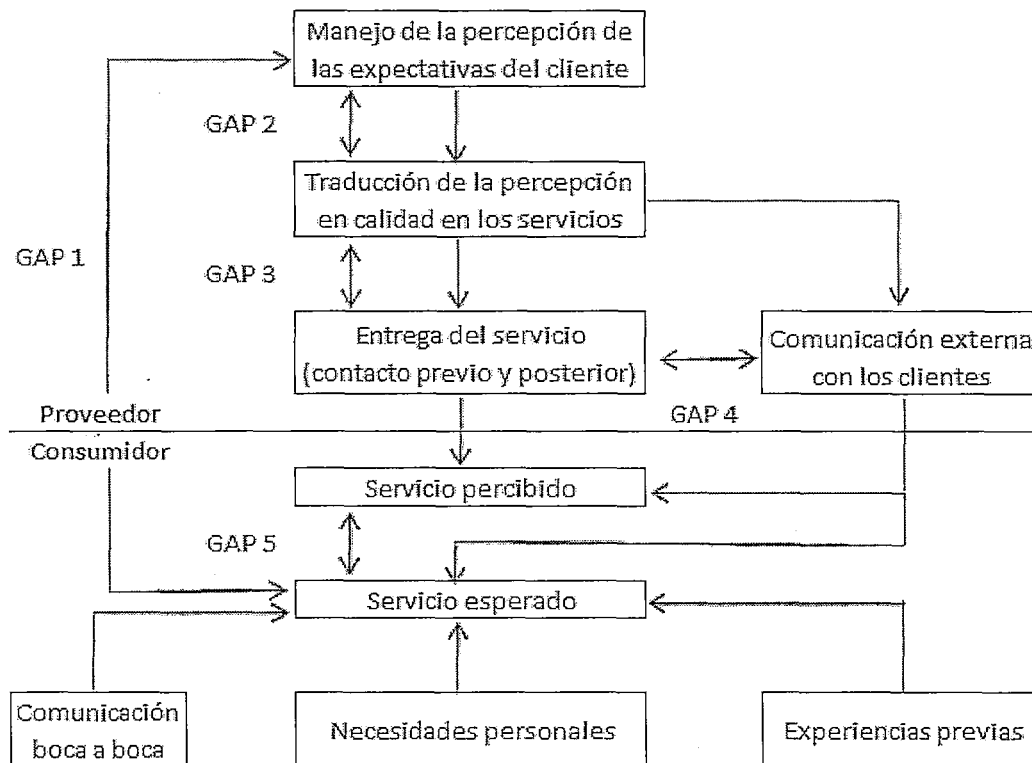
El producto más sobresaliente de la investigación fue la identificación de un conjunto de diferencias ("gaps" por su nombre en inglés) entre las percepciones que tenían los proveedores acerca de lo que es calidad en el servicio (partiendo de su interpretación de las necesidades de los clientes) y lo que los clientes esperaban recibir como servicio de calidad (partiendo de su percepción de calidad en el servicio y de sus expectativas).

Estas diferencias o "gaps" permitieron establecer un modelo conceptual sobre la calidad en los servicios, expuesto en la figura 2.2. Según Parasuraman et al. (1985), los "gaps" o diferencias del modelo conceptual se interpretan de la siguiente forma:

Gap1-Diferencia entre las percepciones de los proveedores sobre lo que los clientes esperan en un servicio de calidad y las expectativas de los clientes: los ejecutivos no siempre logran identificar, previamente, los elementos considerados de alta calidad por los clientes, los elementos que el servicio debe tener para cumplir con sus necesidades y los niveles de esos elementos, durante la prestación del servicio.

Gap2-Diferencia entre la percepción que tienen los proveedores de las expectativas del cliente y las especificaciones de calidad en los servicios: debido a varios factores como falta de recursos, condiciones de mercado e indiferencia de los administradores, se dificulta que la prestación del servicio coincida con los que los clientes esperan recibir.

Figura 2.2
Modelo Gaps



Fuente: Parasuraman, Zeithaml & Berry, 1985, pág.44

Gap3-Diferencia entre las especificaciones de calidad en el servicio y la prestación del mismo: los empleados que brindan el servicio producen una gran influencia sobre la evaluación de la calidad percibida por los clientes. Aunque existan lineamientos que guían su actuación, ésta no puede ser siempre estandarizada, es decir, es variable.

Gap4-Diferencia entre la prestación del servicio y la comunicación externa: los medios de comunicación y de publicidad utilizados por los proveedores pueden afectar las expectativas y percepciones de los clientes sobre el servicio a recibir. En su publicidad, los proveedores no deben prometer más de lo que realmente pueden brindar, así como tampoco deben omitir información sobre el servicio.

Gap5-Diferencia entre las expectativas de los usuarios y la percepción del servicio recibido: la calidad percibida por un cliente, al recibir un servicio, está en función de la magnitud y dirección de la diferencia entre lo que espera recibir del servicio y lo que percibe que recibió.

La propuesta de Parasuraman et al. (1985) consiste en que las discrepancias entre las expectativas y percepciones de los usuarios, que determinan si el servicio recibido es de calidad o no, se ven afectadas por las otras diferencias indicadas en el modelo. Específicamente, se propone que el gap 5 está en función de los demás "gaps".

Las diferencias o "gaps" permiten enfocarse en puntos clave del proceso de prestación del servicio e identificar debilidades y fortalezas en dicho proceso. El análisis de los "gaps" permite proponer mejoras al servicio para ser considerado de calidad por los clientes. El punto principal es que, la magnitud y dirección de cada "gap" tiene un impacto directo en la calidad del servicio percibida.

La conclusión de Parasuraman et al. (1985) es que la calidad percibida en los servicios depende de la naturaleza de la discrepancia entre el servicio esperado (S.E.) y el servicio percibido (S.P.), de la siguiente forma:

- si $S.E. > S.P. \Rightarrow$ la calidad percibida no es satisfactoria.
- si $S.E. = S.P. \Rightarrow$ la calidad percibida es satisfactoria.
- si $S.E. < S.P. \Rightarrow$ la calidad percibida es más que satisfactoria.

Los resultados de esta investigación revelaron, además, que los criterios utilizados por los usuarios de los servicios, al evaluar la calidad, pueden ser clasificados en 10 dimensiones, denominadas por los autores "determinantes de calidad en el servicio" (Parasuraman et al., 1985, pág.46). Estas dimensiones son:

- **Confiable:** involucra consistencia en la actuación y confianza. Significa que el proveedor brinda el servicio bien desde la primera vez.
- **Responsabilidad:** concerniente al compromiso de los empleados al momento de brindar el servicio. Involucra los tiempos en las etapas del servicio.

- Competencia: significa poseer las habilidades adecuadas y el conocimiento necesario para proveer el servicio.
- Acceso: involucra accesibilidad y facilidad de contacto. El servicio puede ser accedido por varios medios y de forma conveniente.
- Cortesía: incluye amabilidad, respeto, consideración y simpatía del personal de contacto (para con los clientes).
- Comunicación: significa mantener a los clientes informados en un lenguaje fácil de entender y escuchar lo que ellos opinan.
- Credibilidad: concierne a integridad, honestidad, ser fidedigno. Significa tener presente los intereses de los clientes.
- Seguridad: es estar libres de peligro, riesgo o duda. Ser de fiar. Involucra confidencialidad.
- Entendimiento / conocimiento del usuario: significa hacer un esfuerzo para entender las necesidades del cliente, conocer sus requerimientos.
- Tangibles: incluye la evidencia física del servicio. Todo aquello que puede ser percibido físicamente.

2.2.4 SERVQUAL: Service Quality (Calidad del Servicio)

Los investigadores Parasuraman et al. (1988) realizaron una segunda investigación con el objetivo de desarrollar una escala estandarizada para medir las expectativas y percepciones de los usuarios sobre calidad en los servicios utilizados. El estudio permitió proponer un método para medir el "gap 5" del modelo conceptual propuesto en la primera investigación (Modelo GAPS de la figura 2.2).

La investigación incluyó dos etapas de análisis: en la primera etapa se seleccionaron los ítems con mejor capacidad para discriminar entre usuarios con diferentes percepciones de calidad, además se analizó la dimensionalidad de la escala. La segunda etapa consistió en comprobar la confiabilidad y la validez de la escala por medio de análisis de factores exploratorios y alfa de Cronbach (Parasuraman et al., 1988).

El estudio se inició con un instrumento de 97 ítems que pretendía medir expectativas y percepciones de 200 usuarios de cinco tipos de servicios diferentes, incluyendo aspectos de las 10 dimensiones de calidad del servicio.

Posterior a un análisis iterativo de confiabilidad y validez (análisis de factores exploratorios y alfa de Cronbach), la escala se redujo a 34 ítems y las dimensiones se redujeron a 7. En una segunda etapa, la escala reducida se puso a prueba en cuatro muestras de 200 usuarios de cuatro servicios distintos.

Según Parasuraman et al. (1988), los resultados de la segunda etapa fueron "consistentes". El análisis de confiabilidad recomendó la eliminación de 12 ítems. De forma análoga, los análisis de factores recomendaron la unificación de dos pares de dimensiones.

El resultado final, después de los análisis, fue una escala depurada de 22 ítems distribuidos en 5 dimensiones. Las nuevas dimensiones de calidad resultantes fueron (Parasuraman et al., 1988):

- Tangibles: "facilidades físicas, materiales, equipo y apariencia del personal". La apariencia física indica interés de los proveedores de servicio para brindar el mejor ambiente.
- Confiabilidad: "habilidad para brindar el servicio prometido de forma fiable y precisa". Indica que el servicio es brindado a tiempo todo el tiempo, de la misma forma y sin errores.
- Responsabilidad: "voluntad de ayudar a los usuarios y proporcionar un servicio rápido". Se refiere a no dejar a los usuarios esperando sin alguna razón aparente.
- Aseguramiento: "conocimiento y cortesía de los empleados y su habilidad para inspirar seguridad y confianza". Hace referencia a la competencia para brindar el servicio, la amabilidad, respeto y simpatía para con los usuarios y una comunicación efectiva.
- Empatía: "cuidado y atención individualizada brindada a los usuarios". Incluye aspectos de accesibilidad al servicio y el esfuerzo por entender las necesidades de los usuarios.

Según Parasuraman et al. (1988) "ServQual es una escala concisa de ítems múltiples con buena confiabilidad y validez que los proveedores (detallistas) pueden utilizar para comprender mejor las expectativas y percepciones de los clientes sobre el servicio" (pág.30).

En resumen, el modelo "GAPS" corresponde a un modelo conceptual que permite obtener una definición del concepto calidad percibida en los servicios y la escala "SERVQUAL" a una escala multidimensional creada para medir dicho concepto.

2.3 Indicadores de calidad

El monitoreo realizado por la Administración del Gobierno sobre la calidad percibida por los usuarios de servicios públicos, como parte de un proceso continuo de autoevaluación y mejora, usualmente termina con el establecimiento de indicadores de calidad (Sancho, 1991). Los indicadores de calidad permiten cuantificar la calidad percibida por los usuarios según las dimensiones involucradas en la prestación del servicio (Mendoza, 1993, citado por Sancho, 1991).

A continuación se exponen las características indicadas por Alvira (1991, citado por Sancho, 1991) que deben poseer los indicadores para el monitoreo y control de calidad de los servicios públicos:

- La adecuación del sistema de indicadores a la finalidad del análisis: los indicadores tienen que estar directamente relacionados con el objetivo de la prestación de servicios públicos, a saber: la satisfacción de necesidades ciudadanas. Los indicadores deben estar diseñados conceptualmente, de manera que puedan medir diferentes aspectos como: calidad, eficiencia, eficacia, economía o alguna otra dimensión relacionada con la prestación del servicio.
- La referencia del sistema de indicadores a un período temporal: los resultados reflejados por los indicadores deben referirse a un período de tiempo determinado. Este aspecto es importante para mantener actualizado el sistema de indicadores.
- La claridad y la mensurabilidad de los indicadores: los indicadores deben estar definidos de forma clara y simple, con el fin de que se facilite su lectura y comprensión. El método de medición de los mismos también debe ser sencillo.
- La fiabilidad y la oportunidad de los indicadores: los indicadores deben ser válidos y confiables de acuerdo a los análisis correspondientes. Por otro lado, la

información que aporten debe de estar en función de las necesidades dentro del proceso de toma de decisiones.

- La posibilidad de realizar comparación entre los indicadores: los valores de los indicadores deben ser comparables en tiempo y espacio. Las comparaciones han de poder realizarse entre indicadores en momentos diferentes así como referidos a lugares de prestación distintos.

Los indicadores de calidad dentro del sector público, se pueden clasificar de diversas formas de acuerdo con los objetivos. Según Sancho (1991) los criterios más comunes los clasifican en función del formato del indicador y en función de su contenido.

2.3.1 Indicadores en función del formato

2.3.1.1. Por el modo de expresar su valor

- a) Indicadores cuantitativos: aquellos cuyo valor se expresa en términos numéricos.
- b) Indicadores cualitativos: aquellos que no son expresados en términos numéricos. Generalmente responden a una escala ordinal.

2.3.1.2. Por el grado de complejidad

- a) Indicadores simples o primarios: aquellos de carácter cuantitativo, cuyos valores son absolutos y se obtienen de manera directa a partir de la observación o medición de la característica o fenómeno.
- b) Indicadores compuestos o derivados: aquellos que se obtienen a partir de dos o más indicadores simples. Representan relaciones (usualmente cocientes) entre dos o más indicadores o variables.

2.3.2 Indicadores en función del contenido

2.3.2.1. Evaluación externa

- a) Indicadores de impacto: miden el cambio que se quiere producir externamente mediante la realización de actuaciones. Estos indicadores miden el resultado conseguido con la implementación de los servicios, es decir, la eficacia de la actuación pública.
- b) Indicadores de percepción: miden cómo la ciudadanía percibe y valora la actuación del sector público. Usualmente se generan a partir de la aplicación de encuestas de opinión sobre el servicio público. Diferencian las valoraciones entre usuarios y no usuarios de los servicios.
- c) Indicadores de entorno: miden las condiciones referentes a los elementos externos a la organización que tienen vínculo con ella: sus actuaciones y destinatarios, actuaciones de la misma organización y con la ciudadanía.

2.3.2.2. Evaluación interna

- a) Indicadores de volumen de actuación: miden los resultados de las organizaciones públicas según el volumen de servicios ofrecidos y de objetivos cumplidos.
- b) Indicadores de recursos: miden los recursos empleados en la realización de las actividades públicas. Se subdividen en indicadores de economía e indicadores de eficiencia.
 - b1) Indicadores de economía: son el resultado de comparar los costos de los recursos de la Administración invertidos en las actuaciones públicas y los costos de estos mismos recursos a precios de mercado.

- b2) Indicadores de eficiencia: razones que resultan al comparar las actuaciones públicas realizadas y el volumen y/o el costo de los recursos empleados para llevarlas a cabo.

Para efectos del presente trabajo, siguiendo la clasificación de Sancho (1991), el índice propuesto para medir calidad en servicios públicos, corresponde a un índice global de percepción, cuantitativo y derivado, compuesto de varios subíndices también cuantitativos derivados.

El índice de calidad percibida aplicado a servicios públicos, se propuso con el fin de evaluar la calidad de los servicios públicos, desde el punto de vista externo de la organización, mediante el examen de las percepciones de los usuarios.

III Antecedentes

En nuestro país se han propuesto distintas iniciativas que promueven la mejora de los servicios públicos, mediante el monitoreo de las instituciones que los brindan y la aplicación de estudios de medición de la calidad de dichos servicios, así como el seguimiento a las medidas correctivas propuestas, con el fin de optimizar la prestación de servicios la ciudadanía.

En este capítulo se pone de manifiesto el interés de las Administraciones del Gobierno de Costa Rica por darle un impulso al monitoreo y mejoramiento de los servicios públicos. Adicionalmente, se mencionan los instrumentos actualmente utilizados para medir y analizar la calidad en los servicios públicos en Costa Rica.

3.1 Interés en la calidad en los servicios públicos nacionales

Los servicios públicos son prestados por el Estado para la satisfacción de necesidades sociales, en respuesta a demandas ciudadanas. Debido a esta condición, las Administraciones del Gobierno se han preocupado por monitorear los servicios públicos, a fin mejorar su prestación. Es posible identificar instrumentos de planificación y directrices gubernamentales con metas y objetivos de mejoramiento de servicios públicos.

El PND 2006-2010 de la Administración Arias Sánchez, incluye una meta para “agilizar, simplificar y elevar la calidad de servicios públicos críticos (...)” (MIDEPLAN, 2007b, pág. 92). En dicho documento se impulsa la realización de mejoras en áreas críticas de la prestación de servicios públicos para alcanzar el objetivo de incrementar la calidad de este tipo de servicios.

Según se indica en el PND 2011-2014 (MIDEPLAN, 2010b) que corresponde a la Administración Chinchilla Miranda, las instituciones del Estado tienen el desafío de mejorar la calidad y ampliar la cobertura de los servicios prestados a la ciudadanía. Así se establece en las metas del apartado de Modernización del Estado, incluidas en el PND 2011-2014.

Otras propuestas dirigidas al mejoramiento de los servicios públicos se mencionan en los desafíos planteados por la CEARE, incluidos dentro de su Propuesta de Acciones de Reforma del Estado. Dos de estos desafíos se refieren a reducir el tiempo de respuesta del Estado y mejorar la calidad de los servicios públicos, además de mejorar la capacidad de los funcionarios públicos para proveer servicios de calidad.

Con el objetivo de cumplir con los desafíos planteados por la CEARE, se propusieron acciones concretas para el mejoramiento de la prestación de servicios públicos, entre las que sobresalen el establecimiento de un programa de mejora basado en resultados del monitoreo de los servicios públicos y realizar periódicamente estudios de percepción entre usuarios de servicios públicos (CEARE, 2008).

La Comisión de Eficiencia Administrativa (CEA), en su Plan de Trabajo para 2011, propuso el objetivo de “promover el mejoramiento continuo en el acceso y la calidad del servicio público” (CEA, 2010, pág.8). Como parte de las acciones, se propuso identificar instrumentos para medir el impacto de los servicios públicos en los usuarios

A las anteriores iniciativas, se suma la creación del SNCS con la responsabilidad de promover la mejora continua en los servicios públicos, participando a los usuarios en dicha labor (Presidencia y MIDEPLAN, 2009).

En el plan 2010-2014 del PME se propuso implementar la mejora continua en la prestación de servicios públicos, con el fin de promover la calidad en dichos servicios (MIDEPLAN, 2010c).

Con el establecimiento del SINE, la Administración del Gobierno puso de manifiesto la necesidad de promover la eficiencia en la prestación de servicios públicos, con el fin de mejorar la gestión pública de forma continua (Presidencia y MIDEPLAN, 2010^a).

Las anteriores referencias, ponen en evidencia el interés que posee el Estado en monitorear la prestación los servicios públicos, con propuestas dirigidas a medir, evaluar, mejorar e impulsar la calidad en dichos servicios. En todas ellas, se demuestra la necesidad de contar con mediciones válidas y confiables sobre calidad en los servicios prestados.

3.2 Monitoreo de calidad en los servicios públicos nacionales

Con el fin de darle seguimiento a la prestación de servicios públicos y realizar una evaluación de la gestión de las instituciones que brindan dichos servicios, se han desarrollado varios instrumentos con el objetivo de medir el trabajo de las instituciones públicas que brindan servicios así como la calidad de dichos servicios.

Entre estos instrumentos, los más representativos son el diagnóstico de situación de del SNCS sobre las contralorías de servicios institucionales y los índices de Gestión Institucional y Gestión para Resultados en el Desarrollo. En este apartado se realiza una descripción general y se mencionan las características básicas de estos instrumentos.

3.2.1 Diagnóstico de la Secretaría Técnica del SNCS

El SNCS se creó bajo el principio que los usuarios tienen el derecho de recibir servicios de óptima calidad (Presidencia y MIDEPLAN, 2009). La función primordial del SNCS es impulsar el mejoramiento continuo en la prestación de servicios públicos

La evaluación que realiza la Secretaría Técnica del SNCS sobre los servicios públicos en Costa Rica, consiste en un diagnóstico de la situación de los servicios que brindan las instituciones del Estado. Este diagnóstico se basa en un informe anual que las contralorías de servicios institucionales le remiten a la Secretaría Técnica, sobre el estado de dichas contralorías y de los servicios brindados.

El principal objetivo de los informes anuales es el de contar con información continua sobre los resultados de la gestión realizada por las contralorías de servicios de las distintas instituciones que prestan servicios a la ciudadanía (MIDEPLAN, 2009). Los informes anuales se basan en las inconformidades, sugerencias, solicitudes y denuncias que realizan los usuarios de los servicios, las cuales se clasifican en cinco categorías (MIDEPLAN, 2011c):

- Información: accesibilidad a instalaciones y calidad de la información brindada.
- Trato a los usuarios: amabilidad, cortesía, educación, interés, equidad, ayuda y conocimiento de los funcionarios.

- Calidad del servicio: simplicidad del procedimiento, tiempos de espera, petición de documentación, recurso humano, horario y cumplimiento de expectativas.
- Instalaciones: accesibilidad, barreras físicas, condiciones ambientales y físicas, recursos tecnológicos, mobiliario y espacio.
- Otras: cumplimiento de labores y de horarios, discriminación, uso de los recursos

El diagnóstico realizado por la Secretaría Técnica del SNCS se considera un instrumento de retroalimentación utilizado para impulsar la mejora continua en los servicios, a partir de las inconformidades presentadas por los usuarios (MIDEPLAN, 2009). Una limitación de este diagnóstico es que no considera la opinión de los usuarios que, por distintos motivos, no presentan inconformidades ante las contralorías de servicios institucionales.

Se debe considerar que la interpretación y análisis de las inconformidades de los usuarios, es realizado por el Contralor(a) de Servicios, de forma subjetiva y desde el punto de vista interno de la gestión y dinámica de la institución a la que pertenece.

Otro aspecto que limita el alcance del análisis del SNCS corresponde a que la información indicada en el informe se centra en las condiciones y limitaciones de las contralorías y no tanto en la evaluación de los servicios propiamente. En relación a los servicios, el informe se limita a indicar la distribución porcentual de las inconformidades presentadas por los usuarios, según la clasificación vista anteriormente.

La Secretaría Técnica del SNCS ha resaltado en sus diagnósticos, una serie de limitaciones que poseen los informes anuales de las contralorías de servicios institucionales, lo que compromete la información presentada. Entre las limitaciones se mencionan (MIDEPLAN, 2011b):

- Información incompleta en cuanto las inconformidades con la prestación de los servicios, no se clasifican en las categorías propuestas para dicho fin, lo que dificulta su análisis.
- Los informes no contienen datos estadísticos confiables, se presentan pocos datos, algunos de ellos incompletos o con errores. Esta situación afecta la validez del análisis y las conclusiones de los diagnósticos del SNCS.

- La mayor limitación es la no presentación del informe anual por parte de las contralorías de servicios de las distintas instituciones del Estado, impidiendo dar seguimiento a la labor de prestación de servicios públicos.

El diagnóstico realizado por la Secretaría Técnica del SNCS corresponde a un esfuerzo por evaluar la gestión de las instituciones públicas en la prestación de servicios a la sociedad; sin embargo, no logra realizar una medición estandarizada, objetiva y numérica que permita cuantificar el nivel de calidad de los servicios brindados.

3.2.2 Índice de Gestión Institucional (IGI)

De acuerdo con su labor fiscalizadora, la Contraloría General de la República (CGR) desarrolló, en 2010, el proyecto denominado "Índice de Gestión Institucional" con el objetivo establecer un indicador que permitiera brindar una visión del fortalecimiento de la gestión de un conjunto de instituciones públicas que son de su interés (DFOE, 2011b).

La aplicación del índice consiste en el envío de un cuestionario digital a las instituciones públicas incluidas en el estudio, destinado a medir cinco factores de la gestión institucional: planificación, gestión financiera y presupuestaria, contratación administrativa, control interno y servicio al usuario (DFOE, 2011b).

El factor "servicio al usuario" que evalúa el tema de servicios públicos, se compone de tan solo 10 preguntas (de 60 a 69) dentro del cuestionario, aplicado a partir de 2011. Por tanto es poco lo aporta sobre el análisis de calidad en servicios.

Las preguntas indagan si la institución cuenta con: mecanismos para admitir inquietudes, contraloría de servicios y políticas sobre el tiempo de atención, si cumple con la Ley 7600 y la Ley 8220 y si se evalúa la satisfacción de los usuarios. La mayor atención se centra en si se dispone o no de una página web, el tipo de información contenida en ella y la posibilidad de realizar trámites en línea (DFOE, 2011b).

El IGI incluye algunos aspectos que tienen influencia en la percepción de la calidad de los servicios, pero éstos no se miden de forma directa; la evaluación se limita a determinar si se dispone o no con cada aspecto.

El objetivo del IGI no es medir calidad en los servicios, sino la gestión institucional de forma general. El IGI evalúa si la institución dispone de elementos clave que conforman los cinco factores de la gestión institucional, pero no si son aplicados o en qué medida se cumple con ellos. El IGI es en una herramienta de recolección de datos sobre las medidas adoptadas por las instituciones para fortalecer su gestión (DFOE, 2012a).

3.2.3 Índice GpRD

El Índice GpRD corresponde a un indicador que forma parte del Sistema de Evaluación PRODEV⁴ (SEP) creado en 2005 por Banco Interamericano de Desarrollo (BID) con el propósito de evaluar el desarrollo y aplicación de las prácticas e instrumentos de gestión para resultados en el desarrollo (GpRD) en el sector público (García, 2011).

La evaluación de la capacidad de GpRD de un país se realiza en cinco pilares: Planificación para resultados, Presupuesto por resultados, Gestión financiera, auditoría y adquisiciones, Gestión de programas y proyectos y Monitoreo y evaluación, los cuales agrupan los elementos del ciclo de gestión.

Los cinco pilares se descomponen en tres niveles: 16 componentes, 37 indicadores y 141 variables. Las variables, llamadas requisitos mínimos, son medidas en una escala de 0 a 5, donde 0 ("inexistente") indica que el requisito no se cumple y 5 ("consolidado") significa que el requisito se encuentra operando óptimamente. El Índice de GpRD consiste en un promedio de los 37 indicadores de desempeño, utilizados para monitorear los planes de gobierno (García & García, 2010).

Por medio del Índice de GpRD se realiza un análisis a nivel país de la capacidad para implementar la GpRD y determinar cuáles países miembros requieren apoyo por parte del BID para lograrlo. Desde su concepción, el Índice de GpRD requiere asociarse a estrategias de mejoramiento continuo de la calidad de los bienes y servicios brindados por el Estado (García & García, 2010).

⁴ Programa de Implementación del Pilar Externo del Plan de Acción a Mediano Plazo para la Efectividad en el Desarrollo.

El componente "Gestión de la producción de bienes y servicios" incluido en el pilar "Gestión de programas y proyectos" incorpora la evaluación de servicios públicos. Esta evaluación se realiza por medio de 3 indicadores: visión sectorial a mediano plazo, GpRD en la producción de bienes y servicios y sistemas sectoriales de información.

La evaluación se limita a determinar si se en el sector público se establecen contratos de gestión, convenios de administración por resultados o esquemas de remuneración de personal ligados a la obtención de resultados institucionales, los cuales se considera, ayudan a mejorar la calidad y cobertura de los bienes y servicios públicos.

3.4 Estudio: "Percepción en la Calidad de los Servicios Públicos"

Las Administraciones de Gobierno han realizado diversas propuestas que evidencian el interés por medir, evaluar y monitorear la calidad de los servicios prestados por las entidades públicas, en respuesta a las demandas ciudadanas.

Como parte de las iniciativas impulsadas por el Gobierno, MIDEPLAN en conjunto con la Universidad de Costa Rica (UCR), desarrollaron un estudio que permitió capturar la opinión de los usuarios de cinco servicios públicos sobre la calidad percibida del servicio recibido y el cambio en dichas percepciones al cabo de un año.

El estudio denominado "Percepción en la Calidad de los Servicios Públicos" tuvo como objetivo determinar la percepción de los usuarios de algunos servicios públicos, catalogados como estratégicos, sobre la calidad con que son brindados (Madrigal Pana, 2009). El fin fue apoyar a las instituciones del Estado que se encargan de realizar la evaluación de los servicios públicos.

El estudio consistió en aplicar dos rondas de cinco encuestas dirigidas a evaluar los servicios de cinco instituciones públicas. En cada ronda, se aplicó una encuesta para cada servicio evaluado. Las encuestas fueron aplicadas con un año de diferencia, en los meses de noviembre a marzo entre 2008-2009 y 2009-2010. Se utilizaron muestras compuestas por distintas personas en cada aplicación.

Las instituciones evaluadas se escogieron siguiendo los lineamientos expuestos en el PND 2006-2010, bajo la consideración que estas instituciones ofrecen servicios sensibles a la sociedad (Madrigal Pana, 2009). Las instituciones consideradas en el estudio así como los servicios evaluados se detallan a continuación:

- Consultas y certificaciones de la Plataforma de Servicios del Registro Nacional (RN).
- Emisión de licencias de conducir por primera vez del Ministerio de Obras Públicas y Transportes (MOPT).
- Emisión y renovación de pasaportes de la Dirección General de Migración y Extranjería (DGME).
- Reclutamiento y selección de personal de la Dirección General de Servicio Civil (DGSC).
- Aprobación de estudios de impacto ambiental de la Secretaría Técnica Nacional Ambiental (SETENA).

En la emisión de licencias de conducir, se evaluó todo el proceso para adquirir una licencia, considerando el examen teórico, el examen práctico y la entrega de la licencia. De esta forma se obtuvieron tres calificaciones, una para cada etapa del proceso, las cuales se promediaron para obtener una calificación global del servicio.

La población de estudio, en todas las encuestas, se refirió a las personas de 18 o más años que, al momento de la evaluación, habían utilizado los servicios al menos una vez. Las inferencias del estudio se refieren a la totalidad de usuarios de cada servicio a nivel nacional.

Las cinco muestras de cada ronda estuvieron compuestas por personas físicas y jurídicas de 18 años y más que habían demandado algún trámite o utilizado alguno de los servicios anteriormente citados, en al menos una ocasión.

Se aplicaron entrevistas personales (cara a cara) para evaluar los servicios del Registro Nacional y entrevistas telefónicas para evaluar los servicios de las demás instituciones. Se aplicó un proceso de selección probabilístico, donde todos los usuarios tenían la misma probabilidad de ser incluidos en la muestra (Madrigal, 2009).

Para aplicar las entrevistas telefónicas, los individuos fueron seleccionados de forma sistemática a partir de un listado otorgado por las instituciones respectivas, extraído de los registros que dichas instituciones guardan sobre los usuarios atendidos (Madrigal, 2009). Estos listados conformaron los marcos muestrales.

Las entrevistas personales se aplicaron en el sitio, seleccionando a los individuos después de haber recibido el servicio. La unidad de selección estuvo compuesta por el día, hora y ventanilla en que fueron atendidos los usuarios, aplicando tres entrevistas por hora. Se utilizó un diseño de probabilidades proporcionales al tamaño, asignando la misma probabilidad de selección a todos los usuarios, considerando la proporción de personas que utilizan el servicio en cada horario y ventanilla (Madrigal, 2009).

En cada ronda de cinco encuestas, se propuso un tamaño de muestra de 100 usuarios para evaluar el servicio de la SETENA y de 300 usuarios para evaluar los servicios prestados por las demás instituciones.

Los tamaños se propusieron asumiendo un muestreo simple al azar con una varianza máxima para variables categóricas ($p = 0,5$), una confianza de 95% y un error máximo permisible de alrededor de 5%. A los tamaños de muestra definidos se les sumó un 20% correspondiente a no respuesta (rechazos y no localizables), asumido por el investigador. Los tamaños de muestra finales se exponen en el Cuadro 3.1.

**Cuadro 3.1. Tamaños de muestra finales.
Estudio Percepción de la Calidad
en Servicios Públicos. 2008-2010.**

Institución	2008-2009	2009-2010
RN	348	360
MOPT	300	308
DGME	306	300
DGSC	302	340
SETENA	103	100
Total	1.359	1.408

Fuente: Madrigal, 2009-2010.

Se utilizaron cuestionarios independientes para evaluar cada servicio, derivados de un cuestionario maestro que incluía preguntas sobre aspectos comunes a todos los

servicios. Se aplicaron cinco cuestionarios diferenciados entre sí por el orden de las preguntas y aspectos específicos de los servicios consultados, con un total de 133 preguntas distribuidas en 24 apartados, además de las preguntas socio-demográficas.

Los cuestionarios mantuvieron una estructura similar guiada por la estrategia conceptual propuesta por el investigador, la cual define calidad en servicios en términos de necesidades satisfechas. Según el modelo conceptual utilizado, los servicios públicos se deben ajustar a los requerimientos de los usuarios con el objetivo de satisfacer sus necesidades (Madrigal, 2009).

En el estudio se analizó la calidad percibida en servicios públicos desde tres ámbitos: satisfacción de las necesidades de los usuarios, calidad de la atención y duración del servicio (Madrigal, 2009).

En el estudio se aplicó un análisis descriptivo de los ámbitos evaluados, estableciendo un punto de partida sobre la percepción de los usuarios en los servicios públicos brindados por las instituciones estudiadas. El análisis permitió dar a conocer aspectos que las instituciones deberían corregir con el fin de mejorar la percepción de los usuarios.

El análisis descriptivo realizado no reprodujo alguna medida estandarizada que agrupara en un solo valor numérico todos los aspectos evaluados, de tal forma que permita resumir la información de las encuestas aplicadas, sintetizar los resultados, realizar comparaciones entre las instituciones y facilitar la toma de decisiones.

De acuerdo con MIDEPLAN, el estudio de la UCR fue el primer paso para desarrollar un sistema de monitoreo de los servicios públicos a ofrecer a las instituciones del Estado como un mecanismo para alcanzar la excelencia en la gestión pública (Madrigal, 2010). Sin embargo el sistema de monitoreo no se desarrolló.

IV Marco Metodológico

Para el desarrollo de la presente investigación se estimaron y se pusieron a prueba dos modelos alternativos para el cálculo de un índice que permitiera medir la calidad percibida por los usuarios de los servicios públicos. Para el ajuste de los modelos propuestos, se requirió recolectar datos y aplicar técnicas de análisis multivariantes, los cuales se detallan en el presente capítulo.

En este apartado se explica de dónde se extrajeron los datos utilizados en el presente estudio, se detallan los modelos propuestos y sus componentes, las pruebas aplicadas a cada modelo y su evaluación, así como los análisis aplicados a los servicios evaluados.

4.1 Recolección de datos

Para el desarrollo de la presente investigación se utilizaron los datos recolectados en el estudio "Percepción en la Calidad de los Servicios Públicos" ejecutado por la Escuela de Estadística. No se recolectaron datos adicionales a los del estudio antes mencionado.

Los análisis elaborados en el presente trabajo se llevaron a cabo con los archivos de datos confeccionados a partir de las encuestas del estudio "Percepción en la Calidad de los Servicios Públicos". El primer archivo, correspondiente a la aplicación del 2008-2009, se utilizó para realizar análisis exploratorios y ajustar el modelo para generar el índice propuesto (muestra de ajuste y calibración).

El segundo archivo de datos (aplicación 2009-2010) se utilizó para realizar estudios confirmatorios sobre los modelos planteados, como un medio para realizar la validación interna del modelo, del índice y sus resultados (muestra de prueba y validación). Se calcularon dos índices con cada modelo uno para la evaluación de 2008 y otro para la evaluación de 2009.

4.2 Instrumentos de recolección

Los instrumentos de recolección de datos corresponden a los cuestionarios aplicados en el estudio realizado por la Escuela de Estadística.

Se utilizó un cuestionario independiente para evaluar cada uno de los servicios. Cada cuestionario es una derivación del cuestionario maestro que incluía preguntas sobre aspectos evaluados en todos los servicios y preguntas específicas de cada servicio e institución que lo brinda.

La estructura de los cuestionarios proviene del esquema conceptual utilizado, el cual mide la calidad percibida en servicios públicos desde tres ámbitos: satisfacción de las necesidades de los usuarios, calidad de la atención y duración del servicio (Madrigal, 2009).

4.3 Población

La población de estudio correspondió a personas de 18 o más años, de todo el territorio nacional, que al momento de la entrevista, habían utilizado los servicios evaluados al menos una vez. En las cuatro encuestas telefónicas, aplicadas en cada ronda, la población estuvo limitada a los usuarios de los servicios que poseían una línea telefónica (residencial y/o celular).

4.4 Muestra

Las muestra total del presente estudio correspondió a personas de 18 años o más que habían utilizado, en cualquiera de las oficinas disponibles, los servicios de las instituciones incluidas en el estudio de la cual se extrajeron los datos, en al menos una ocasión, y que se encontraban incluidas en los registros de dichas instituciones.

Los registros institucionales correspondieron al marco muestral de cada encuesta realizada en forma telefónica. Las entrevistas personales se basaron en un marco muestral compuesto por el día, hora y ventanilla en que los usuarios fueron atendidos. En todos los casos se utilizaron muestras probabilísticas con igual probabilidad de selección.

Dada la población de estudio y partiendo de que se seleccionaron muestras estadísticamente fidedignas, las inferencias de la presente investigación, abarcan la totalidad de personas, a nivel nacional, que hacen uso de los servicios otorgados por las cinco instituciones evaluadas.

4.5 Conceptualización del índice de calidad percibida

La conceptualización del índice propuesto para medir la calidad percibida en servicios públicos se representa en la figura 4.1. El esquema conceptual propuesto procede del modelo "Gaps" y de la escala "SevQual" propuestos por Parasuraman et al. (1985-1988).

El esquema conceptual de la figura 4.1 plantea que la calidad percibida se compone de cinco dimensiones, a diferencia del esquema conceptual utilizado en el estudio "Percepción en la Calidad de los Servicios Públicos", que midió la calidad en servicios a partir de tres dimensiones.

Figura 4.1 Esquema conceptual de calidad percibida.



Fuente: Elaboración propia

Fuente: Parasuraman et ál., 1988, pág.12

El factor central o factor común de la figura 4.1 corresponde al constructo *calidad percibida*, el cual se ve afectado o influenciado por cinco factores externos o subyacentes, que corresponden a las dimensiones de la calidad percibida. Cada dimensión se encuentra compuesta por una serie de elementos, denominados "descriptores", que la describen y la definen, estableciendo límites entre ellas, permitiendo diferenciarlas.

Las preguntas seleccionadas de los cuestionarios del estudio realizado por la Escuela de Estadística se consideraron una medición de los descriptores de las dimensiones y se utilizaron como las variables observadas de los modelos ajustados (anexo A).

Las variables observadas se agregaron para obtener un puntaje para cada dimensión (subíndice) y los puntajes de las dimensiones se agregaron para obtener un puntaje general del factor central (índice calculado).

De acuerdo con Parasuraman et al. (1988), no hay una correlación entre pares de dimensiones; es decir, una dimensión no afecta a la otra. Por tanto, se asumió que la relación entre las dimensiones se debe al efecto que el factor central ejerce sobre ellas (relación indirecta).

4.6 Variables de estudio

Para seleccionar las variables de estudio se realizó un análisis de las preguntas incluidas en los diferentes cuestionarios, con el fin de determinar y seleccionar aquellas que mejor se ajustaban al esquema conceptual propuesto (figura 4.1) y a los objetivos de la presente investigación.

A continuación se describen los pasos seguidos para realizar el análisis y selección de variables. El procedimiento se aplicó de igual forma a los archivos de datos de las dos muestras (2008-2009 y 2009-2010).

4.6.1 Selección preliminar de variables

Primeramente se compararon los cuestionarios utilizados en el estudio "Percepción en la Calidad de los Servicios Públicos", a fin de identificar y extraer las preguntas presentes en los cinco cuestionarios, con el objetivo de establecer una base común de variables que permitiera la comparabilidad de los análisis.

Seguidamente se seleccionó un primer grupo de variables de acuerdo al criterio de validez aparente, seleccionando aquellas variables que mejor representaban las dimensiones del esquema conceptual propuesto (figura 4.1). Las variables elegidas se consideraron mediciones de los descriptores de las dimensiones, es decir, de los factores subyacentes al factor central.

Es importante aclarar que se utilizaron únicamente las variables disponibles en los archivos de datos del estudio del cual se extrajeron los datos; es decir, no se crearon nuevas variables ni se recolectó información adicional para medir los factores del modelo propuesto.

4.6.2 Análisis de datos faltantes (no respuesta)

Al primer grupo de variables seleccionado se le aplicó un análisis de datos faltantes (no respuesta). Se estudiaron los casos o registros que presentaban preguntas sin respuesta, donde el informante no brindó información (no responde), la pregunta no le correspondía (no aplica) o indicó no saber la respuesta (no sabe).

Se analizó el porcentaje de casos sin respuesta para cada variable y siguiendo convenciones generales (Lohr, 2000), se seleccionaron las variables con la menor cantidad de datos faltantes (5% o menos), las cuales se consideran más estables. Según Lohr (2000) no es prudente aplicar algún tipo de corrección (imputación o eliminación) cuando el número de casos con no respuesta supera el 5% del total de la muestra, debido al posible efecto de distorsión en los resultados.

Tampoco se tomaron en cuenta las variables que no aplicaban a todos los informantes; es decir, preguntas no respondidas debido a saltos en el cuestionario. Esto por cuanto se tendrían tamaños de muestra distintos en los diversos análisis ya que cada variable tendría una cantidad distinta de casos válidos.

Lo que se buscó con este análisis fue obtener un archivo de datos "rectangular", sin "irregularidades" ni "vacíos" en los datos (Lohr, 2000). El fin fue generar un archivo de datos donde todas las variables tuvieran la misma cantidad de casos válidos. Con las variables seleccionadas se conformó el archivo de datos para los análisis.

4.6.3 Tratamiento de datos faltantes (no respuesta)

Se comparó la distribución por género, edad y educación de los entrevistados que respondieron y los que no lo hicieron, con el fin de determinar la posibilidad de aplicar técnicas de corrección de la no respuesta. Si las personas que no responden poseen un

comportamiento similar a los que sí lo hacen, es posible aplicar la imputación de valores o eliminación de casos sin perjudicar la representatividad de la muestra (Lohr, 2000). La imputación consiste en asignar valores estimados de reemplazo a los valores faltantes.

El problema de no respuesta en las variables de control y en la variable calificación de los servicios, se trató mediante la aplicación de técnicas de imputación calculada y lógica. En las variables observadas, utilizadas para ajustar los modelos, se aplicó la eliminación total de los casos con valores faltantes.

La imputación se aplicó con el fin de evitar una reducción en el tamaño de muestra debido a la no respuesta, la eliminación de casos se aplicó para evitar un posible efecto de la imputación sobre el ajuste de los modelos, bajo la consideración de que los valores imputados, aunque pueden ser buenas estimaciones, no son datos reales (Lohr, 2000). Por esta razón la eliminación se aplicó solamente a las variables observadas.

Para la variable *edad* se utilizaron cuatro técnicas de imputación en caliente ("hot deck"): mediana y media del conjunto de datos, media del vecino más próximo e interpolación lineal. El valor imputado final correspondió al promedio de las cuatro estimaciones obtenidas. En la variable *condición laboral* los valores faltantes se asignaron a la categoría "trabaja y estudia", que presentó la mayor frecuencia, procurando no producir un efecto distorsionador sobre la distribución de esta variable.

En la variable *nivel educativo* se aplicó una imputación deductiva, en donde el valor imputado se dedujo de acuerdo a la edad y condición laboral de cada entrevistado. La variable *calificación del servicio* se imputó asignando la categoría "regular" a los valores faltantes. Se asignaron a la categoría que ubicada a la mitad de la escala de respuesta, evitando cargar alguno de los extremos de la misma.

Posterior a la corrección de la no respuesta, se comparó la distribución de las variables control con y sin valores imputados y los tamaños de muestra, en cada aplicación de los cuestionarios, con y sin eliminación de casos. La comparación se aplicó con el fin de determinar el efecto de la imputación y de la eliminación de casos sobre las distribuciones de las variables.

4.6.4 Estandarización de variables

Con el fin de homogenizar las variables de tal forma que el valor más alto en la escala de respuesta representara el mayor nivel de la variable, brindándole una interpretación positiva, se recodificaron las variables dicotómicas, es decir, se cambiaron los códigos de respuesta. De esta forma se dio el mismo sentido o dirección a todas las variables.

Esta acción no fue necesaria para los ítems que se respondieron con una escala ordinal ya que éstos fueron redactados de forma positiva, en donde el mayor valor en la escala de respuesta representó la mayor calificación posible de la variable. Así, entre más alto el valor de la variable, mayor es la presencia del rasgo que miden.

Posteriormente se realizó una estandarización de las escalas de respuesta de las variables, de forma tal que todas se midieran en una misma escala numérica. Primero se agruparon las variables correspondientes a un mismo tema mediante una suma de los códigos de respuesta, luego, estas sumas se transformaron para que todas tuvieran una variación entre 0 y 1.

Para brindarle la misma variación a todas las variables se aplicó la transformación de máximos y mínimos (PNUD & UCR, 2011) expuesta en la ecuación 4.1

$$\frac{y_i - \text{Min.}(y_i)}{\text{Max.}(y_i) - \text{Min.}(y_i)} \quad (\text{Ecuación 4.1})$$

donde y_i es el valor observado de cada variable, $\text{Min}(y_i)$ es el límite inferior del rango de variación de la variable y_i y $\text{Max}(y_i)$ es el límite superior del rango de variación de la variable y_i .

4.7 Modelos ajustados

Con el fin de estimar un índice para medir calidad percibida en los servicios públicos, se ajustaron dos modelos alternativos: un modelo de Regresión Lineal Múltiple (RLM) y un modelo de Análisis de Factores Confirmatorio (AFC), caso particular de un Modelo de Ecuaciones Estructurales (MEE). Se calculó un índice con cada modelo.

El modelo RLM se propuso como una combinación lineal de los factores subyacentes (dimensiones) al factor central (calidad percibida) y expresa una relación directa entre el factor central y sus dimensiones. El modelo AFC se compuso de múltiples relaciones directas e indirectas entre las variables observadas y los factores subyacentes, medidas en dos niveles. El modelo AFC se propuso para estimar la estructura factorial del concepto calidad percibida.

El fin de comparar un modelo RLM con un modelo AFC, que poseen estructuras distintas, fue el de evaluar la posibilidad de obtener una medición válida y confiable con modelos de distinto nivel de complejidad, en busca del modelo más sencillo que logre cumplir con los objetivos propuestos.

4.7.1 Regresión Lineal Múltiple (RLM)

El modelo de regresión lineal múltiple corresponde a un modelo lineal aditivo probabilístico, en donde la variable de estudio⁵, es igual a la suma de los efectos producidos por cada una de las variables⁶ incluidas en el modelo, las cuales están linealmente relacionadas con la variable de estudio. El modelo mide el efecto combinado de las variables independientes sobre la variable de estudio (Cea, 2002; Catena, Ramos & Trujillo, 2003; Gujarati, 2004).

El modelo RLM correspondió a una suma ponderada de las dimensiones que componen el concepto de calidad percibida, según el esquema conceptual (figura 4.1). Mediante este modelo se midió el efecto de las dimensiones sobre la calidad percibida. La figura 4.2 muestra las dimensiones y los temas que componen el índice de calificación de calidad (ICC), ajustado con el modelo RLM.

Primeramente se calculó un indicador para medir cada tema, luego se promediaron los temas generando un subíndice para medir cada dimensión y finalmente, se agregaron de los cinco subíndices en una medición del factor central, lo que correspondió al índice de calificación de calidad.

⁵ Variable respuesta, dependiente o regresada

⁶ Variables predictoras, explicativas o independientes

**Figura 4.2 Dimensiones y temas.
Modelo de Regresión Lineal Múltiple.**



Fuente: Elaboración propia

Para generar los indicadores que representaran los temas, se obtuvo una suma de los puntajes (valores observados) de las variables que componen cada tema, según la ecuación 4.2. Esta suma representó la calificación que cada entrevistado le asigna al tema respectivo.

$$S_{Tj} = \sum_{i=1}^k y_i \quad (\text{Ecuación 4.2})$$

donde S_{Tj} es la suma de puntajes del tema j , y_i corresponde a cada variable que compone el tema j y k es la cantidad de variables que componen dicho tema.

A cada suma de puntajes se le aplicó la transformación de mínimos y máximos (ecuación 4.1). El resultado de la ecuación 4.3 correspondió al indicador utilizado para medir cada tema dentro de cada dimensión, el cual es una suma estandarizada (con variación entre 0 y 1) de las variables que lo integran.

$$I_{Tj} = \frac{S_{Tj} - \text{Min.}(S_{Tj})}{\text{Max.}(S_{Tj}) - \text{Min.}(S_{Tj})} \quad (\text{Ecuación 4.3})$$

donde I_{Tj} corresponde al indicador del tema j , S_{Tj} es la suma de puntajes del tema j , $\text{Min}(S_{Tj})$ es el mínimo valor que puede resultar de la suma de puntajes y $\text{Max}(S_{Tj})$ es el valor máximo posible de la suma de puntajes.

Para cada dimensión se calculó un subíndice, agregando los indicadores por tema que las componen. Los subíndices de las dimensiones *confiabilidad* (SC) y *responsabilidad* (SR) corresponden al promedio de sus indicadores de tema respectivos (ecuación 4.4).

$$SC = \frac{I_{T1.1} + I_{T1.2}}{2} \quad SR = \frac{I_{T2.1} + I_{T2.2}}{2} \quad (\text{Ecuación})$$

Los subíndices de las restantes dimensiones (*aseguramiento* -SA-, *empatía* -SE- y *tangibles* -ST-), adquieren el mismo valor que el indicador de tema (ecuación 4.5).

$$SA = I_{T3.1} \quad SE = I_{T4.1} \quad ST = I_{T5.1} \quad (\text{Ecuación})$$

Siguiendo el esquema conceptual (figura 4.1) y al no disponer de información que permitiera asignar ponderaciones distintas, se le asignó el mismo peso relativo a todas las variables que componen cada indicador, lo mismo que a los indicadores que componen cada subíndice y por tanto, también, a los subíndices que componen el índice general.

A partir de las 25 variables observadas se calcularon 7 indicadores de tema y 5 subíndices de dimensión, con un rango de variación entre 0 y 1.

La ecuación de regresión ajustada, utilizada para medir el efecto de las dimensiones de la calidad percibida sobre el índice de calificación de calidad, se expone en la ecuación 4.6. Esta ecuación se estimó mediante el proceso de mínimos cuadrados ordinarios (MCO).

$$ICC = \alpha + \beta_1 C + \beta_2 R + \beta_3 A + \beta_4 E + \beta_5 T + \varepsilon \quad (\text{Ecuación 4.6})$$

donde *ICC* es índice de calificación de calidad, *C* es confiabilidad, *R* es responsabilidad, *A* es aseguramiento, *E* es empatía y *T* es tangibles; α corresponde al intercepto o pendiente, los coeficientes β_i corresponden a los pesos o ponderaciones de las dimensiones de la calidad percibida (miden el efecto de cada dimensión sobre la calificación de calidad) y ε corresponde al término de error aleatorio o residuo (normalmente distribuido con media cero y variancia constante).

Las variables independientes del modelo correspondieron a los subíndices de dimensión. Cada variable independiente (subíndice) se consideró una calificación

asignada a cada dimensión. La variable dependiente (índice de calificación de calidad observado) correspondió a una estandarización de la variable "calificación del servicio recibido", por medio de la transformación de máximos y mínimos (ecuación 4.1).

La variable "calificación del servicio recibido" posee un nivel de medición ordinal, pero para efectos del presente análisis, se modeló como una variable métrica. Esta decisión se tomó siguiendo la idea que las variables ordinales "representan variables con una escala latente" lo que "posibilita su tratamiento como variables continuas" (Cea, 2002, pág. 21).

En el caso de la evaluación al MOPT, se solicitó la calificación de tres servicios específicos: prueba teórica, prueba práctica y entrega de licencia. En este caso, la calificación global del servicio brindado por el MOPT correspondió al promedio de las tres calificaciones particulares.

4.7.2 Análisis de Factores Confirmatorio (AFC)

Con el fin de estimar, conjuntamente, las relaciones existentes entre las variables observadas (indicadores medidos) y las variables latentes (factores subyacentes representados por las dimensiones), así como las ponderaciones y efectos directos e indirectos de dichas relaciones, según el modelo conceptual (figura 4.1), se propuso un AFC de segundo orden.

El AFC es un caso particular de los MEE que permiten recolectar evidencias que apoyan la existencia de una relación causal entre variables latentes o factores y entre éstos y variables observadas o indicadores medidos (Cea, 2002; Catena et al., 2003).

A partir de los resultados del modelo AFC (ponderaciones o pesos factoriales de las variables observadas y variables latentes) se calcularon los subíndices que representaron a cada dimensión (factor subyacente). Estos subíndices indicaron el valor, asignado por cada individuo, a cada dimensión del factor central.

Los subíndices de dimensión se obtuvieron mediante la suma de las variables observadas que los integran, ponderada por sus cargas factoriales respectivas y estandarizada por medio de la transformación de máximos y mínimos (ecuación 4.1).

Las ecuaciones 4.7 y 4.8 expresan el cálculo de las sumas ponderadas y de los subíndices de dimensión, respectivamente.

$$\eta_j = \sum_{i=1}^k y_i * \lambda_i \quad (\text{Ecuación 4.7})$$

donde η_j es la suma ponderada de puntajes de las variables que componen la dimensión j , y_i es el valor observado de la variable i que compone la dimensión j , λ_i es la carga factorial de la variable y_i obtenida del modelo AFC y k es el número de variables observadas incluidas en la dimensión j .

$$SD_j = \frac{\eta_j - \text{Min.}(\eta_j)}{\text{Max.}(\eta_j) - \text{Min.}(\eta_j)} * 100 \quad (\text{Ecuación 4.8})$$

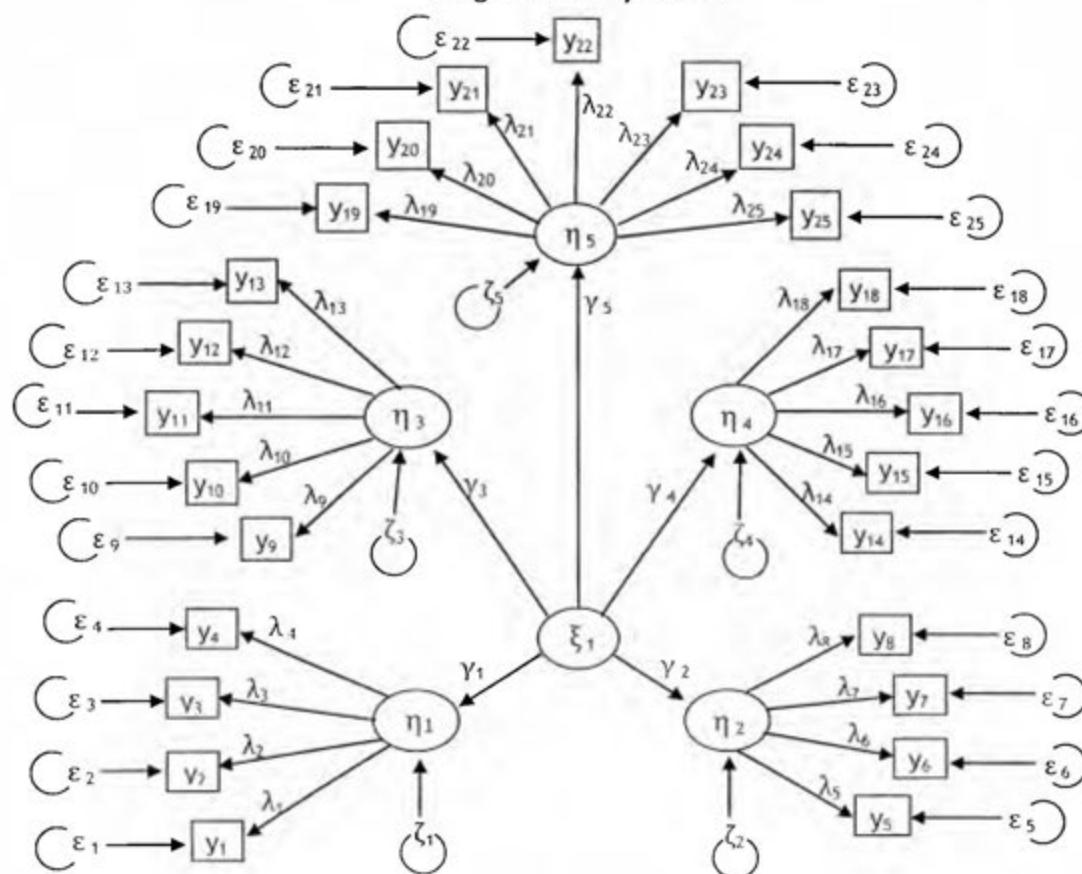
donde SD_j es el subíndice de la dimensión j , η_j es la suma ponderada de puntajes de las variables que componen la dimensión j , $\text{Min}(\eta_j)$ es el valor mínimo posible de la suma ponderada de puntajes de la dimensión j y $\text{Max}(\eta_j)$ es el valor máximo posible de la suma ponderada de puntajes de la dimensión j .

En los MEE, el modelo estructural expresa la estructura del planteamiento causal (esquema conceptual). El modelo estructural del AFC ajustado midió las relaciones lineales y los efectos del factor central (calidad percibida) sobre los factores subyacentes (dimensiones). El modelo estructural correspondió al segundo nivel del modelo AFC.

Los modelos de medición representan la relación entre cada variable latente y las variables observadas asociadas a ella. Estos modelos (primer nivel del AFC) permitieron medir las relaciones lineales y efectos de cada dimensión sobre las variables observadas.

La figura 4.3 representa el diagrama de trayectoria del modelo AFC de segundo orden ajustado mediante mínimos cuadrados ponderados (MCP), compuesto por una variable latente exógena (calidad percibida), cinco variables latentes endógenas (dimensiones) y veinticinco variables observadas (indicadores medidos).

Figura 4.3 Modelo AFC de segundo orden.
Diagrama de trayectoria.



Fuente: Elaboración propia

En la figura 4.3, cada dimensión (η) produce un efecto (λ) sobre las variables observadas asociadas a ella (y), lo que representa el primer nivel del modelo. En el segundo nivel, el factor central o variable latente exógena (ξ), que representa la calidad percibida, produce un efecto (γ) sobre cada variable latente endógena (η) que representa cada dimensión. A partir de estas relaciones y efectos se deriva el índice de calidad percibida propuesto.

En el anexo B se expone la simbología genérica de los MEE, utilizada para interpretar los coeficientes de la figura 4.3. El anexo B también incluye la interpretación de cada una de las variables latentes (exógenas y endógenas) del modelo AFC. Las variables observadas (endógenas) se encuentran identificadas en el cuadro A2 (anexo A).

Para ajustar el AFC se utilizaron directamente las variables observadas, sin ser agregadas en indicadores de tema. Se procedió de esta forma debido a que se disponía de pocos indicadores de tema: uno o dos para cada variable latente (figura 4.2) y se requiere variabilidad en las mediciones, por lo que se recomienda tener al menos 4 indicadores por cada variable latente (Mulaik, 2009).

El cuadro 4.1 muestra las ecuaciones estructurales generadas a partir del modelo AFC. Por medio de las ecuaciones estructurales se midió el efecto de las variables latentes endógenas (dimensiones) sobre las variables observadas (primer nivel) y de la variable latente exógena (calidad percibida) sobre las variables latentes endógenas (segundo nivel).

**Cuadro 4.1 Ecuaciones estructurales.
Modelo AFC de segundo orden.**

Ecuaciones de primer nivel	
$y_1 = \lambda_1\eta_1 + \varepsilon_1$	$y_{14} = \lambda_{14}\eta_4 + \varepsilon_{14}$
$y_2 = \lambda_2\eta_1 + \varepsilon_2$	$y_{15} = \lambda_{15}\eta_4 + \varepsilon_{15}$
$y_3 = \lambda_3\eta_1 + \varepsilon_3$	$y_{16} = \lambda_{16}\eta_4 + \varepsilon_{16}$
$y_4 = \lambda_4\eta_1 + \varepsilon_4$	$y_{17} = \lambda_{17}\eta_4 + \varepsilon_{17}$
$y_5 = \lambda_5\eta_2 + \varepsilon_5$	$y_{18} = \lambda_{18}\eta_4 + \varepsilon_{18}$
$y_6 = \lambda_6\eta_2 + \varepsilon_6$	$y_{19} = \lambda_{19}\eta_5 + \varepsilon_{19}$
$y_7 = \lambda_7\eta_2 + \varepsilon_7$	$y_{20} = \lambda_{20}\eta_5 + \varepsilon_{20}$
$y_8 = \lambda_8\eta_2 + \varepsilon_8$	$y_{21} = \lambda_{21}\eta_5 + \varepsilon_{21}$
$y_9 = \lambda_9\eta_3 + \varepsilon_9$	$y_{22} = \lambda_{22}\eta_5 + \varepsilon_{22}$
$y_{10} = \lambda_{10}\eta_3 + \varepsilon_{10}$	$y_{23} = \lambda_{23}\eta_5 + \varepsilon_{23}$
$y_{11} = \lambda_{11}\eta_3 + \varepsilon_{11}$	$y_{24} = \lambda_{24}\eta_5 + \varepsilon_{24}$
$y_{12} = \lambda_{12}\eta_3 + \varepsilon_{12}$	$y_{25} = \lambda_{25}\eta_5 + \varepsilon_{25}$
$y_{13} = \lambda_{13}\eta_3 + \varepsilon_{13}$	
Ecuaciones de segundo nivel	
$\eta_1 = \gamma_{11}\xi_1 + \zeta_1$	$\eta_4 = \gamma_{14}\xi_4 + \zeta_4$
$\eta_2 = \gamma_{12}\xi_2 + \zeta_2$	$\eta_5 = \gamma_{15}\xi_5 + \zeta_5$
$\eta_3 = \gamma_{13}\xi_3 + \zeta_3$	

Fuente: Elaboración propia

4.8 Evaluación del modelo

Con el fin de determinar si cada modelo propuesto es adecuado para medir el concepto *calidad percibida* por medio de las relaciones con sus dimensiones, se puso a prueba la bondad de ajuste de cada modelo. El fin de la evaluación es comprobar que el modelo propuesto es el que genera los errores más bajos en las estimaciones y que además, logra cumplir con las condiciones necesarias para modelar la variable de estudio.

4.8.1 Regresión Lineal Múltiple (RLM)

En los modelos RLM, la bondad de ajuste se evalúa mediante el análisis de los errores de ajuste o residuos (diferencias entre los valores observados en la muestra y los valores predichos o pronosticados por el modelo). El objetivo es buscar estimaciones de los coeficientes que minimicen los errores de ajuste (Ruiz, Pardo & San Martín, 2010).

Como medida de ajuste se utilizó el coeficiente de determinación (R^2), que mide la proporción de la variación total de la variable respuesta (ICC) que se debe a la variación de (o que es explicada por) las variables predictoras, en este caso, las dimensiones de la calificación de calidad (Cea, 2002; Catena et al., 2003; Gujarati, 2004).

También, se realizó un análisis de varianza (ANDEVA) para probar la significancia del modelo, por medio del estadístico F de Fisher (cociente entre la varianza explicada por el modelo y la varianza residual). Por medio del ANDEVA se puso a prueba si el efecto conjunto de las variables independientes (dimensiones) sobre la variable respuesta (ICC) es significativamente diferente de cero (Cea, 2002).

Adicionalmente, se aplicó una prueba de significancia a los coeficientes individuales, mediante el estadístico t de Student (cociente entre cada coeficiente y su error estándar). Con esta prueba se analizó si el efecto individual de cada variable independiente sobre la variable respuesta es significativamente diferente de cero (Cea, 2002).

El ANDEVA y las pruebas t se utilizaron para comprobar que las ponderaciones de las variables independientes (dimensiones) son no nulas (diferentes de cero), lo que indicaría que las dimensiones poseen una relación lineal con el ICC y por tanto, producen un efecto sobre el índice y aportan información para explicar su variación.

Para evaluar el modelo, también se realizaron los diagnósticos de prueba o verificación de supuestos, por medio de análisis gráficos y pruebas formales con estadísticos de contraste, siguiendo recomendaciones indicadas en la bibliografía consultada (Cea, 2002; Catena et al., 2003; Gujarati, 2004). Los supuestos verificados y sus respectivas pruebas se mencionan a continuación:

- Linealidad: comprobar que existe relación lineal entre las variables predictoras y la variable respuesta. Prueba: significancia de los coeficientes de regresión, gráfico de dispersión entre los valores pronosticados y los residuos (errores de pronóstico).
- Normalidad de los residuos: comprobar que los residuos o errores de pronóstico siguen una distribución normal [$\varepsilon_i \sim N(0, \sigma^2)$]. Prueba: gráfico de normalidad (gráfico P-P) y pruebas de normalidad Kolmogorov-Smirnov y Shapiro-Wilk.
- Media cero de los residuos: comprobar que los residuos poseen una media o esperanza matemática nula [$E(\varepsilon_i) = 0$]. Prueba: se probó junto al diagnóstico anterior.
- Igualdad de varianza de los residuos (homoscedasticidad): comprobar que las varianzas de los residuos son iguales [$Var(\varepsilon_i) = \sigma^2$]. Prueba: gráfico de residuos versus valores pronosticados y prueba de igualdad de varianzas de Leven.
- Ausencia de autocorrelación en residuos: comprobar que los residuos son independientes; es decir, no hay relación entre ellos [$Cov(\varepsilon_i, \varepsilon_j) = 0$]. Prueba: gráfico de dispersión entre el orden temporal de los datos y los errores de predicción y prueba de independencia mediante el estadístico Durbin-Watson.
- Ausencia de correlación entre residuos y variables independientes: comprobar que los residuos son independientes de las variables predictoras; es decir, no hay relación entre ellos. Prueba: gráfico de dispersión matricial entre variables predictoras y residuos, además del gráfico de residuos versus valores pronosticados.
- Ausencia de multicolinealidad: comprobar que no existe correlación alta entre pares de variables predictoras; es decir, que son independientes. Prueba: factor de incremento de la varianza (FIV) e índice de tolerancia (IT).
- Tamaño de muestra suficiente: comprobar que se cuenta con la cantidad suficiente de observaciones para ajustar el modelo propuesto. Prueba: que el tamaño de muestra cumpla con los criterios mencionadas en la literatura (Cea, 2002; Catena et al., 2003; Gujarati, 2004): multiplicar la cantidad de variables independientes por un factor que va desde 5 hasta 20.

4.8.2 Análisis de Factores Confirmatorio (AFC)

El modelo AFC se evaluó en cada una de las etapas de su construcción, según la metodología de los modelos de ecuaciones estructurales y de acuerdo con las recomendaciones expuestas en la bibliografía consultada (Bollen, 1989; Diamantopoulos

& Siguaaw, 2000; Cea, 2002; Catena et al., 2003; Arias, 2008; Orgaz, 2008; Ruiz, 2008; Kaplan, 2009; Mulaik, 2009; Freiberg, Stover & Fernández, 2013).

4.8.2.1-Especificación

La especificación del modelo consiste en proponer un modelo hipotético con base en teorías existentes en el campo de análisis y en investigaciones previas sobre el tema, el cual se contrasta con el modelo ajustado con los datos observados. En esta etapa se establecieron las variables incluidas en el modelo ajustado y las relaciones entre ellas; es decir, se especificó el modelo estructural y los modelos de medición.

La evaluación de esta etapa consistió en analizar las razones para escoger el planteamiento conceptual utilizado (figura 4.1) y las variables observadas del modelo ajustado (figura 4.3), con el fin de determinar su idoneidad.

4.8.2.2-Identificación

En la identificación del modelo se determina la posibilidad de que exista uno o varios conjuntos de valores para los parámetros del modelo, a partir de los datos observados. Para estimar el modelo, los parámetros libres deben estar al menos identificados (Mulaik, 2009). La identificación de los parámetros es necesaria para reproducir la matriz de covarianzas del modelo, lo que permite dar solución a las ecuaciones estructurales.

Para que los parámetros del modelo estén identificados, la cantidad de parámetros libres a estimarse en el modelo debe ser menor o igual al número de unidades únicas de información de la matriz de varianzas-covarianzas, ajustada con los datos observados en la muestra (Orgaz, 2008). Para estimar la cantidad de unidades únicas de información de la matriz de varianzas-covarianzas muestral, se utilizó la ecuación 4.9.

$$u = \frac{p(p+1)}{2} \quad (\text{Ecuación 4.9})$$

donde u es la cantidad de unidades únicas de información de la matriz de varianzas-covarianzas y p es la cantidad de variables observadas, utilizadas en el ajuste del modelo.

Adicionalmente, para obtener estimaciones válidas de los parámetros, se requiere fijar la métrica (unidades de medida) de las variables latentes o factores subyacentes (Mulaik, 2009). Una forma para establecer la métrica de los factores consiste en fijar uno de los coeficientes estructurales de las variables latentes (λ o β) en 1, de esta forma la escala de la variable latente se hace igual a la escala de la variable observada (x o y) correspondiente (Catena et al., 2003).

Para ajustar el modelo AFC, se fijó una de las cargas factoriales de las variables observadas asociadas a cada variable latente, en uno; es decir, se impuso una restricción para cada factor. El cuadro 4.2 muestra las variables cuyas cargas se fijaron en uno.

Cuadro 4.2
Restricciones impuestas: variables con cargas fijadas en 1

Dimensión	Variable
Confiabilidad	y_1 Solicitud de información que ya había entregado
Responsabilidad	y_8 Tiempo de atención
Aseguramiento	y_{10} Preparación de los funcionarios
Empatía	y_{15} Compromiso con el trabajo
Tangibles	y_{22} Estado del mobiliario

Fuente: Elaboración propia

La escogencia de estas variables se realizó de acuerdo con el nivel en que representaban a cada dimensión, según los descriptores de cada una de ellas (figura 4.1); es decir, se escogieron las variables que mejor representaban cada dimensión.

4.8.2.3-Estimación

El objetivo de esta etapa es identificar una función de discrepancia para estimar los parámetros (coeficientes) del modelo, que permita minimizar la diferencia entre la matriz de covarianzas muestral (estimada) y la matriz de covarianzas poblacional (hipotetizada).

Los parámetros del modelo AFC se estimaron con la función de discrepancia basada en el método de mínimos cuadrados ponderados (MCP). Como matriz de ponderación se utilizó la inversa de la matriz de covarianzas asintótica estimada de las correlaciones policóricas.

Para evaluar la significancia de los parámetros estimados se aplicó la prueba por medio del estadístico t de Student, que se obtiene como el cociente entre cada coeficiente y su respectivo error estándar. Con esta prueba se contrasta la hipótesis que cada coeficiente (λ , γ , ε ó ζ) es igual a cero con la hipótesis que cada uno es diferente de cero.

Los coeficientes estimados corresponden a las cargas factoriales de primer (λ) y segundo orden (γ) y a los términos de error de variables observadas (ε) y variables latentes (ζ). Un coeficiente no significativamente diferente de cero indicaría que el factor subyacente no produce ningún efecto sobre la variable latente u observada correspondiente.

Adicionalmente, se comprobó que las estimaciones tuvieran valores consistentes; es decir, que no tomaran valores impropios (casos Heywood). Se verificó que los coeficientes tuvieran la dirección adecuada (signo correcto) según el esquema conceptual, que las varianzas y errores estándar resultaran positivos, que los coeficientes estandarizados no resultaran mayores a uno y que las correlaciones estuvieran entre sus límites (-1 y 1).

4.8.2.4- Bondad de ajuste

En los MEE y en consecuencia en los modelos AFC, el ajuste se evalúa comparando la matriz de covarianzas teórica, propuesta por el modelo, con la matriz de covarianzas de los datos observados, calculada a partir de la muestra. (Bollen, 1989; Diamantopoulos & Siguaw, 2000; Cea, 2002; Catena et al., 2003; Kaplan, 2009; Mulaik, 2009).

En los MEE los residuos corresponden a la diferencia entre las covarianzas pronosticadas por el modelo teórico y las covarianzas observadas en la muestra (Ruiz, 2008). El mejor ajuste se logra cuando las estimaciones de los parámetros producen los residuos más bajos posibles. Para evaluar el ajuste del modelo se puso a prueba la hipótesis de igualdad entre la matriz de covarianzas poblacional (Σ) y la matriz de covarianzas estimada ($\Sigma[\theta]$).

La hipótesis de igualdad de matrices se probó mediante distintos índices de ajuste, los cuales analizan la capacidad del modelo estimado para reproducir la matriz de

parámetros muestrales, esto es, la matriz de varianzas y covarianzas calculadas en la muestra (Ruiz, 2008). Los índices utilizados y sus umbrales se indican en la cuadro 4.3.

Cuadro 4.3. Índices de ajuste. Modelo AFC

Tipo	Nombre del índice	Siglas*	Límites de evaluación
Absoluto	Prueba con estadístico chi-cuadrado	p	> 0,05
Absoluto	Razón chi-cuadrado / grados de libertad	χ^2/df	< 3
Absoluto	Índice de bondad de ajuste	GFI	> 0,90
Absoluto	Raíz del residuo cuadrático medio	RMR	< 0,05
Absoluto	Raíz del residuo cuadrático medio estandarizado	SRMR	< 0,09
Absoluto	Índice de Centralidad	CI	> 0,90
Absoluto	Raíz del error cuadrático medio de aproximación	RMSEA	< 0,05
Comparativo	Índice de ajuste comparativo	CFI	> 0,95
Comparativo	Índice de ajuste normado	NFI	> 0,90
Comparativo	Índice de ajuste no normado	NNFI	> 0,95
Comparativo	Índice de ajuste incremental	IFI	> 0,95
Comparativo	Índice de ajuste relativo	RFI	> 0,95
Parsimonia	Índice parsimonioso de ajuste normado	PNFI	> 0,80
Parsimonia	Índice parsimonioso de bondad de ajuste	PGFI	> 0,80
Parsimonia	Índice ajustado de bondad de ajuste	AGFI	> 0,90

*Siglas del nombre en inglés

Fuente: Elaboración propia.

Los límites establecidos para cada índice se propusieron siguiendo las reglas mencionadas en la bibliografía consultada (Diamantopoulos & Siguaw, 2000; Cea, 2002; Catena et al., 2003; Orgaz, 2008; Ruiz, 2008; Mulaik, 2009).

Otro aspecto importante para evaluar la bondad de ajuste del modelo es utilizar más de un índice de ajuste (Catena et al., 2003) con el fin de realizar un análisis más robusto. De acuerdo con Diamantopoulos & Siguaw (2000) la prueba chi-cuadrado y los índices RMSEA, SRMR, GFI y CFI son más que suficientes para comprobar el ajuste del modelo.

4.8.2.5- Supuestos

Como parte de la evaluación del modelo AFC se aplicaron los diagnósticos de prueba (verificación de los supuestos del modelo), por medio de análisis gráficos y cálculo de estadísticos, siguiendo reglas mencionadas en la bibliografía (Diamantopoulos & Siguaw,

2000; Cea, 2002; Catena et al., 2003; Orgaz, 2008; Ruiz et al., 2010; Morales, 2011; Freiberg et al., 2013). Los supuestos sometidos a prueba fueron los siguientes:

- Linealidad: existencia de relaciones lineales entre las variables observadas y las variables latentes y entre pares de variables latentes. Prueba: se comprobó con la especificación del modelo según diagrama de trayectoria (figura 4.3), la significancia de los parámetros estimados y por medio de gráficos de dispersión entre variables observadas y subíndices, entre variables e índice y entre subíndices e índice.
- Ausencia de colinealidad y singularidad: correlación nula entre pares de variables observadas y no singularidad de la matriz de correlaciones. Prueba: se analizó la matriz de correlaciones policóricas y las estimaciones de los coeficientes.
- Tamaño de muestra suficiente: contar con la cantidad suficiente de observaciones para realizar las estimaciones de los parámetros. Prueba: se aplicaron reglas heurísticas mencionadas en la bibliografía consultada para determinar el tamaño de muestra mínimo: desde 5 hasta 15 sujetos por cada variable observada, de 8 a 10 veces más sujetos que variables (observadas y latentes), utilizar una muestra de 1.000 sujetos para modelos con más de 15 variables observadas y métodos de estimación de distribución libre asintótica (Diamantopoulos & Siguaaw, 2000; Cea, 2002; Catena et al., 2003; Orgaz, 2008).
- Aditividad de los efectos: los efectos de las variables latentes sobre las variables observadas pueden sumarse entre sí. Prueba: se analizó el ajuste de las ecuaciones estructurales.
- Indicadores múltiples: disponer de más de un indicador o variable observada relacionada con cada variable latente, procurando una distribución del error. Prueba: aplicar las recomendaciones mencionadas en la bibliografía consultada: utilizar 5 o 6 variables latentes con 3 o 4 indicadores asociados para conjuntos de 20 variables observadas o utilizar 4 o 5 indicadores por cada variable latente para generar modelos sobre-identificados (Bollen, 1989; Diamantopoulos & Siguaaw, 2000; Kaplan, 2009; Mulaik, 2009).
- Variables métricas: utilizar variables medidas en escala de intervalo o razón para ajustar el modelo. Prueba: ajuste del modelo según recomendaciones mencionadas en la bibliografía consultada: utilizar la matriz de correlaciones policóricas para ajustar modelos con variables ordinales (Bollen, 1989; Diamantopoulos & Siguaaw, 2000; Orgaz, 2008; Ruiz et al., 2010).

No se puso a prueba el supuesto de normalidad multivariada debido a que el modelo se ajustó por medio de mínimos cuadrado ponderados, los cuales no requieren del supuesto de normalidad (Bollen, 1989; Diamantopoulos & Siguaw, 2000; Arias, 2008). De acuerdo con Freiberg et al. (2013) los resultados obtenidos con mínimos cuadrado ponderados no se afectan sustancialmente cuando se viola el supuesto de normalidad.

4.9 Índice calculado

Para obtener la estimación del índice para medir calidad percibida en servicios públicos se aplicaron los modelos ajustados a los datos muestrales con lo cual se obtuvieron dos puntajes, uno con cada modelo: el índice de calificación de calidad (ICC), calculado a partir del modelo RLM y el índice de calidad percibida (ICP), calculado a partir del modelo AFC. En este apartado se detalla el cálculo de cada índice.

4.9.1 Índice de calificación de calidad

El índice de calificación de calidad observado se calculó con base en los datos de la muestra y se utilizó como variable respuesta del modelo RLM. El resultado de la ecuación de regresión corresponde a una estimación del ICC observado.

El ICC observado en la muestra se calculó a partir de la variable "calificación del servicio recibido" que corresponde a una variable ordinal de siete categorías, medida mediante una escala Likert y que se analizó como variable métrica. A la variable "calificación del servicio recibido" se le aplicó la transformación de máximos y mínimos y se multiplicó por 100 (ecuación 4.10), dando como resultado un puntaje estandarizado (con variación entre 0 y 100), correspondiente al valor observado del ICC.

$$ICC = \frac{C_i - \text{Min}(C_i)}{\text{Max}(C_i) - \text{Min}(C_i)} = \frac{C_i - 1}{6} * 100 \quad (\text{Ecuación 4.10})$$

donde ICC es el índice de calificación de calidad observado y C_i es la calificación del servicio recibido, asignada por el sujeto i

4.9.2 Índice de calidad percibida

El índice de calidad percibida se calculó a partir de los coeficientes del segundo nivel del modelo AFC de segundo orden ajustado y corresponde a un puntaje estandarizado (que varía entre 0 y 100) que permite cuantificar el nivel de calidad que los usuarios de los servicios públicos han percibido al momento de recibir el servicio.

El ICP se obtuvo como la suma de los puntajes de las dimensiones (ecuación 4.7), ponderada por sus respectivas cargas factoriales de segundo orden (γ), según la ecuación 4.11 y estandarizada con la transformación de máximos y mínimos, multiplicada por cien (ecuación 4.12).

$$\xi_i = \sum_{j=1}^5 \eta_j * \gamma_j \quad (\text{Ecuación 4.11})$$

donde ξ_i es la suma ponderada de las dimensiones de calidad percibida para el sujeto i , η_j es la suma ponderada de puntajes de la dimensión j y γ_j es la carga factorial de segundo orden de la dimensión η_j obtenida del modelo AFC.

$$ICP = \frac{\xi_i - \text{Min}(\xi_i)}{\text{Max}(\xi_i) - \text{Min}(\xi_i)} * 100 = \frac{\xi_i - 11,12}{53,23} * 100 \quad (\text{Ecuación 4.12})$$

donde ICP es el índice de calidad percibida, ξ_i es la suma ponderada de las dimensiones de calidad percibida para el sujeto i , $\text{Min}(\xi_i)$ es el valor mínimo posible de ξ_i y $\text{Max}(\xi_i)$ es el máximo valor posible de ξ_i .

4.10 Fiabilidad

La fiabilidad de una medición hace referencia a la confianza que se pueda tener en el resultado de dicha medición. Cuando un instrumento de medición se aplica en forma repetida y las mediciones obtenidas son estables, consistentes y precisas, se dice que son fiables (Alavaro, 1997). La fiabilidad implica obtener los mismos resultados en mediciones repetidas de la misma muestra, bajo las mismas condiciones (Prieto & Delgado, 2010).

La fiabilidad no es una característica de los instrumentos de medición (tests, índices o escalas) sino que es una propiedad de la interpretación del resultado de la medición y del uso que se le pueda dar a la medida resultante (Prieto & Delgado, 2010).

Para medir el grado de fiabilidad de los índices calculados, se estimó el coeficiente de fiabilidad, aplicando dos enfoques diferentes, propuestos en la Teoría Clásica de los Test (TCT) y mencionados en la bibliografía consultada (Alvaro, 1997; Aiken, 2003; Abad, Garrido, Olea & Ponsada, 2006; Morales, 2007; Prieto & Delgado, 2010).

- 1) Dos mitades: consiste en dividir el instrumento de medición en dos mitades, consideradas dos subconjuntos de ítems diferentes pero equivalentes. Este enfoque se aplicó en ambas muestras (calibración y validación) por separado.

En el modelo RLM, el índice calculado se obtuvo como una combinación lineal de los subíndices de dimensión. Para evaluar su fiabilidad, se agruparon los 5 subíndices de dimensión: las dimensiones confiabilidad, aseguramiento y tangibles en un grupo y las dimensiones responsabilidad y empatía, en otro.

En el caso del modelo AFC, el índice calculado se obtuvo a partir de las 25 variables observada. Para evaluar su fiabilidad, se distribuyeron las 25 variables en dos grupos: las variables pares en un grupo y las variables impares en otro grupo.

Con cada grupo de subíndices o variables observadas se obtuvo un puntaje, calculado como una suma de los puntajes de los subíndices o variables agrupadas, ponderadas por sus respectivas cargas factoriales. Posteriormente, se calculó el coeficiente de correlación de Pearson entre las puntuaciones de ambas mitades.

Debido a que en cada caso no se trabaja con el instrumento completo, sino con la mitad del mismo, al coeficiente de correlación de Pearson se le aplicaron las correcciones de Spearman-Brown (ecuación 4.13) y Guttman-Flanagan (ecuación 4.14), las cuales se consideran coeficientes de fiabilidad.

$$r_{SB} = \frac{2r_p}{1 + r_p} \quad (\text{Ecuación 4.13})$$

donde r_{SB} es el coeficiente de fiabilidad Spearman-Brown y r_p es el coeficiente de correlación de Pearson entre las dos mitades.

$$r_{GF} = 2 \left(1 - \frac{[S_1^2 + S_2^2]}{S_T^2} \right) \quad (\text{Ecuación 4.14})$$

donde r_{GF} es el coeficiente de fiabilidad Guttman-Flanagan, S_1^2 es la varianza del puntaje calculado con la primera mitad, S_2^2 es la varianza del puntaje calculado con la segunda mitad y S_T^2 es la varianza del puntaje calculado con todos los ítems del instrumento.

- 2) Consistencia interna: en este enfoque se analizan todos los ítems del instrumento de medición, ya que se considera que todos ellos miden el mismo rasgo de interés. Al analizar todos los ítems se supone que se están analizando todas las posibles mitades que se pueden construir a partir de los ítems del instrumento.

El coeficiente de fiabilidad, en este caso llamado coeficiente de consistencia interna, mide la correlación múltiple entre cada variable y el índice calculado. La consistencia interna del índice calculado se midió por medio del coeficiente alfa de Cronbach (ecuación 4.15), que representa a la media de los coeficientes de fiabilidad de todas las posibles mitades.

$$\alpha = \frac{N}{N-1} \left(1 - \frac{\sum_{i=1}^N S_i^2}{S_T^2} \right) \quad (\text{Ecuación 4.15})$$

donde N es el número total de ítems, S_i^2 es la varianza de cada ítem y S_T^2 es la varianza del índice calculado.

Con el fin de realizar una medición más robusta, se calculó el coeficiente alfa de Cronbach aplicando el enfoque de dos mitades, evaluando la consistencia interna de cada mitad. Adicionalmente, como los modelos se ajustaron con variables ordinales, también se calculó el coeficiente alfa ordinal (ecuación 4.16) el cual corresponde a la versión del coeficiente alfa de Cronbach cuando las variables son ordinales.

$$\alpha = \frac{N}{N-1} \left(\frac{N(\bar{\lambda})^2 - \bar{\lambda}^2}{N(\bar{\lambda})^2 + \bar{u}^2} \right) \quad (\text{Ecuación 4.16})$$

donde N es el número total de ítems, $\bar{\lambda}$ es el promedio de las cargas factoriales, $\overline{\lambda^2}$ es el promedio de las cargas factoriales cuadráticas o comunalidades y $\overline{u^2}$ es el promedio de las unicidades o errores.

Debido a que los coeficientes de fiabilidad proceden del coeficiente de correlación, su interpretación es similar: valores del coeficiente cercanos a 1 indican alta fiabilidad de las puntuaciones. En esta caso se utilizó como umbral 0,60 por tanto, valores del coeficiente de 0,60 o mayores se interpretan como alta fiabilidad en las puntuaciones del índice (Aiken, 2003; Morales, 2007; Prieto & Delgado, 2010).

4.11 Validez

La validez es la característica que permite la interpretación de las puntuaciones obtenidas con los instrumentos de medición (índices, tests y escalas) sobre un concepto determinado, respaldadas por evidencias empíricas y teorías sobre dichas mediciones (Prieto & Delgado, 2010). El análisis de validez se centra en la interpretación de los puntajes y no en el instrumento de medición.

La validación de las puntuaciones se realiza por medio de recolección de pruebas que apoyen su interpretación y su uso. Evidencias importantes de validez se refieren al contenido y a la estructura interna del instrumento, a los procesos de respuesta y a la relación con otras variables o criterios externos (Alvaro, 1997; Prieto & Delgado, 2010). Para Mesik (1980, citado por Pérez, Chacón & Moreno, 2000), la principal evidencia de validez se refiere al constructo, pues integra aspectos de criterio, de contenido y de estructura interna.

Por medio de las evidencias de constructo se prueban las hipótesis sobre las relaciones teóricas importantes entre los conceptos, llegando a la conclusión que el instrumento logra medir el concepto que pretende medir.

4.11.1 Evidencias de contenido

La evidencia de contenido hace referencia a la importancia de la composición del instrumento de medición (Alvaro, 1997). Las mediciones de un instrumento son válidas si

los ítems que lo componen son una muestra del contenido a evaluar (Prieto & Delgado, 2010).

Para recolectar evidencias de contenido se compararon las variables incluidas en cada modelo ajustado con los descriptores que definen las dimensiones del concepto. El análisis consistió en determinar si la selección de variables está apoyada por el esquema conceptual (figura 4.1). Si las variables seleccionadas coinciden con los descriptores se concluye que miden los conceptos que se supone que miden.

Adicionalmente se analizó la estructura interna de los índices calculados, con el fin de comprobar su dimensionalidad. Para ello se aplicaron análisis de factores exploratorios (AFE), con el método de extracción de factores de ejes principales (FEP), recomendado para variables no normales, y la rotación Varimax, para generar factores ortogonales.

Se realizaron dos análisis de factores, primero sin especificar la cantidad de factores a ajustar y luego, definiendo previamente el número de factores, de acuerdo a cada modelo ajustado. Se compararon las cargas factoriales y la varianza explicada de ambos análisis. Sólo las variables con cargas factoriales mayores a 0,50 se consideraron como pertenecientes a un factor (Cea, 2002; Morales, 2011; Freiberg et al., 2013).

En el modelo RLM, el AFE se ajustó con los subíndices de dimensión (variables independientes del modelo). En el caso del modelo AFC, el AFE se aplicó en dos niveles, primero con las variables observadas (primer nivel) y luego con los subíndices de dimensión (segundo nivel). Los análisis se aplicaron con los datos de la muestra de validación.

La dimensionalidad se confirma si en el primer nivel del análisis resultan cinco factores diferenciados, correspondientes a las dimensiones del esquema conceptual, y un solo factor en el segundo nivel, que corresponde al concepto medido. En el caso del modelo RLM, se espera un solo factor correspondiente al concepto medido.

También, se analizaron las correlaciones ítem-puntaje, por medio de los coeficientes de Spearman y de contingencia para variables ordinales y dicotómicas, respectivamente. Se analizaron las correlaciones entre las variables observadas y los subíndices por

dimensión, entre los subíndices y el índice general y entre las variable observadas y el índice general.

Se consideraron altas, las correlaciones superiores a 0,60 y menores a 0,40 se consideraron bajas (Cea, 2002; Morales, 2011). Si las cargas factoriales del AFE y las correlaciones superan el umbral establecido, se consideran evidencias de contenido a favor las puntuaciones de los índices calculados.

4.11.2 Evidencias de criterio

La evidencia de criterio se demuestra si se logra establecer una relación entre las puntuaciones obtenidas con el instrumento de medición y otras variables externas al modelo que miden el mismo concepto. El criterio externo es otra medida de la variable o concepto de interés (Prieto & Delgado, 2010).

Las evidencias de criterio se obtuvieron comparando los resultados de los índices calculados (ICC e ICP) con el índice de gestión institucional (IGI). Esta evidencia fue de tipo predictiva, dado que el criterio (IGI) se obtuvo posterior a la aplicación del instrumento con el cual se calcularon los índices evaluados. La validez se confirma si se logran establecer evidencias de relación entre lo índices calculados y el criterio externo.

En el análisis se compararon los índices calculados con ambas muestras (calibración y validación), correspondientes a mediciones 2008 y 2009, con el IGI 2010 y 2011. Cabe aclarar que el IGI se comenzó a aplicar en 2010 y en 2011 se cambió su metodología.

En el cálculo del IGI 2010 solamente se incluyó al RN y al MOPT. A las instituciones faltantes se les asignó el valor del IGI del grupo al cual pertenecen, según la desagregación de dicho índice: a la DGSC se le asignó el puntaje del sector trabajo, a la DGME el puntaje del sector seguridad ciudadana y justicia y a la SETENA se le asignó el puntaje del sector ordenamiento territorial y vivienda.

Como medida de validez se utilizó el coeficiente de correlación de Pearson entre el ICC, el ICP y el IGI, calculado a partir de los puntajes de cada índice en cada uno de los

cinco servicios evaluados. Un coeficiente de correlación mayor a 0,60 se consideró como muestra de evidencias de criterio (Cea, 2002; Catena et al., 2003).

A nivel total, el cálculo del IGI incluye a más de 100 instituciones públicas. Debido a lo anterior, para homologar criterios y realizar las comparaciones entre los índices, se calculó el promedio de las cinco instituciones analizadas en la presente investigación.

Como un segundo criterio externo se utilizaron los diagnósticos aplicados por la Secretaría Técnica del SNCS, tomando en cuenta la limitante que los diagnósticos solamente incluyeron a tres de los cinco servicios evaluados.

El análisis consistió en comparar los resultados de los subíndices de dimensión del ICC y del ICP con los resultados de las dimensiones de los diagnósticos. En tanto los resultados con ambos instrumentos sean congruentes, es decir, que permitan llegar a conclusiones similares, se recolectan evidencia de criterio.

Para realizar el análisis comparativo se utilizaron dos indicadores, calculados a partir de las inconformidades presentadas por los clientes externos: el complemento del porcentaje de inconformidades recibidas y el porcentaje de inconformidades resueltas. Se parte de la premisa que a menor cantidad de inconformidades presentadas y a mayor cantidad de inconformidades resueltas, es mayor la calidad del servicio evaluado.

4.11.3 Evidencias de constructo

Las evidencias de constructo se presentan cuando hay una concordancia entre los resultados del instrumento de medición y los planteamientos teóricos sobre el concepto evaluado (Alvaro, 1997). De acuerdo con Pérez (2000) el AFC es la mejor técnica que se puede utilizar para recolectar evidencias de constructo.

La evidencia de constructo se obtuvo mediante el análisis de los resultados y el ajuste del modelo AFC de segundo orden. Un buen ajuste del modelo confirma la estructura factorial o subyacente del constructo (calidad percibida) y la estructura del planteamiento teórico (relaciones entre las variables).

Como evidencias de constructo se utilizaron la fiabilidad de constructo, la varianza extraída y la fiabilidad de las variables. La fiabilidad de constructo (ecuación 4.17), mide la consistencia interna de las variables observadas dentro de cada variable latente (Arias, 2008). Este indicador determina si las variables observadas son consistentes en la medida del concepto de la forma en que se especificó en el modelo (Fernández, 2008).

$$FC = \frac{(\sum_{i=1}^p \lambda_i)^2}{(\sum_{i=1}^p \lambda_i)^2 + \sum_{i=1}^p \varepsilon_i} \quad (\text{Ecuación 4.17})$$

donde FC es la fiabilidad de constructo, λ_i es la carga factorial estandarizada de la variable observada y_i , p es la cantidad de variables independientes asociadas a cada factor o dimensión y ε_i es la carga factorial estandarizada del error asociado a la variable y_i .

La varianza extraída (ecuación 4.18) corresponde al porcentaje de la varianza de las variables observadas que explican cada variable latente (Fernández, 2008). Este indicador compara la cantidad de varianza de las variables observadas extraída por cada factor con la varianza extraída por el error de medida (Arias, 2008). La varianza extraída es una medida complementaria a la de fiabilidad de constructo.

$$S^2 = \frac{\sum(\lambda_i^2)}{p} \quad (\text{Ecuación 4.18})$$

donde S^2 es la varianza extraída, λ_i es la carga factorial estandarizada de la variable observada y_i y p es la cantidad de variables independientes asociadas a cada factor o dimensión.

Se calculó la fiabilidad de constructo y la varianza extraída para cada dimensión, analizando las variables observadas que las componen y para el índice general, utilizando las cargas factoriales de segundo orden, correspondientes a las dimensiones.

La fiabilidad de las variables corresponde a la correlación cuadrada entre las variables observadas y la variable latente asociada (correlación múltiple cuadrada) e indica la proporción de la varianza de la variable observada que es explicada por la variable latente que se supone que mide (Arias, 2008).

Para evaluar la estructura factorial del índice, se utilizaron los criterios planteados por Cea (2002), Morales (2007), Arias (2008) y Fernández (2008): cargas factoriales y varianzas extraídas superiores a 0,50, fiabilidad de las variables de 0,60 o mayor y fiabilidad de constructo de al menos 0,70.

4.12 Invarianza factorial

La evaluación de la invarianza factorial se realizó con el fin de determinar la posibilidad interpretar las puntuaciones resultantes de las mediciones en grupos distintos, de tal modo que el significado sustantivo de la medición sea independiente de los grupos evaluados. Con la comprobación de invarianza factorial se confirma que las puntuaciones observadas dependen únicamente de las variables latentes que miden (Elosua, 2005).

Al establecerse la invarianza factorial se concluye que el modelo logra medir el mismo concepto entre grupos distintos (Byrne, 2008). Lo anterior evidencia que las puntuaciones tienen el mismo significado en los distintos grupos evaluados, por lo que pueden ser interpretadas de la misma forma, permitiendo realizar comparaciones.

El análisis de la invarianza factorial consistió en ajustar el modelo de análisis de factores confirmatorio en diferentes grupos de poblaciones o subpoblaciones e imponiendo restricciones al modelo. El fin del análisis fue determinar cuáles componentes del modelo se mantienen sin variación a través de los grupos, es decir, hasta qué punto se mantiene la misma estructura factorial entre los grupos (Barrera, Navarro & Peris, 2015).

La evaluación de la invarianza factorial se realizó de forma progresiva, ajustando el modelo en las dos muestras de estudio (muestra de calibración y muestra de validación), aplicando restricciones a los parámetros en cada nivel y comparando los resultados de ambos grupos. Seguidamente se menciona el proceso de análisis de evaluación de la invarianza factorial aplicado (Byrne, 2008; Timmos, 2010; Barrera et al., 2015):

- 1- Invarianza configural: se ajustó el mismo modelo en cada muestra manteniendo el mismo planteamiento conceptual, es decir, el mismo patrón de parámetros fijos y libres con la misma cantidad de factores y relaciones en cada muestra. Con lo

anterior se obtienen dos grupos de estimaciones para todos los parámetros del modelo.

- 2- Invarianza débil o invarianza métrica sin restricción: se ajustó el modelo en la muestra de calibración y las cargas factoriales estimadas se mantuvieron invariantes en el ajuste del modelo en la muestra de validación. Las varianzas de los errores de las variables observadas y las varianzas de las variables latentes se mantuvieron libres, es decir, se estimaron en ambas muestras. Con esta restricción se obtiene solamente un grupo de estimadores para las cargas factoriales de primer y segundo orden.
- 3- Invarianza débil o invarianza métrica con restricción: adicionalmente a la restricción de las cargas factoriales impuesta en el modelo anterior, también se impusieron restricciones en las varianzas de los errores. De este modo las estimaciones obtenidas de la muestra de calibración de cargas factoriales y varianzas se mantienen invariantes en la muestra de validación.
- 4- Invarianza fuerte o estructural: en este modelo se mantuvieron las restricciones sobre las cargas factoriales y se impusieron restricciones sobre las medias y los interceptos de las variables observadas. Esto quiere decir que las cargas factoriales, los interceptos y las medias se estimaron solamente en la muestra de calibración y se mantuvieron sin variación en la muestra de validación.
- 5- Invarianza estricta o completa: para este modelo se mantuvo la estructura del modelo de invarianza estructural y además se agregaron restricciones sobre las varianzas de los errores o residuos, es decir, todos los parámetros se estimaron solamente con los datos de la muestra de calibración y se replicaron en la muestra de validación.

Para realizar el contraste de hipótesis entre los modelos en cada nivel se utilizaron los siguientes criterios (Elosua, 2005; Byrne, 2008; Timmos, 2010; Barrera et al., 2015):

- 1- Invarianza configural: se analizaron los índices de ajuste conjunto χ^2 , RMSEA, CFI, NNFI y GFI. Esta prueba se utiliza para determinar si la estructura factorial es la misma entre los grupos.
- 2- Invarianza métrica: se compararon los indicadores χ^2 , RMSEA y CFI del modelo de invarianza métrica con los valores del modelo de invarianza configural. Si los valores de RMSEA de un modelo se incluyen dentro del intervalo de confianza del

otro modelo se confirma la invarianza métrica. Si el cambio en el CFI de ambos modelos es inferior a 0,01 también se confirma la invarianza métrica. Con esta prueba se determina si ambas muestras poseen la misma estructura factorial y los mismos parámetros de cargas factoriales.

- 3- Invarianza estructural: se aplicaron los mismos criterios utilizados en la prueba de invarianza métrica, pero esta vez comparando los cambios en los índices de ajuste RMSEA y CFI del modelo de invarianza métrica con el modelo de invarianza estructural. Un resultado positivo comprueba que ambas muestras poseen la misma estructura factorial, los mismos parámetros de cargas factoriales y los mismos interceptos y medias.
- 4- Invarianza completa: nuevamente se aplicaron los mismos criterios utilizados en la prueba de invarianza métrica, pero esta vez comparando los cambios en los índices de ajuste del modelo de invarianza completa con el modelo de invarianza estructural. Con esta prueba se demuestra el mayor nivel de invarianza, en el cual ambas muestras presentan igualdad en la estructura factorial y en todos los parámetros ajustados (cargas factoriales, interceptos, medias y covarianzas y varianzas de los errores).

Adicionalmente se aplicaron análisis de validación cruzada, en donde el modelo ajustado se replica en dos o más muestras y se aplican pruebas de significancia estadística, por medio del estadístico chi-cuadrado para las diferencias de los coeficientes estimados en los modelos de invarianza ajustados. Las muestras pueden extraerse de la misma población, de poblaciones diferentes o bien, se pueden utilizar submuestras de una misma muestra (enfoque de muestra dividida).

Con las pruebas de invarianza se analiza la capacidad del modelo de mantenerse invariante entre las muestras, lo que comprueba la estabilidad y la validez del modelo (Diamantopoulos & Siguaaw, 2000). Estas pruebas permiten determinar si los ítems del instrumento presentan un comportamiento diferencial en las subpoblaciones estudiadas, en cuanto a su estructura factorial (Abad et al., 2006).

El análisis de validación cruzada se aplicó en las distintas formas de replicación:

- 1- Replicación libre: el modelo se ajustó en cada muestra, esto quiere decir que todos los parámetros se dejaron libres en ambas muestras. Esta replicación equivale al modelo de invarianza configural.
- 2- Replicación moderada: el modelo se ajustó en el muestra de calibración y se replicó en la muestra de validación, pero dejando algunos parámetros libres. Este tipo de replicación es consistente con el modelo de invarianza métrica e invarianza estructural.
- 3- Replicación ajustada: el modelo ajustado en la muestra de calibración se replicó por completo en la muestra de validación, lo que corresponde al modelo de invarianza completa.

En las pruebas de invarianza aplicadas se contrastó la hipótesis nula (H_0) que los parámetros estimados son iguales en las muestras versus la hipótesis alternativa que las estimaciones de los parámetros son distintas (H_1).

Específicamente se comparó el modelo de invarianza configural (H_1) con los demás modelos de invarianza (H_0): modelo de invarianza métrica (validación débil), modelo de inavarianza estructural (validación moderada) y modelo de invarianza completa (validación fuerte). Para ello se utilizó la diferencia en el estadístico chi-cuadrado (X^2) global y la diferencia en los grados de libertad.

Este tipo de análisis permite apoyar o desaprobar la estructura factorial propuesta desde el modelo conceptual, por lo cual, es posible utilizarlo como evidencia de constructo (Abad et al., 2006).

4.13 Análisis de usuarios

Posterior al ajuste de los modelos estudiados, al cálculo de los índices propuestos y a la evaluación de la fiabilidad y validez de las mediciones, se aplicó un análisis de los usuarios de los cinco servicios evaluados según las puntuaciones obtenidas en el índice calculado, considerando los sujetos incluidos en la muestra de calibración.

Se aplicaron técnicas de análisis multivariantes para evaluar los puntajes del ICP en los diferentes servicios, realizando comparaciones de los usuarios por nivel educativo, por

género, por grupos de edad y por condición laboral, con el fin de establecer perfiles y segmentar a la población.

4.13.1 Segmentación de usuarios

La segmentación de los usuarios de los servicios se realizó aplicando un análisis de correspondencias, con el cual se establecieron asociaciones o correspondencias entre las características sociodemográficas de los usuarios y los niveles de calidad percibida, medidos por el ICP y los subíndices de dimensión.

El análisis de correspondencias es homólogo al análisis de componentes principales, pero aplicado a variables categóricas. Este análisis permite describir las relaciones entre las categorías de cada variable y entre las variables (Bartholomew, Steele, Moustaki & Galbraith, 2002). El análisis permitió identificar los segmentos de la población que están asociados a puntajes altos y bajos de calidad percibida.

Para aplicar el análisis de correspondencias se agrupó el puntaje del ICP y de los subíndices de dimensión en tres categorías o niveles: nivel bajo (puntajes inferiores a 71), nivel medio (puntajes entre 71 y 85) y nivel alto (puntajes mayores a 85). El análisis se aplicó al índice general (ICP) y a los subíndices de dimensión, de forma independiente.

Las variables control utilizadas en los análisis fueron: género, edad, nivel educativo y condición laboral, las cuales se categorizaron de la siguiente manera:

- Género: hombre-mujer
- Edad: jóvenes (18-29 años), adultos (30-49 años) y mayores (50 o más años).
- Nivel educativo: primaria o menos, secundaria y universitaria.
- Condición laboral: trabaja (trabaja, trabaja y estudia) y otro (estudia, rentista, ama de casa).

El análisis partió de la premisa de que las personas de más edad y menor nivel educativo, que no trabajan y principalmente mujeres, son propensos a brindar puntajes altos de calidad percibida; es decir, son más persuasivos al evaluar la calidad.

4.13.2 Perfiles de usuarios

Con el fin de establecer perfiles de usuarios que permitieran categorizar a la población de estudio, se aplicó un análisis de conglomerados, el cual permite distribuir los objetos de observación en grupos homogéneos, en función de la similitud entre ellos, de acuerdo a ciertas características observadas (Aldenferder & Blashfield, 1984).

El objetivo del análisis fue generar grupos homogéneos de usuarios de acuerdo a los puntajes en el ICP y en los subíndices de dimensión, según sus características sociodemográficas. Estos grupos se consideraron perfiles o tipologías de los usuarios de los servicios públicos según la percepción que poseen de la calidad de dichos servicios. El análisis permitió caracterizar a los usuarios que están asociados a puntajes altos y bajos de calidad percibida.

Específicamente, se aplicó un análisis de conglomerados de individuos mediante el método no jerárquico con el procedimiento k-medias, utilizando el coeficiente de correlación como medida de similitud. Como variables de agrupación se utilizaron el ICP y los subíndices de dimensión (de forma independiente), junto con la edad.

Para generar los perfiles de usuarios se utilizó el nivel educativo de los entrevistados, considerado como la cantidad de años de estudio, además de las variables de control: género y condición laboral. Al ser éstas variables cegógicas, se codificaron, asignándoles un valor numérico a cada categoría, lo que posibilita ser incluidas en el análisis. A continuación se muestran la codificación utilizada:

- Género: si es hombre se le asignó un 100 y si es mujer se le asignó un 0
- Condición laboral: si trabaja se le asignó un 100 y si tiene otra condición laboral se le asignó un 0.

La premisa propuesta consistió en que las personas que están asociadas a puntajes altos de calidad percibida poseen diferentes características que las personas asociadas a puntajes bajos. Se puso a prueba la hipótesis que los puntajes bajos en los índices están asociados con personas de menor edad, con altos niveles de educación, que trabajan y principalmente hombres; siendo más exigentes al evaluar la calidad en los servicios.

V Resultados

5.1 Selección de variables

En total, los cuestionarios analizados se compusieron de 144 preguntas distribuidas en 27 temas, de las cuales, 56 preguntas se encontraban presentes en los 5 instrumentos aplicados (anexo A). De éstas, se seleccionó un grupo inicial de 29 preguntas, más las preguntas de control, según el criterio de validez aparente, seleccionando las preguntas asociadas a los descriptores que definen las dimensiones del constructo central.

Del primer grupo de preguntas se eliminaron 4, las cuales presentaron más de 5% de valores faltantes. Finalmente se seleccionó un grupo de 25 preguntas (variables de estudio) que representaron las dimensiones del constructo (anexo A). Estas variables presentaron un porcentaje de valores faltantes por debajo del umbral definido: desde 0,2% hasta 4,7%.

Las 25 variables seleccionadas, utilizadas para ajustar los modelos propuestos, se distribuyeron en 7 temas, además se escogieron 4 variables de control (sociodemográficas) y una variable general sobre la calificación del servicio recibido (Cuadro A2, Anexo A).

5.2 Análisis de datos faltantes (no respuesta)

El análisis de datos faltantes o no respuesta, reveló que no hay diferencias en las distribuciones por género, edad y nivel educativo de las 25 variables observadas (anexo C), lo que indica que las personas que sí respondieron poseen las mismas características o muy similares a las personas que no respondieron.

La mayoría de los que respondieron y de los que no lo hicieron son hombres, concentrados en el grupo etario de 18 a 49 años (jóvenes y adultos) y con niveles educativos altos (educación universitaria). La cantidad de casos en las variables con patrones distintos a los mencionados es muy baja: 1,3% o menos.

Debido a los resultados obtenidos, se concluye que la no respuesta no se debe a características de los entrevistados, lo que posibilita la aplicación de técnicas de control de la no respuesta: eliminación de casos con datos faltantes o imputación de valores.

5.3 Tratamiento de datos faltantes

En la primera muestra de calibración (aplicación 2008-2009), solamente las variables control edad y educación presentaron casos con valores faltantes: 0,7% y 0,4% respectivamente. En la muestra de validación (aplicación 2009-2010) los valores faltantes aumentaron a 0,9% y 0,6% respectivamente. Además, en la muestra de validación, la variable *calificación del servicio* presentó 0,4% de casos con no respuesta.

Al ser baja (menor a 1%) la cantidad de casos con datos faltantes se aplicaron reglas de imputación sin que esto tuviera un efecto de distorsión en las distribuciones de las variables de control y de calificación imputadas. De esta forma, menos de 1% de los casos poseen valores imputados en dos variables.

A nivel general, la distribución de las variables control con imputación de casos no difiere de la distribución de las mismas variables sin casos imputados, no hay diferencias significativas al 5% a lo largo de las categorías (cuadro D1, anexo D). Estos resultados confirman que la imputación no produjo distorsión en las distribuciones de las variables tratadas.

De forma similar, la distribución de las variables observadas se mantiene sin variaciones después de aplicar la eliminación de casos, tampoco se presentaron diferencias significativas al 5% en las categorías de cada variable antes y después de aplicar la eliminación.

La distribución del tamaño de muestra por servicio y a nivel total, tampoco se vio afectada por la eliminación de casos con valores faltantes, no se presentaron diferencias significativas en los tamaños de muestra con y sin eliminación. Las muestras totales de las dos aplicaciones se redujeron en 23% y 26% (cuadro D2, anexo D).

5.4 Modelo RLM

5.4.1 Ecuación de regresión

La siguiente corresponde a la ecuación de regresión que resultó de ajustar el modelo de regresión lineal múltiple, utilizado para predecir la calificación de calidad percibida, a los datos de la muestra de calibración (aplicación 2008-2009).

$$ICC = 8,19 + 0,16C + 0,21R - 0,06A + 0,16E + 0,33T + \varepsilon \quad (\text{Ecuación 5.1})$$

donde ICC es índice de calificación de calidad, C es confiabilidad, R es responsabilidad, A es aseguramiento, E es empatía y T es tangibles y ε corresponde al término de error aleatorio o residuo (normalmente distribuido con media cero y variancia constante).

En la ecuación 5.1 sobresale que el coeficiente de la dimensión *aseguramiento* resultó negativo, cuando se esperaba que fuese positivo. El signo negativo indica que dicha dimensión le resta valor a la calificación de la calidad, lo que iría en contra del esquema conceptual (figura 4.1) y del modelo ajustado (figura 4.2). Esto es un indicio que la ecuación de regresión presenta algún problema.

5.4.2 Ajuste del modelo

Los estadísticos de bondad de ajuste del modelo de regresión se muestran en el cuadro 5.1. El coeficiente de determinación indica una baja explicación del modelo: las dimensiones, en conjunto, logran explicar apenas el 18% de la variabilidad del índice de calificación de calidad percibida ajustado. Se obtuvo un error de estimación elevado, el cual representa 1/5 de la escala de variación del índice (0 a 100).

En el análisis de varianza (ANDEVA) se obtuvo un estadístico F de Fisher significativo con $\alpha = 5\%$, lo que indica que las variables independientes del modelo (dimensiones), en conjunto, poseen una relación lineal con el ICC y que al menos una de ellas posee un efecto directo sobre dicho índice.

**Cuadro 5.1 Estadísticos de ajuste.
Modelo RLM. Muestra calibración.**

Estadístico	Valor
R	0,43
R ²	0,18
R ² ajustada	0,18
Error estándar de estimación	20,72
Estadístico F (ANDEVA)	46,75

Fuente: elaboración propia

El cuadro 5.2 muestra los resultados de las pruebas de significancia sobre las variables independientes (dimensiones), mediante el estadístico *t* de Student. Las variables independientes resultaron con un coeficiente significativamente diferente de cero, con $\alpha = 5\%$, a excepción de la dimensión *aseguramiento*. El resultado evidencia que la dimensión *aseguramiento* no está linealmente relacionada con el ICC, por lo que no produce un efecto sobre el índice.

Cuadro 5.2 Significancia de coeficientes de regresión. Muestra de calibración.

Variable	Coeficiente β	Error Estándar	Coeficiente β estandarizado	Estadístico <i>t</i>	Valor p	Límites de confianza 95%		IT	FIV
						Inferior	Superior		
Constante	8,19	4,82	0,00	1,70	0,09*	-0,01	0,18	0,00	0,00
Confiabilidad	0,16	0,04	0,11	3,65	0,00	0,07	0,24	0,88	1,14
Responsabilidad	0,21	0,04	0,18	5,55	0,00	0,14	0,29	0,78	1,29
Aseguramiento	-0,06	0,05	-0,05	-1,06	0,29*	-0,16	0,05	0,40	2,48
Empatía	0,16	0,05	0,13	2,96	0,00	0,05	0,26	0,39	2,55
Tangibles	0,33	0,04	0,23	7,50	0,00	0,25	0,42	0,83	1,21

*Coeficiente estadísticamente no significativo al 5 %

Siglas: IT = índice de tolerancia

FIV: factor de incremento de la varianza

Fuente: elaboración propia

Las correlaciones entre el ICC y los subíndices calculados con los datos observados estuvieron entre 0,20 y 0,39, por debajo del umbral definido (0,40), lo que indica una baja relación entre las variables independientes y la variable respuesta del modelo. Entre las variables independientes se presentaron correlaciones bajas, excepto en la correlación entre *aseguramiento* y *tangibles* (Cuadro E3, anexo E). Esta relación podría tener efecto en la significancia y dirección de coeficiente de *aseguramiento*.

Finalmente, la bondad de ajuste del modelo RLM se evaluó por medio de la comprobación de los siguientes supuestos del modelo (realización de diagnósticos):

- Linealidad: cuatro de las cinco dimensiones presentan coeficientes significativamente diferentes de cero, lo que indica que estas variables están linealmente relacionadas con el índice calculado, no así la dimensión que resultó con un coeficiente no significativo (dimensión *aseguramiento*).

Adicionalmente, el gráfico E1 (anexo E) muestra la dispersión entre los valores pronosticados o predichos y los residuos o errores de pronóstico. En el mismo se nota una tendencia decreciente y una dispersión concentrada en un rango entre 50 y 90. En el gráfico no se muestra con claridad la presencia de relación lineal entre las variables independientes y la variable dependiente. No se cumple el supuesto.

- Normalidad: el gráfico E2 del anexo E, muestra el gráfico P-P de los residuos estandarizados, en el cual se nota que la mayoría de valores se posicionan por encima de la recta de tendencia y en los extremos los puntos se posicionan debajo de la línea. Esto es muestra de falta de ajuste de los residuos a una distribución normal.

Lo anterior se confirma con las pruebas formales de normalidad aplicadas (Kolmogorov-Smirnov y Shapiro-Wilk). Las pruebas resultaron estadísticamente significativas al 5%, con estadísticos t de 0,08 y 0,94 respectivamente. Por tanto se descarta que los residuos se distribuyan normalmente.

- Media cero: la distribución de los residuos posee una media de 0 con una desviación estándar de 21. El rango de variación es de 128 (entre -85 y 43), sobrepasando la variación del índice (entre 0 y 100). Adicionalmente, los residuos presentan una asimetría negativa, según el coeficiente de asimetría (-1,06). Los datos indican que la media real de los residuos difiere de cero.

El incumplimiento del supuesto de normalidad refuerza estos resultados, pues al no tener una distribución normal, se descarta que los residuos posean media 0.

- Homoscedasticidad: en el gráfico E1 se nota una tendencia decreciente entre los valores pronosticados o predichos y los residuos; además, se observa una mayor cantidad de residuos con valores positivos que negativos, lo que da un indicio que las varianzas no son iguales.

La prueba formal mediante el estadístico de Levene confirma el análisis gráfico. Dicha prueba resultó significativa al 5% con un estadístico F de 4,84. Se rechaza la hipótesis nula de igualdad de varianzas, por lo que no se cumple con este supuesto.

- Autocorrelación: el gráfico E3 (anexo E) muestra la relación entre los errores de pronóstico (residuos) y el orden temporal en el cual fueron adquiridos los datos, con el fin de analizar la presencia/ausencia de autocorrelación. En él se no se nota ningún patrón que sugiera algún tipo de relación entre ambas variables.

Según el estadístico Durbin-Watson ($d = 1,92$ y $4-d = 2,08$) no hay evidencia de correlación serial entre los residuos. En ambos casos, el estadístico fue mayor al límite superior del valor crítico (1,90) obtenido para cinco variables independientes y un tamaño de muestra de 1.050 observaciones. Este supuesto sí se cumple.

- Independencia entre residuos y variables: el gráfico E4 (anexo E) muestra la relación de los residuos con cada una de las cinco dimensiones, utilizadas como variables independientes del modelo. No se logra apreciar una tendencia que indique la presencia de relación entre los residuos y las variables predictoras. Por su parte en el gráfico de residuos y valores pronosticados (gráfico E1) no se nota con claridad una relación lineal. Sí se cumple con este supuesto.
- Multicolinealidad: en el cuadro 5.2 se indican los valores del factor de incremento de la varianza (FIV) y del índice de tolerancia (IT), para las variables independientes. Todas las variables presentan valores del FIV inferiores a 5 y tolerancias mayores a 0,2 lo que significa que hay ausencia de colinealidad. Adicionalmente los errores estándar son bajos en relación a la escala de medida de las variables independientes. Se cumple con el supuesto

- **Tamaño de muestra:** no hay un consenso respecto al tamaño de muestra que se considera suficiente para ajustar modelos de regresión. Una regla general, consiste en multiplicar la cantidad de variables independientes (p) por un factor (f) que puede variar entre 5 y 20. Así el tamaño de muestra suficiente (n') es igual a f multiplicado por p . Para un modelo de cinco variables independientes ($p=5$), resultaría un tamaño de muestra suficiente (n') entre 25 y 100.

El modelo RLM se ajustó con un tamaño de muestra total mayor a 1.000 superando el tamaño muestral recomendado (100). Sucede lo mismo al considerar el tamaño de muestra por servicio, el cual es superior a 200, excepto en el servicio de la SETENA, cuyo tamaño de muestra fue de 91. Se cumple con el supuesto de tamaño de muestra suficiente.

El modelo RLM ajustado no logra cumplir con todos los supuestos puestos a prueba, por lo que el modelo carece de bondad de ajuste.

5.4.3 Índice de Calificación Calidad (ICC)

En el cuadro E1 (anexo E) se comparan los estadísticos descriptivos del ICC observado y pronosticado por el modelo de regresión. A nivel total, ambas mediciones poseen la misma media (74,24) y la variación de los datos observados es mayor a la de los datos pronosticados. En ambos casos, la distribución del índice es asimétrica negativa y platicúrtica (posee baja concentración de datos).

El comportamiento por servicio es diferente: los valores observados son mayores a los pronosticados en los servicios del RN, del MOPT y de la DGME, con diferencias de 1,8 3,1 y 7,5 puntos, respectivamente. En los servicios de la DGSC y de la SETENA, el índice medio pronosticado es mayor al observado, con diferencias de 1,7 y 12,5 puntos, respectivamente.

Según el indicador de curtosis, la distribución del índice observado se acerca a la normal para los servicios de la DGME, en los restantes servicios hay una baja concentración de los datos. La baja concentración de los datos se repite en los valores

pronosticados, en cuatro servicios, excepto para el servicio de la DGSC, cuya concentración de datos es alta.

En todos los servicios, los valores observados poseen una variación mayor a los valores pronosticados. Ambos valores (observado y pronosticado) poseen una asimetría negativa en todos los servicios. Los límites de confianza (al 95%) para la media de los valores observados se encuentran entre 70,1 y 76,0 para todos los servicios y entre 72,1 y 74,7 para valores pronosticados de todos los servicios.

5.4.4 Fiabilidad

El cuadro 5.3 muestra los coeficientes de fiabilidad de las puntuaciones del índice estimado, obtenidas con el modelo de RLM, para ambas aplicaciones del instrumento.

**Cuadro 5.3 Coeficientes de fiabilidad. Modelo RLM.
Ambas muestras.**

Técnica	Coeficiente de fiabilidad	Muestra	
		Calibración	Validación
Dos mitades	Spearman-Brown (r)	0,82	0,82
	Guttman-Flanagan (r)	0,82	0,82
	Alfa (mitad 1) (α)	0,49	0,54
	Alfa (mitad 2) (α)	0,53	0,56
Consistencia	Cronbach total (α)	0,73	0,75
	Ordinal (α)	0,72	0,75

Fuente: elaboración propia

Los coeficientes obtenidos con la técnica de dos mitades resultaron superiores al umbral de 0,60 e iguales con las correcciones Spearman-Brown y Guttman-Flanagan. El coeficiente alfa de cada mitad resultó por debajo del umbral, pero el coeficiente alfa de todo el instrumento y el coeficiente alfa ordinal, resultaron por encima de dicho umbral. El coeficiente alfa se ve disminuido si se elimina alguna de las variables del modelo (dimensiones). Este comportamiento se mantuvo en las dos muestras (calibración y evaluación).

Según los anteriores resultados, las puntuaciones del ICC son consistentes y estables. Las puntuaciones del índice obtenido mediante el modelo de regresión lineal múltiple poseen fiabilidad en su medida.

5.4.5 Validez

5.4.5.1 Evidencias de contenido

Las variables utilizadas en el modelo RLM (figura 4.2) corresponden a las dimensiones del modelo conceptual (figura 4.1), el cual se utilizó para realizar la propuesta del ICC.

De acuerdo con el modelo conceptual, las variables incluidas en el modelo son correctas y válidas.

El cuadro 5.4 muestra los resultados del análisis de factores exploratorio (AFE), utilizado para probar la dimensionalidad del ICC. Los resultados indican un único factor con un autovalor superior a 2, que acumula el 51% de la varianza explicada por dicho factor. Los demás autovalores son inferiores a 1. Las cargas factoriales de tres dimensiones superaron el umbral establecido (0,50), considerándose representantes de un mismo factor.

**Cuadro 5.4 Análisis de Factores Exploratorio. Solución rotada.
Modelo RLM. Muestra de validación.**

Dimensión	Cargas factoriales	Comunalidades		Factor	Autovalores	Varianza explicada
		Inicial	Extracción			
Confiabilidad	0,42	0,15	0,18	1	2,56	51,12
Responsabilidad	0,53	0,24	0,28	2	0,81	16,24
Aseguramiento	0,83	0,61	0,70	3	0,76	15,28
Empatía	0,86	0,62	0,74	4	0,65	12,91
Tangibles	0,44	0,18	0,20	5	0,22	4,46

Fuente: elaboración propia

Las dimensiones *aseguramiento* y *empatía*, fueron las más representativas o de mayor influencia sobre el factor único, el cual representa al ICC.

Las correlaciones entre las variables observadas y los subíndices de las dimensiones *confiabilidad* y *responsabilidad* (cuadro E2, anexo E) resultaron positivas, entre moderadas y altas (de 0,50 a 0,69). En las restantes dimensiones, las correlaciones entre cada variable y el subíndice al cual pertenecen indican una asociación lineal positiva y alta (superior al umbral de 0,60), en la muestra de validación.

Las correlaciones entre el índice y los subíndices de dimensión, en ambas muestras (cuadro E3, anexo E), resultaron entre bajas y moderadas (por debajo del umbral). La correlación más baja (0,43) se presentó con la dimensión *confiabilidad*. Entre dimensiones, las correlaciones resultaron bajas (menores a 0,35), excepto entre *empatía* y *aseguramiento*, superando el umbral de 0,60.

Es importante destacar que todas las correlaciones analizadas resultaron significativamente diferentes de cero, con $\alpha = 5\%$.

5.4.5.2 Evidencias de criterio

El cuadro 5.5 muestra los puntajes del índice de calificación de calidad y del índice de gestión institucional, para cada aplicación, total y por servicio, además incluye el promedio de los cinco servicios evaluados.

El promedio del ICC (2008 y 2009) es inferior al promedio del IGI (2010 y 2011). A nivel total ocurre lo contrario: el ICC es mayor al IGI en ambos años, pero las diferencias son pequeñas. Al comparar el ICC contra el factor servicio al usuario del IGI, las diferencias se hacen mayores, tanto a nivel promedio como en el total.

El RN obtuvo la calificación más alta en el ICC 2008 y 2009 y en el factor servicio al usuario del IGI 2010, la SETENA obtuvo la menor calificación con ambos índices, en los dos años y en el factor servicio al usuario.

El coeficiente de correlación indica una asociación lineal fuerte (mayor a 0,60) y directa entre el ICC de ambos años y la dimensión servicio al usuario del IGI 2010, el más cercano a la aplicación del ICC; no así con el IGI general. Por tanto, el ICC es consecuente con el factor servicio al usuario del IGI.

Cuadro 5.5 Puntaje del ICC y del IGI por servicio y coeficiente de validez, según año de aplicación. Modelo RLM. 2008 a 2011.

Servicio	Índice					
	ICC		IGI		SU	
	2008	2009	2010	2011	2010	2011
Media	73,42	71,05	75,37	78,09	58,44	75,86
RN	79,76	75,28	77,78	80,00	81,25	78,95
MOPT	71,73	65,21	77,48	86,05	43,75	78,95
DGME	74,54	75,45	70,93	77,40	54,69	73,68
DGSC	71,62	71,18	86,91	78,79	72,92	88,89
SETENA	69,43	68,15	63,77	68,21	39,58	58,82
Total	74,24	71,93	70,02	72,41	58,60	59,21
Coeficiente validez*						
ICC ₂₀₀₈	1,00	0,74	0,23	0,35	0,74	0,32
ICC ₂₀₀₉	0,74	1,00	0,09	-0,11	0,69	0,19

*Coeficiente de correlación de Pearson

Siglas: ICC = índice de calificación de calidad, IGI = índice de gestión institucional,

SU = factor servicio al usuario del IGI

Fuente: elaboración propia

El cuadro 5.6 muestra el porcentaje de inconformidades "no recibidas" y el porcentaje de inconformidades resueltas por dimensión, según los diagnósticos de la Secretaría Técnica del SNCS, así como los resultados del ICC total y por dimensión. Los datos se presentaron para las aplicaciones de los diagnósticos 2009 y 2010, coincidiendo con la aplicación de los instrumentos utilizados para el cálculo del ICC.

Cuadro 5.6 Puntajes de inconformidades externas e ICC estimado por dimensión. Modelo RLM. 2008 a 2010.

Dimensión	Diagnóstico				Dimensión	ICC	
	2009		2010			2008	2009
	no recibidas	resueltas	no recibidas	resueltas			
Total*	89,06	70,17	92,59	80,92	Total	74,24	71,93
Calidad	91,69	26,53	93,99	32,56	Confiabilidad	92,10	91,06
Información	88,36	27,55	90,11	31,10	Responsabilidad	86,52	82,69
Trato al usuario	81,65	20,41	91,86	20,73	Aseguramiento	86,51	89,68
Instalaciones	85,25	9,18	74,36	7,09	Empatía	86,04	87,67
Otras	91,16	16,33	95,97	8,51	Tangibles	73,29	69,08

*Se refiere al total para órganos adscritos a ministerios

Fuente: elaboración propia

La dimensión *información* del diagnóstico se comparó con la dimensión *confiabilidad* del ICC, la dimensión *trato al usuario* con la dimensión *empatía* y la dimensión *instalaciones* con la dimensión *tangibles*. En el análisis no se consideró la categoría "otras" (presente en los diagnósticos) por no detallar su contenido.

En la dimensión *calidad* de los diagnósticos, se obtuvieron valores altos en ambos años, mientras que en el ICC se obtuvieron valores más bajos, pero igualmente positivos. En los diagnósticos, las mayores calificaciones se concentraron en la dimensión *información* y la dimensión *instalaciones* concentró las calificaciones más bajas. Este comportamiento se replica en el ICC con las dimensiones *confiabilidad* y *tangibles*.

La congruencia de los resultados del ICC con los resultados del IGI y de los diagnósticos del SNCS, en donde se pueden llegar a conclusiones similares, se considera una evidencia de criterio a favor del ICC.

5.4.5.3 Evidencias de constructo

La principal evidencia de constructo se evaluó mediante el ajuste del AFC, que corresponde al segundo modelo propuesto en esta investigación para el cálculo del índice de calidad percibida.

Las evidencias de constructo del modelo RLM se limitaron al análisis de la fiabilidad compuesta (0,61) y de la varianza extraída (0,02) del ICC estimado. En ambos casos no se logró cumplir con el umbral propuesto: 0,70 y 0,50 respectivamente. Dados estos resultados, el ICC obtenido por medio del modelo de regresión lineal múltiple carece de evidencias de constructo.

En resumen, el ICC tiende a medir el concepto de calidad percibida en servicios con las variables correctas (evidencia de contenido), es congruente con otras mediciones externas del mismo concepto (evidencia de criterio), pero no logra realizar una medición de la estructura interna del concepto calidad percibida (evidencia de constructo).

5.5 Modelo AFC

5.5.1 Ecuaciones estructurales

A partir de la estimación de los parámetros (coeficientes) del modelo AFC se conformaron las ecuaciones estructurales, las cuales representan las relaciones entre las variables del modelo y miden sus efectos. Se ajustó una ecuación para cada variable observada endógena y una para cada variable latente endógena. Las ecuaciones resultantes se detallan en el cuadro 5.7.

En su mayoría, los efectos de las dimensiones, que representan las variables latentes del modelo, sobre las variables observadas son altos (mayores al umbral de 0,60). La dimensión con los efectos más bajos es *confiabilidad* (C) con dos variables con una carga factorial por debajo del umbral. El menor efecto del factor central (CP) se da sobre la dimensión *tangibles* (T), por debajo del umbral.

Cuadro 5.7 Ecuaciones estructurales estimadas.
Modelo AFC.

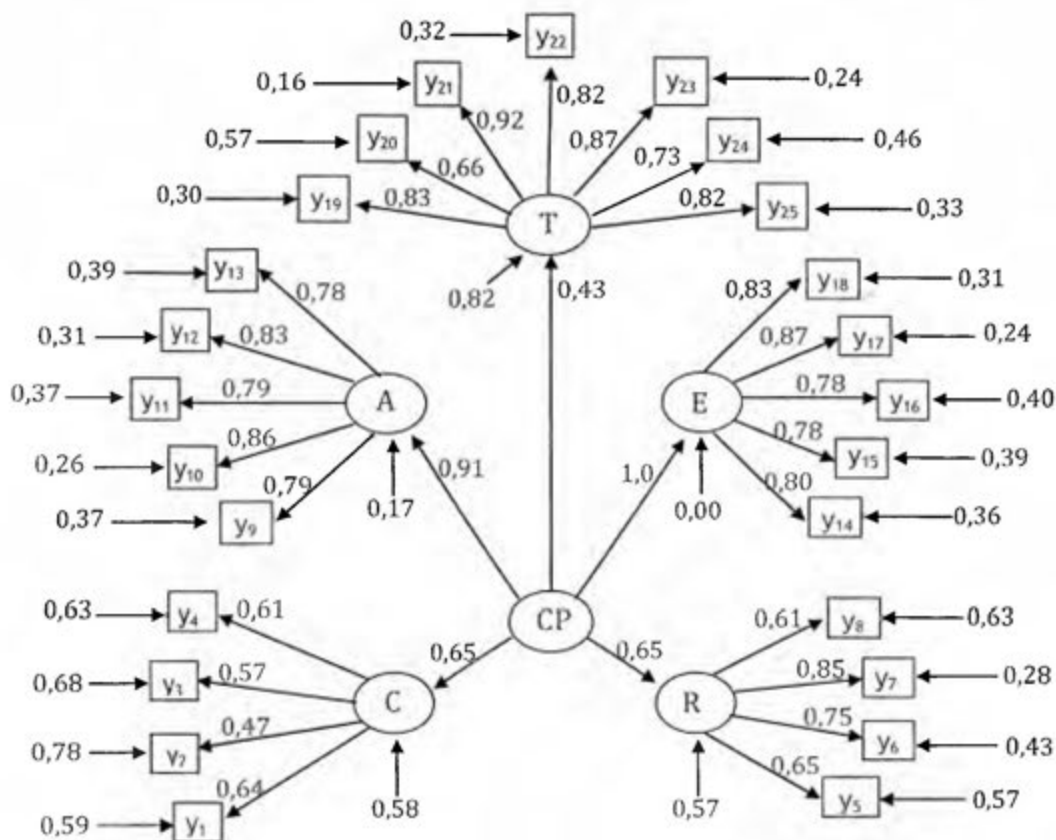
Ecuaciones de primer nivel	
$y_1 = 0,64 * C + 0,59$	$y_{14} = 0,80 * E + 0,36$
$y_2 = 0,47 * C + 0,78$	$y_{15} = 0,78 * E + 0,39$
$y_3 = 0,57 * C + 0,68$	$y_{16} = 0,78 * E + 0,40$
$y_4 = 0,61 * C + 0,63$	$y_{17} = 0,87 * E + 0,24$
$y_5 = 0,65 * R + 0,57$	$y_{18} = 0,83 * E + 0,31$
$y_6 = 0,75 * R + 0,43$	$y_{19} = 0,83 * T + 0,30$
$y_7 = 0,85 * R + 0,28$	$y_{20} = 0,66 * T + 0,57$
$y_8 = 0,61 * R + 0,63$	$y_{21} = 0,92 * T + 0,16$
$y_9 = 0,79 * A + 0,37$	$y_{22} = 0,82 * T + 0,32$
$y_{10} = 0,86 * A + 0,26$	$y_{23} = 0,87 * T + 0,24$
$y_{11} = 0,76 * A + 0,37$	$y_{24} = 0,73 * T + 0,46$
$y_{12} = 0,83 * A + 0,31$	$y_{25} = 0,82 * T + 0,33$
$y_{13} = 0,78 * A + 0,39$	
Ecuaciones de segundo nivel	
$\eta_1 = 0,65 * CP + 0,58$	$\eta_4 = 1,00 * CP + 0,00$
$\eta_2 = 0,65 * CP + 0,57$	$\eta_5 = 0,43 * CP + 0,82$
$\eta_3 = 0,91 * CP + 0,17$	

Fuente: Elaboración propia

La solución completamente estandarizada del modelo AFC se muestra en la figura 5.1. En esta solución se han estandarizado tanto las variables latentes como las variables

observadas, incluyendo los errores, lo que difiere de la solución estandarizada (simple) donde únicamente se estandarizan las variables latentes.

Figura 5.1 Modelo AFC de segundo orden.
Solución completamente estandarizada.



Fuente: elaboración propia

En la figura 5.1 se nota que el factor calidad percibida (CP) carga mayoritariamente (1,00) en la dimensión *empatía* (E), mientras la dimensión *tangibles* (T) es la que presenta la menor carga factorial de segundo orden (0,43), por debajo del umbral de 0,50. Por su parte, las dimensiones *confiabilidad* (C) y *responsabilidad* (R) poseen la misma carga factorial (0,65). Las cargas factoriales de segundo orden corresponden a la correlación entre el factor central (CP) y los factores subyacentes (dimensiones).

Las cargas factoriales de primer orden (λ), representan el efecto de las variables latentes sobre las variables observadas (primer nivel). La dimensión *confiabilidad* (C) es la

que produce los efectos de más baja intensidad, es decir, posee cargas factoriales bajas. Las cargas factoriales de todas las variables variaron entre 0,47 y 0,92. Solamente una carga factorial (variable y_2) estuvo por debajo del umbral de 0,50.

5.5.2 Ajuste del modelo AFC

5.5.2.1 Especificación

El modelo AFC se formuló siguiendo el esquema conceptual propuesto (figura 4.1), ajustado a las condiciones específicas del presente estudio y basado en el planteamiento de Parasuraman et al. (1985, 1988), plasmado en el Modelo Gaps (figura 2.2) y en la escala ServQual. La especificación del modelo AFC está respaldada por la teoría conceptual de la escala ServQual.

Se utilizó el planteamiento conceptual de Parasuraman et al. (1988) debido a que es el modelo que ha estudiado más exhaustivamente el concepto de calidad percibida en los servicios, así como su relación con otros conceptos. Adicionalmente, el cuestionario aplicado en la recolección de datos se construyó basado en este mismo planteamiento.

Según el modelo conceptual utilizado, la calidad percibida (factor general) está constituida de cinco dimensiones (factores específicos) cada una de ellas compuesta por una serie de variables medibles (descriptores), que describen características en cada una.

Las variables del modelo se seleccionaron bajo el criterio de validez aparente, se eligieron las variables que mejor representaron cada una de las cinco dimensiones del esquema conceptual (figura 4.1), de acuerdo con los descriptores de cada dimensión. Siguiendo el modelo conceptual, se propuso el modelo AFC (figura 4.3). Por lo anterior el modelo AFC cumple con el criterio de especificación.

5.5.2.2 Identificación

El modelo ajustado se compone de 60 coeficientes o parámetros: 25 cargas factoriales de primer orden asignadas a las variables observadas, junto con sus

respectivos errores y 5 cargas factoriales de segundo orden establecidas para las variables latentes, junto a sus errores respectivos.

Para ajustar el modelo se fijaron cinco cargas factoriales de primer orden en uno, con el fin de identificar la métrica de las variables latentes, quedando 55 parámetros libres a ser estimados en el modelo, a partir de las 25 variables observadas. Aplicando la ecuación 4.9 se obtuvo la cantidad de unidades únicas de información de la matriz de varianzas-covarianzas: $25(25+1)/2 = 325$.

Dado que la cantidad de piezas únicas de información de la matriz de varianzas-covarianzas ajustada con los datos observados (325) es mayor a la cantidad de parámetros libres estimados en el modelo (55), los parámetros del modelo resultaron sobre-identificados. Este resultado implica que el modelo planteado posee más de una solución.

5.5.2.3 Estimación

El cuadro F1 (anexo F) muestra las estimaciones ajustadas de los parámetros (coeficientes) junto con sus respectivos errores estándar y el valor del estadístico t , utilizado para evaluar su significancia. Además, incluye las estimaciones completamente estandarizadas y las correlaciones múltiples cuadradas, que representan las comunalidades.

Según el cuadro F1, las cargas factoriales de primer orden correspondientes a las variables y_2 , y_3 , y_4 , y_5 e y_6 (tabla A3, anexo A) resultaron no significativamente diferentes de cero, con un nivel de significancia de 5%, de acuerdo con el valor de su estadístico t , el cual resultó menor a 1,96.

Las primeras tres variables pertenecen a la dimensión *confiabilidad* mientras las dos restantes a la dimensión *responsabilidad*. Las demás variables del modelo resultaron con cargas factoriales significativamente diferentes de cero, al 5% de significancia.

Las variables con cargas factoriales no significativamente diferentes de cero indican que la dimensión a la cual están asociadas no ejerce efecto sobre ellas. No obstante,

estas variables y sus relaciones se mantienen en el modelo ajustado con el fin de respetar el planteamiento causal y teórico del cual proceden.

La carga factorial del factor general sobre la dimensión *confiabilidad* resultó no significativamente diferente de cero, al 5% ($t = 1,26$). Similar al caso anterior, esto significa que la *calidad percibida* no produce efecto sobre la dimensión *confiabilidad*. En contraste, el planteamiento conceptual utilizado indica que *confiabilidad* es una de las dimensiones relacionadas con el concepto de calidad, por tanto se mantiene en el modelo.

En ningún caso se obtuvo una estimación impropia (caso Heywood), esto es, no se presentaron coeficientes estimados estandarizados mayores a 1, todas las varianzas de error resultaron positivas, no se presentó ninguna correlación fuera de los límites establecidos (de -1 a 1) y todas las estimaciones resultaron en la dirección esperada (coeficientes positivos).

5.5.2.4 Bondad de ajuste

El análisis de los residuos, diferencia entre la matriz de covarianzas poblacional (Σ) y la matriz de covarianzas estimada ($\Sigma[\theta]$), se realizó por medio de varios índices que permitieron evaluar el ajuste de los datos al modelo. El valor de estos índices se muestra en el cuadro 5.8.

Los índices absolutos miden el ajuste global del modelo; es decir, miden las diferencias entre la matriz de varianzas-covarianzas observada y la matriz predicha por el modelo. De estos índices, no se cumplió con el umbral establecido para RMR y SRMR. La prueba mediante el estadístico chi-cuadrado (la más utilizada) resultó no significativa con un contraste al 5%, lo que indica que ambas matrices no difieren.

Los índices comparativos contrastan el modelo ajustado con el modelo de independencia (todos los parámetros estructurales son fijados en 0). En todos los casos se cumple con el umbral establecido. Los índices de parsimonia evalúan el modelo considerando el número de parámetros estimados. En este caso se compara contra el

modelo saturado (todas las posibles relaciones entre las variables). También se cumple con el umbral establecido para los índices de parsimonia.

Cuadro 5.8 Índices de ajuste estimados. Modelo AFC. Muestra de calibración.

Tipo	Nombre del índice	Siglas	Umbral	Valor estimado
Absoluto	Prueba con estadístico chi-cuadrado (X^2)	p	> 0,05	1,00
Absoluto	Razón chi-cuadrado / gl	X^2/df	< 3	0,11
Absoluto	Índice de bondad de ajuste	GFI	> 0,90	0,98
Absoluto	Raíz del residuo cuadrático medio	RMR	< 0,05	0,07*
Absoluto	Raíz del residuo cuadrático medio estandarizado	SRMR	< 0,09	0,13*
Absoluto	Raíz del error cuadrático medio de aproximación	RMSEA	< 0,05	0,00
Comparativo	Índice de ajuste comparativo	CFI	> 0,95	1,00
Comparativo	Índice de ajuste normado	NFI	> 0,90	0,96
Comparativo	Índice de ajuste no normado	NNFI	> 0,95	1,69
Comparativo	Índice de ajuste incremental	IFI	> 0,95	1,58
Comparativo	Índice de ajuste relativo	RFI	> 0,95	0,95
Parsimonia	Índice parsimonioso de ajuste normado	PNFI	> 0,80	0,86
Parsimonia	Índice parsimonioso de bondad de ajuste	PGFI	> 0,80	0,82
Parsimonia	Índice ajustado de bondad de ajuste	AGFI	> 0,90	0,98

*Valores mayores al umbral

Fuente: elaboración propia

Aplicando el criterio propuesto por Diamantopoulos & Siguaw (2000), se logra cumplir con el umbral establecido para cinco de los seis criterios indicados. Al cumplirse con los criterios de esta regla y con los criterios de los demás índices de ajuste evaluados, los datos se ajustan al modelo planteado y se cumple con el criterio de ajuste.

5.5.2.5 Supuestos

- Linealidad: el modelo propuesto consiste en un modelo recursivo (relaciones directas en una vía) en el cual se exponen las relaciones lineales entre las dimensiones y las variables observadas (primer nivel) así como las relaciones lineales entre el constructo general y las dimensiones (segundo nivel). Desde la especificación del modelo se cumple con la condición de linealidad.

La significancia de los coeficientes viene a reforzar el establecimiento de las relaciones lineales entre variables, lo que permitió establecer las ecuaciones estructurales del modelo ajustado.

Adicionalmente, en el anexo F, se incluyen los gráficos de dispersión entre el índice calculado y las variables observadas (Gráfico F1), entre los subíndices de dimensión y las variables observadas (Gráfico F2) y entre los subíndices de dimensión y el índice calculado (Gráfico F2). Debido a la gran cantidad de variables y combinaciones posibles, en el análisis se utilizó una variable observada por cada variable latente, la que mejor representaba a cada dimensión.

A pesar de mostrarse aumentos discretos debido a que se utilizaron variables ordinales, en todos los gráficos se nota una relación directa y creciente entre pares de variables e índices, coincidiendo con el modelo teórico. Los gráficos respaldan la presencia de relaciones lineales entre variables observadas y variables latentes.

- Ausencia de colinealidad: en el cuadro F2 (anexo F), se muestran las correlaciones policóricas entre pares de variables observadas, la cuales varían entre 0,01 y 0,81. Las correlaciones más bajas se presentaron entre las variables de las dimensiones *confiabilidad* y *responsabilidad*, en las demás dimensiones las correlaciones superaron el umbral de 0,60 y resultaron significativas al 5%.

Debido a que se pudo obtener una estimación para cada variable observada, tanto de su efecto como de su error, se entiende que la matriz es no singular, lo que su vez refleja que las variables no son una combinación lineal de otras variables, es decir, no son multicolineales. Por tanto, se cumple con este supuesto.

- Tamaño de muestra: aplicando los criterios y reglas heurísticas mencionadas en la bibliografía consultada (Diamantopoulos & Sigauw, 2000; Cea, 2002; Catena et al., 2003; Orgaz, 2008) para un modelo con 25 variables observadas y 5 variables latentes se obtiene un tamaño de muestra de entre 125 y 375 individuos.

En el presente caso, el modelo se ajustó con un tamaño de muestra total mayor a 1.000 sujetos, superando el resultado de las reglas heurísticas y las recomendaciones de los autores en casos de modelos complejos con muchas variables, por lo que se cumple con el supuesto de tamaño de muestra.

- **Aditividad de los efectos:** haber obtenido una solución del modelo; es decir, obtener estimaciones de los parámetros según el modelo definido, es muestra que los efectos de las variables pueden sumarse. Esta condición se ve reflejada en las ecuaciones estructurales estimadas (cuadro 5.7). Se cumple con este supuesto.
- **Indicadores múltiples:** los autores consultados (Bollen, 1989; Diamantopoulos & Siguaw, 2000; Kaplan, 2009; Mulaik, 2009) concuerdan en la necesidad de utilizar más de una variable observada asociada a cada variable latente, por eso proponen utilizan entre 3 y 6 variables observadas por cada variable latente.

El modelo AFC se ajustó con las variables observadas, sin ser agregadas en indicadores, lo que permitió contar con 4, 5 y hasta 7 variables observadas por cada variable latente, cumpliendo con las recomendaciones de los autores mencionados. Se cumple con la condición de disponer de indicadores múltiples para las variables latentes.

- **Variables métricas:** aunque es una condición en los modelos AFC, el uso de variables métricas raramente se cumple, sobre todo en ciencias sociales, donde lo más común es el uso de variables ordinales medidas con escalas tipo Likert (Bollen, 1989; Ruiz et al., 2010). Las variables utilizadas en el modelo ajustado correspondieron a variables ordinales y nominales, por lo que no se cumple con esta condición.

El problema de este incumplimiento es que se viola la hipótesis de estructura de covarianzas poblacionales, la cual no se mantiene para la matriz de covarianzas ajustada con variables ordinales. Esta condición produce que los datos no ajusten al modelo.

Para tratar este problema se utilizó la matriz de correlaciones policóricas, la cual es un estimador consistente de la matriz de covarianzas de las variables continuas subyacentes. Esta última se utilizó en lugar de la matriz de covarianzas poblacional, ya que ésta sí mantiene la hipótesis sobre la estructura de covarianzas poblacionales (Bollen, 1989).

Con esta medida se corrigió el problema de utilizar variables ordinales, permitiendo obtener estimaciones válidas para los parámetros y logrando un buen ajuste del modelo. El incumplimiento de esta condición no afectó los resultados del modelo.

5.5.3 Índice de Calidad Percibida (ICP)

El cuadro F3 (anexo F) muestra los estadísticos descriptivos del índice global ICP y de los subíndices de dimensión, por servicio evaluado y a nivel total, calculados con datos de la muestra de calibración. El menor puntaje del índice global fue de 13 puntos mientras que en los subíndices fue de 0, el máximo en todos los índices fue 100 puntos. Lo anterior indica que ningún individuo calificó con 0 a todas las dimensiones, a la vez.

Al comparar por servicio evaluado, el servicio brindado por el RN es el que obtuvo el mejor puntaje promedio de calidad percibida (87,20), mientras que el servicio de la SETENA fue el que obtuvo el menor puntaje promedio (75,23). La diferencia entre ambas calificaciones extremas fue de 12 puntos.

En la muestra se presentaron calificaciones máximas de 100 puntos, en el ICP, correspondientes a los servicios del RN y de la DGME. En los restantes servicios, las calificaciones máximas fueron de 98 y 99 puntos. Los servicios de la DGSC y de la DGME presentaron las calificaciones más bajas (13 puntos) en el ICP. Las 5 instituciones obtuvieron calificaciones de 0 en al menos una dimensión.

Los puntajes más frecuentes en los subíndices fueron de 100 puntos, tanto a nivel total como por servicio. Los puntajes promedio más bajos se presentaron en el subíndice de la dimensión *tangibles*, entre 68 y 84 puntos, en los servicios brindados por el MOPT, la DGME, la DGSC y la SETENA. La dimensión *confiabilidad* fue la mejor evaluada en los 5 servicios estudiados, con los puntajes más altos (de 81 a 96 puntos).

Tanto la distribución del índice general como de los subíndices, presentan asimetrías negativas, a nivel total y por servicio. La dimensión *tangibles* es la que presenta el menor grado asimetría entre los distintos servicios evaluados.

5.5.4 Fiabilidad

El cuadro 5.9 contiene los coeficientes de fiabilidad calculados a partir de las variables incluidas en el modelo AFC; obtenidos para el ICP, en ambas muestras. Todos los coeficientes de fiabilidad calculados superaron el umbral establecido (0,60) tanto con la técnica de dos mitades como con la evaluación de consistencia interna.

Cuadro 5.9. Coeficientes de fiabilidad. Modelo AFC.

Técnica	Coeficiente de fiabilidad	Muestra	
		Calibración	Validación
Dos mitades	Spearman-Brown (r)	0,94	0,93
	Guttman-Flanagan (r)	0,93	0,93
	Alfa (mitad 1) (α)	0,79	0,79
	Alfa (mitad 2) (α)	0,76	0,78
Consistencia	Cronbach (α)	0,88	0,89
	Ordinal (α)	0,88	0,89

Fuente: elaboración propia

El coeficiente Alfa de todo el instrumentó se mantiene igual o disminuye si se elimina alguna de las variables observadas, lo que indica que son consistentes con el índice. Se mantuvo el mismo comportamiento en ambas muestras evaluadas.

De acuerdo con los índices de fiabilidad, las puntuaciones del ICP son consistentes y estables, por lo que el índice general y los subíndices de dimensión, obtenidos con el modelo AFC, poseen fiabilidad.

5.5.5 Validez

5.5.5.1 Evidencias de contenido

La selección de variables utilizadas para ajustar el modelo AFC se sustentó en una teoría sobre medición de calidad percibida en servicios (modelo Gaps y ServQual) por medio de la cual se estableció un esquema conceptual (figura 4.1). Este esquema se utilizó para proponer el modelo a ajustar (figura 4.3) con el cual se calculó el índice de calidad percibida propuesto. Las variables utilizadas son correctas y válidas.

Se compararon las variables utilizadas en el modelo ajustado con las características que describen cada dimensión del esquema conceptual. En la tabla F1 del anexo F se nota que las variables del modelo ajustado logran representar las dimensiones del esquema conceptual. Este análisis suma evidencias a favor del contenido del índice.

El cuadro F4 (anexo F) muestra la solución rotada del análisis de factores exploratorio, ajustado con las variables observadas en la muestra de validación, utilizado para evaluar la dimensionalidad del ICP. Resultaron seis factores con un autovalor mayor a 1, los cuales logran explicar el 64% de la varianza del factor subyacente.

Un mismo factor se compone de las variables de las dimensiones *aseguramiento* y *empatía*, las variables de las dimensiones *responsabilidad* y *confiabilidad* se distribuyen en tres factores. Las variables de la dimensión *tangibles* cargan en un mismo factor. Dos variables de la dimensión *aseguramiento* cargan en dos factores diferentes y dos variables de la dimensión *confiabilidad* poseen cargas inferiores al umbral (0,50).

Al restringir la selección a cinco factores se logra acumular el 60% de varianza explicada. Un mismo factor agrupa las dimensiones *aseguramiento* y *empatía*, un factor representa la dimensión *tangibles*, un factor incluye el tema "instrucciones recibidas" (dimensión *responsabilidad*) y otro el tema "información solicitada" (dimensión *confiabilidad*).

Las variables de los temas "atrasos en el servicio" (dimensión *confiabilidad*) y "tiempos" (dimensión *responsabilidad*) poseen cargas factoriales inferiores al umbral (0,50), pero cargan mayoritariamente en un mismo factor, con cargas entre 0,36 y 0,47.

El cuadro 5.10 muestra el resultado del AFE ajustado con los subíndices de dimensión. Resultó un único factor que acumula el 51% de la varianza explicada y un autovalor superior a 2. Los otros autovalores varían entre 0,8 y 0,2. Solamente dos subíndices (dimensiones *aseguramiento* y *empatía*) presentaron cargas factoriales por encima del umbral establecido (0,50).

Cuadro 5.10 Análisis de Factores Exploratorio. Solución rotada. Muestra de validación.

Dimensión	Cargas factoriales	Comunalidades		Factor	Autovalor	Varianza explicada
		Inicial	Extracción			
Confiabilidad	0,43	0,16	0,18	1	2,53	50,59
Responsabilidad	0,49	0,21	0,24	2	0,82	16,38
Aseguramiento	0,82	0,61	0,68	3	0,77	15,47
Empatía	0,87	0,63	0,76	4	0,66	13,12
Tangibles	0,44	0,18	0,19	5	0,22	4,45

Fuente: elaboración propia

Según los resultados del AFE, el primer nivel del modelo AFC se compone de 5 factores, pero con una estructura distinta de variables a la planteada en el modelo ajustado (figura 4.3). El segundo nivel presenta un único factor conformado por las dimensiones del modelo, lo que confirma la unidimensionalidad del índice.

Finalmente, se analizaron las correlaciones entre cada variable y el índice global, así como entre cada variable y el subíndice al cual pertenecen. Se presentaron correlaciones altas (mayores al umbral de 0,60) y positivas entre variables observadas y subíndices. Dichas correlaciones resultaron significativas al 5% (cuadro F5, anexo F).

Las correlaciones entre las variables y el ICP se muestran en el cuadro F6 (anexo F), las cuales estuvieron entre 0,52 y 0,82 (moderadas y altas). Las variables que pertenecen a las dimensiones *confiabilidad* y *responsabilidad* son las que presentan mayores correlaciones lineales con el índice global.

Los resultados analizados indican que las variables utilizadas para calcular los subíndices están relacionadas con dichos subíndices, por lo que representan las dimensiones. Así mismo, las variables utilizadas en el modelo AFC están relacionadas linealmente con el índice global, por lo que son una representación de este concepto. Estos resultados representan evidencias de contenido del índice calculado.

5.5.5.2 Evidencias de criterio

Los resultados del índice de calidad percibida y del índice de gestión institucional se exponen en el cuadro 5.11. La relación más importante se da entre el ICP y el factor

servicio al usuario del IGI 2010, por ser el indicador que más se ajusta al concepto que mide el ICP y por ser la medición más cercana a la aplicación del ICP (2008 y 2009).

Cuadro 5.11 Puntaje del ICP y del IGI por servicio y coeficiente de validez, según año de aplicación. Modelo AFC. 2008 a 2011.

Servicio	ICP		Índice			
	2008	2009	IGI		SU	
			2010	2011	2010	2011
Media	80,81	80,34	75,37	78,09	58,44	75,86
RN	87,20	84,27	77,78	80,00	81,25	78,95
MOPT	78,39	71,30	77,48	86,05	43,75	78,95
DGME	83,21	87,17	70,93	77,40	54,69	73,68
DGSC	80,02	82,06	86,91	78,79	72,92	88,89
SETENA	75,23	76,91	63,77	68,21	39,58	58,82
Total	81,87	81,40	70,02	72,41	58,60	59,21
Coeficiente validez*						
ICC ₂₀₀₈	1,00	0,70	0,35	0,39	0,80	0,45
ICC ₂₀₀₉	0,70	1,00	0,09	-0,21	0,63	0,18

*Coeficiente de correlación de Pearson

Siglas: ICP = índice de calidad percibida, IGI = índice de gestión institucional, SU = factor servicio al usuario del IGI

Fuente: elaboración propia

El promedio de los subíndices de las cinco instituciones evaluadas es mayor en el ICP que en el IGI y que en el factor servicio al usuario, con diferencias entre 2 y 5 puntos. Considerando el valor total del IGI (calculado para todas las instituciones) y del factor servicio al usuario, las diferencias con respecto al ICP total se hacen mayores, a favor del ICP (entre 9 y 23 puntos).

Sobre los servicios, la SETENA obtuvo la calificación más baja en ambas mediciones del IGI y del factor servicio al usuario, lo mismo que en el ICP 2008. El RN obtuvo la calificación más alta en el ICP 2008 y en el factor servicio al usuario del IGI 2010.

El coeficiente de correlación, utilizado como coeficiente de validez, indica una asociación lineal fuerte (mayor a 0,60), directa y significativa entre el ICP de ambas aplicaciones y la dimensión servicio al usuario del IGI 2010, el más cercano a la aplicación del ICP. Los resultados indican que el ICP y el IGI presentan un comportamiento similar a lo largo de los servicios evaluados, así como un coeficiente de validez alto.

Al relacionar el ICP con los diagnósticos realizados por la Secretaría Técnica del SNCS, se obtienen resultados similares (cuadro 5.12). Con ambas mediciones se alcanzan calificaciones favorables a nivel total, mayores a 70. En la dimensión "calidad" se obtuvo calificaciones altas en los diagnósticos, lo que se equipara a las calificaciones favorables del ICP.

**Cuadro 5.12 Inconformidades externas e ICP estimado por dimensión.
Modelo AFC. 2008 a 2010.**

Dimensión	Diagnóstico				Dimensión	ICP	
	2009		2010			2008	2009
	no recibidas	resueltas	no recibidas	resueltas			
Total*	89,06	70,17	92,59	80,92	Total	81,87	81,40
Calidad	91,69	26,53	93,99	32,56	Confiabilidad	92,27	91,14
Información	88,36	27,55	90,11	31,10	Responsabilidad	83,69	79,52
Trato al usuario	81,65	20,41	91,86	20,73	Aseguramiento	86,49	89,66
Instalaciones	85,25	9,18	74,36	7,09	Empatía	86,03	87,65
Otras	91,16	16,33	95,97	8,51	Tangibles	73,45	69,27

*Se refiere al total para órganos adscritos a ministerios

Fuente: elaboración propia

Las dimensiones "información" (diagnósticos) y *confiabilidad* (ICP) miden aspectos similares y ambas resultaron las mejores evaluadas. Por el contrario, las dimensiones en las que predominaron las calificaciones más bajas fueron "instalaciones" (diagnósticos) y *tangibles* (ICP). Estas dos dimensiones, también miden aspectos similares.

Las evidencias de criterio se reflejan en que los resultados del ICP son congruentes con los resultados del IGI y de los diagnósticos; hay una correspondencia en el resultado de los criterios y se pueden llegar a conclusiones similares.

5.5.5.3 Evidencias de constructo

A excepción de la variable y_2 de la dimensión *confiabilidad*, las cargas factoriales completamente estandarizadas de las variables observadas (primer orden) resultaron mayores al umbral (0,50). Solamente la carga factorial correspondiente a la variable latente *tangibles* (segundo orden) resultó menor al umbral establecido, pero significativamente diferente de cero (con un nivel de significancia de 5%).

La correlación múltiple cuadrada (fiabilidad de las variables) resultó mayor al umbral de 0,60 para las variables asociadas a las dimensiones *aseguramiento*, *empatía* y *tangibles* (cuadro F1, anexo F). Adicionalmente, el modelo AFC posee bondad de ajuste, al cumplirse con los umbrales establecidos para los distintos índices de ajuste evaluados (cuadro 5.8).

El cuadro 5.13 muestra los resultados de los coeficientes de validez utilizados. Solamente en la dimensión *confiabilidad* se obtuvo una fiabilidad compuesta y una varianza extraída menor a los umbrales definidos: 0,70 y 0,50, respectivamente. En las restantes dimensiones y en el índice calculado se logró cumplir con dichos umbrales, recolectando evidencias de validez a favor del ICP.

**Cuadro 5.13 Coeficientes de validez.
Modelo AFC. Muestra de calibración**

Dimensión	FC	S²
Confiabilidad	0,66	0,33
Responsabilidad	0,81	0,52
Aseguramiento	0,91	0,66
Empatía	0,91	0,66
Tangibles	0,93	0,66
ICP	0,97	0,59
Dimensiones	0,86	0,57

FC = fiabilidad compuesta

S² = varianza extraída

Fuente: elaboración propia

La fiabilidad compuesta y la varianza extraída, calculadas con los subíndices de las variables latentes (dimensiones) también superan los umbrales (0,70 y 0,50). De acuerdo con estos resultados, los subíndices de dimensión y el índice general poseen evidencias de constructo.

5.6 Invarianza factorial

El cuadro 5.14 muestra los resultados de los modelos de invarianza configural, métrica sin restricción y métrica con restricción. Los modelos de invarianza estructural e invarianza completa resultaron sin ajuste, por lo que el alcance del análisis de invarianza llega hasta el análisis de las cargas factoriales y las varianzas de error.

Cuadro 5.14 Análisis de invarianza factorial. Modelo AFC.

Índice	Umbral	Configural	Métrica sin restricción	Métrica con restricción
X^2	---	52,25	65,74	70,05
gl	---	561	585	595
valor p de X^2	> 0,05	1,00	1,00	1,00
RMSEA	< 0,05	0,00	0,00	0,00
Li RMSEA	---	0,00	0,00	0,00
Ls RMSEA	---	0,00	0,00	0,00
CFI	> 0,95	1,00	1,00	1,00
Δ CFI	< 0,01	---	0,00	0,00
NNFI	> 0,95	2,04	2,02	2,01
GFI	> 0,90	0,98	0,98	0,98
GFI	> 0,91	0,99	0,99	0,99
AIC	menor valor	230,25	195,74	180,05

X^2 = estadístico chi-cuadrado

gl = grados de libertad

Li = límite inferior / Ls = límite superior (límites al 90% de confianza)

Fuente: elaboración propia

Se cumple con los umbrales de los índices de ajuste y se cumple con los criterios de evaluación definidos: el valor de RMSEA de un modelo se incluye en el intervalo de confianza del otro modelo y el cambio en CFI entre dos modelos es menor a 0,01. Se demuestra que el modelo AFC posee invarianza métrica. El modelo AFC mantiene la misma estructura factorial, las mismas cargas factoriales y errores entre las dos muestras.

El cuadro F7 (anexo F) muestra los resultados del modelo de validación cruzada ajustado con datos de las dos muestras (calibración y validación). Los límites de confianza (al 95%) de los parámetros estimados en ambas muestras se traslapan, este comportamiento da un indicio de que dichos coeficientes no difieren.

El resultado de las pruebas de invarianza para el análisis de validación cruzada se expone en el cuadro 5.15. Ambas pruebas resultaron estadísticamente no significativas con $\alpha = 5\%$ y un valor p cercano a 1. En ambos casos no se rechaza la hipótesis de igualdad de parámetros de cargas factoriales y de varianzas de error.

Cuadro 5.15 Análisis de validación cruzada. Modelo AFC.

Prueba 1. Cargas Factoriales				
Hipótesis	χ^2	gl	Valor p	ECVI
H ₁ . Invarianza configural	52,25	561	1,00	0,36
H ₀ . Invarianza métrica sin restricción	65,74	585	1,00	0,34
Diferencia de parámetros	13,49	24	0,96	
Prueba 2. Cargas Factoriales y varianzas de error				
Hipótesis	χ^2	gl	Valor p	ECVI
H ₁ . Invarianza configural	52,25	561	1,00	0,36
H ₀ . Invarianza métrica con restricción	70,05	595	1,00	0,34
Diferencia de parámetros	17,8	34	0,99	

Fuente: elaboración propia

En ambas pruebas el Índice de Validación Cruzada Esperado (ECVI por sus siglas en inglés) resultó menor en el modelo ajustado bajo la hipótesis nula, lo que sugiere la aceptación de dicha hipótesis.

Los resultados de las pruebas de invarianza sugieren que las cargas factoriales y las varianzas de error se mantienen invariantes entre las muestras, es decir, los parámetros ajustados en ambas muestras no son significativamente diferentes. Dado que los modelos de invarianza estructural y completa resultaron sin ajuste, no se puede demostrar invarianza a nivel de las medias e interceptos, por lo que el análisis se limita a una validación débil.

Los resultados obtenidos indican que la estructura factorial del modelo ajustado (figura 4.3) se logró mantener en ambas muestras (calibración y validación). Al no haber un comportamiento diferencial en las cargas factoriales, se valida la estructura propuesta en el esquema conceptual (figura 4.1). Este resultado, también constituye una evidencia de constructo del índice ajustado (ICP).

5.7 Análisis de usuarios

El cuadro G1 (anexo G) muestra la distribución de los puntajes del índice de calidad percibida y de los subíndices de dimensión, según los tres niveles definidos (bajo, medio y alto). Con excepción de la dimensión *tangibles*, los subíndices y el índice general se concentran en puntajes altos (mayores a 85).

Sobre las características de los usuarios estudiados (cuadro G2, anexo G), 6 de cada 10 son hombres, mayoritariamente jóvenes y adultos (menores a 50 años), 8 de cada 10 trabajan y 6 de cada 10 poseen altos niveles de escolaridad (educación universitaria).

De acuerdo a los análisis aplicados, se encontró que los puntajes del índice general y de los subíndices se encuentran asociados a la edad y al nivel educativo, principalmente. Las personas de mayor edad y menor educación formal están asociadas a puntajes altos en los índices. Esto demuestra que las personas jóvenes y con mayor nivel educativo son más estrictas para calificar la calidad percibida en los servicios.

Los análisis aplicados a los usuarios abarcaron únicamente al índice obtenido por medio del análisis de factores confirmatorio (AFC), por ser el modelo que superó las pruebas de ajuste y que cumple con las condiciones para medir la calidad percibida en servicios públicos.

5.7.1 Segmentación de usuarios

El cuadro G3 (anexo G) expone los resultados del análisis de correspondencias del índice general y de los subíndices de dimensión con las variables edad y educación. Al realizar el análisis con las variables: género y ocupación, resultó una sola dimensión, lo que indica que no hay diferencias en los puntajes entre las categorías de dichas variables. Por tanto estas variables no se consideraron en la segmentación de los usuarios.

Los puntajes de los subíndices *confiabilidad*, *aseguramiento* y *empatía* no dependen de los grupos etarios, el puntaje del subíndice *aseguramiento* es independiente del nivel educativo (cuadro G3). Todas las asociaciones son de tipo bidimensional (2 dimensiones) y en todos los casos, la dimensión 1 logra explicar la mayoría de la inercia (asociación) entre las categorías de las variables edad y educación.

Los gráficos de asociación dimensional (gráficos G1 a G12) del anexo G, reflejan el comportamiento esperado: los más jóvenes (menores de 30 años) y los de mayor nivel educativo (universitario) están asociados a puntajes intermedios y bajos (85 o menos) en el índice general y en los subíndices de dimensión.

Distribuciones particulares se presentaron en el caso del subíndice de *confiabilidad*, donde los jóvenes están asociados a los puntajes intermedios, pero las personas de mayor edad, están asociadas a los puntajes más altos (gráfico G3). En este caso se mantuvo el comportamiento general de las otras dimensiones.

En el subíndice de *aseguramiento*, las calificaciones intermedias y bajas están asociadas a los que poseen educación universitaria (gráfico G8). En el subíndice de *tangible* los usuarios con secundaria están asociados a los puntajes más altos y los puntajes intermedios están asociados a los usuarios con nivel de primaria (gráfico G12). Este comportamiento también es consistente con el esperado: a mayor educación y menor edad, menor es el puntaje en los índices.

Los resultados obtenidos en los análisis de segmentación indican que hay una relación directa entre la edad y el índice general y los subíndices de dimensión y una relación inversa entre el nivel educativo con respecto al índice y a los subíndices.

5.7.2 Perfiles de usuarios

El cuadro 5.16 muestra el resultado del análisis de conglomerados para el índice general y para los subíndices de dimensión. Resultaron tres conglomerados (grupos de usuarios) con un puntaje promedio entre 80 y 86 del ICP y con puntajes medios en los subíndices de dimensión entre 72 y 95.

Solamente en el subíndice de la dimensión *confiabilidad* no se presentaron diferencias significativas (con $\alpha = 5\%$) en las medias de los tres grupos. Tampoco se presentaron diferencias significativas por género entre los tres grupos, por lo que esta variable no se consideró en la generación de los perfiles.

A partir de las variables edad, nivel educativo y condición laboral (mayor nivel de condición laboral indica que trabaja en mayor medida), se generaron tres perfiles de usuarios de servicios públicos (correspondientes a los tres conglomerados). Todos los perfiles estuvieron asociados a puntajes medios y altos (mayores a 70).

Cuadro 5.16 Perfiles de usuarios. Muestra de calibración.

Variable	Media de conglomerado			Total	Valor p	Eta
	1	2	3			
ICP	80,3	82,7	86,4	81,9	0,00	0,15
Confiabilidad	91,6	92,5	94,7	92,3	0,14	0,06
Responsabilidad	81,3	86,3	87,1	83,7	0,00	0,12
Aseguramiento	85,2	87,5	89,3	86,5	0,04	0,08
Empatía	84,4	86,9	90,8	86,0	0,00	0,11
Tangibles	71,9	73,7	79,5	73,5	0,00	0,15
Género	58,9	59,0	63,2	59,5	0,66	0,03
Edad	25,4	43,2	60,9	35,7	0,00	0,92
Condición laboral	82,0	82,5	59,2	79,4	0,00	0,19
Educación	77,7	72,7	63,2	74,3	0,00	0,14
n	562	354	125	1.041		

Fuente: elaboración propia

Al primer grupo se le denominó "exigente" y se caracterizó por ser el más severo al evaluar los servicios. Este grupo está asociado a los puntajes más bajos en el índice de calidad percibida y en los subíndices de cada dimensión. Los "exigentes" se caracterizan por ser el grupo más joven (menor edad), con trabajo y poseer el mayor nivel educativo.

El segundo grupo está compuesto por los "neutrales", se les denominó así ya corresponde al grupo que brindó las calificaciones intermedias. Son usuarios de mediana edad (adultos), con educación secundaria (nivel intermedio) y poseen trabajo.

El tercer grupo se califica como "tolerante", son más flexibles en la exigencia de calidad en los servicios, asociados a los puntajes más altos en el índice general y en los subíndices de dimensión. Se caracteriza por ser el grupo de mayor edad, que trabaja en menor medida o no trabaja y que posee el nivel educativo más bajo (primaria o menos).

VI Conclusiones y recomendaciones

6.1 Alcances y limitaciones del estudio

A continuación se exponen las principales limitaciones detectadas en la presente investigación, así como su alcance:

- **Recolección de datos:** debido a limitaciones de tiempo y recursos se partió de datos previamente recolectados en otro estudio, específicamente, se utilizaron los datos recolectados en el estudio "Percepción en la Calidad de los Servicios Públicos" el cual se diseñó con un esquema conceptual distinto al utilizado en la presente investigación.

El procedimiento correcto consiste en proponer un modelo teórico-conceptual, construir un instrumento de recolección de acuerdo al planteamiento conceptual, recolectar los datos según dicho planteamiento y ajustar el modelo matemático.

Esta condición produjo un efecto en el ajuste de los modelos causales y una limitante en los alcances de la presente investigación. Específicamente, la afectación de dio en el planteamiento conceptual y el instrumento de recolección.

- **Planteamiento conceptual:** en la presente investigación se utilizó un modelo de 5 dimensiones para medir calidad en servicios, mientras en el estudio "Percepción en la Calidad de los Servicios Públicos", se utilizó un modelo de 3 dimensiones..

Las dimensiones propuestas en el esquema conceptual no se midieron completamente en temas como servicio prometido, accesibilidad, comunicación y presentación del personal, pero sí estuvieron representadas por las variables medidas. Se dispuso de 4 a 7 variables observadas para cada dimensión.

Los datos utilizados fueron adaptados al planteamiento conceptual y al modelo causal, propuestos en esta investigación, sin sufrir ninguna modificación. Para ajustar el esquema de 5 dimensiones se realizó una redistribución de las variables seleccionadas.

Pese a utilizar modelos conceptuales distintos, el estudio del cual se extrajeron los datos y la presente investigación parten de una misma concepción sobre calidad en servicios. Ambas investigaciones consideran que la calidad en servicios se mide por medio de expectativas y percepciones de los usuarios.

- Instrumento de recolección: debido a que los cuestionarios se confeccionaron siguiendo un modelo conceptual distinto, no fue posible utilizar todos los ítems del instrumento de recolección.

Para ajustar los modelos propuestos se tuvo que realizar un filtrado de variables, se seleccionaron las variables que mejor se ajustaban al esquema de 5 dimensiones planteado, resultando en un grupo de 25 variables. Por el contrario, se descartaron 17 variables que no tenían asociación con la definición de las dimensiones del esquema conceptual utilizado.

Adicionalmente, los cuestionarios se diseñaron para un análisis distinto al presentado en esta investigación, para lo cual se utilizaron preguntas medidas en escalas ordinales y nominales. Para ajustar modelos de ecuaciones estructurales se recomienda utilizar variables métricas (escalas de medición intervalo o razón).

Para ajustar el modelo AFC, se utilizó la matriz de correlaciones policóricas y el método de estimación de mínimos cuadrado ponderados. La limitación en la escala de medición de las variables no tuvo un efecto sobre el ajuste del modelo.

- Criterio externo: para realizar la evaluación de la validez, recolectando evidencias en relación a un criterio externo, se dispuso solamente del índice de gestión institucional (IGI) y de los diagnósticos de evaluación de las contralorías de servicios. Estos instrumentos parten de un planteamiento distinto al utilizado en esta investigación, persiguen objetivos diferentes y abarcan otros aspectos no considerados en el índice propuesto.

El IGI es la única medida numérica actual disponible para evaluar servicios públicos, pero desde una perspectiva administrativa, como parte de la gestión de las instituciones que prestan dichos servicios, abarcando otras dimensiones.

El IGI se comenzó a calcular a partir del año 2010 e incluyó solamente dos de las cinco instituciones evaluadas. A partir de 2011 se cambió la estructura del índice y su metodología de cálculo.

Para la evaluación de los índices propuestos, se dispuso de dos valores del IGI, medidos con metodologías distintas. Adicionalmente, no se pudo contar con una medición numérica efectuada en los mismos años en los que se recolectaron los datos de la presente investigación.

- Alcance: el alcance de la presente investigación está condicionado por las limitaciones antes descritas, de las cuales no fue posible medir su impacto. Sí se tomaron en consideración todas las limitaciones mencionadas y se aplicaron medidas correctivas para mitigar sus posibles efectos.

Las inferencias realizadas en la presente investigación abarcan la totalidad de usuarios de los cinco servicios evaluados a nivel nacional y los alcances de las conclusiones obtenidas abarcan a los servicios públicos en general.

6.2 Discusión

De acuerdo a los resultados obtenidos, el modelo RLM posee falta de ajuste, al presentar un coeficiente de determinación bajo, un error de estimación alto y no cumplir con los supuestos para modelos regresión múltiple. Por el contrario, el modelo AFC posee bondad de ajuste ya que logró cumplir con el umbral de los índices de ajuste y con los supuestos correspondientes a los modelo de ecuaciones estructurales.

El modelo AFC se fue correctamente especificado según el esquema conceptual utilizado y resultó sobre identificado, lo que permitió la estimación de los parámetros o coeficientes. No se presentaron estimaciones impropias y se obtuvieron cargas factoriales significativamente diferentes de cero (con $\alpha = 5\%$).

Debido a que los modelos RLM y AFC se plantearon utilizando el mismo enfoque conceptual, las variables incluidas se consideran adecuadas, lo que constituye una evidencia de contenido.

El AFE de ambos modelos respalda la estructura de cinco factores (dimensiones) aunque con una composición diferente de variables. En los dos modelos ajustados se confirma la unidimensionalidad del índice calculado, al resultar un solo factor general que representa la calidad percibida.

En el modelo AFC se obtuvieron correlaciones altas (mayores a 0,60), entre las variables observadas y los subíndices y en el caso del modelo RLM las correlaciones partieron de 0,40 hacia arriba. Las correlaciones entre los subíndices y el índice (modelo RLM) resultaron inferiores a 0,60 y entre variables observadas y el índice (modelo AFC) resultaron mayores a 0,40. Estos resultados son evidencias de contenido.

Los índices obtenidos con los dos modelos ajustados también mostraron evidencias de criterio, ya que se obtuvieron conclusiones similares con los resultados de los índices ICC e ICP y con los resultados de las mediciones externas (IGI y diagnósticos de contralorías de servicios). Estos últimos son utilizados actualmente por el Estado para evaluar la calidad en los servicios. Se obtuvieron coeficientes de validez superiores a 0,60 (correlaciones entre las mediciones significativas, superiores al umbral).

Al no superar el umbral establecido para la fiabilidad compuesta (0,70) y la varianza extraída (0,50), el ICC obtenido mediante el modelo RLM no posee evidencias de constructo. El ICP que se extrajo del modelo AFC logró cumplir con el umbral establecido para los dos coeficientes de validez: fiabilidad compuesta y varianza extraída.

En el modelo AFC se obtuvieron cargas factoriales ajustadas mayores al umbral (0,50) y la fiabilidad de las variables fue superior a 0,60, lo que representan evidencias de constructo. Las puntuaciones de los dos índices calculados (ICC e ICP) se consideran fiables al superar el umbral (0,60) de los coeficientes de fiabilidad evaluados: Spearman-Brown, Guttman-Flanagan, Cronbach y alfa ordinal.

No se detectaron diferencias significativas para las cargas factoriales y las varianzas de error del modelo AFC en las pruebas de invarianza. La estructura factorial del modelo AFC se logró mantener en las dos muestras estudiadas. No hay un comportamiento diferencial en el modelo AFC, validando su estructura conceptual. Estos resultados también se consideran evidencias de constructo a favor del índice ajustado (ICP).

El modelo RLM ajustado no logró cumplir con las condiciones óptimas de los modelos de regresión múltiple. El modelo RLM no se consideró adecuado para generar un índice para medir la calidad en los servicios públicos, por tanto se descartó.

Los puntajes del índice de calidad percibida y de los subíndices de dimensión obtenidos por medio del modelo AFC, son fiables y válidos. Se consideró que el modelo AFC es el más adecuado para medir la calidad percibida en los servicios públicos.

En el análisis de usuarios se detectó que la variable edad está asociada de forma positiva al puntaje de calidad percibida y el nivel educativo está asociado de forma inversa con dicho puntaje. La asociación se presentó en las mismas direcciones en el índice general y en los subíndices de dimensión.

A partir de las asociaciones identificadas, se logró definir tres perfiles de usuarios: los más jóvenes y con mayor nivel educativo asociados a los puntajes más bajos de calidad percibida; los que brindan puntajes intermedios, de mediana edad y educación intermedia, y los que asignan puntajes altos, que poseen mayor edad y menor educación.

6.3 Conclusiones

Con base en los análisis realizados y los resultados obtenidos en el presente trabajo de investigación, se llegó a las siguientes conclusiones:

- El concepto de "calidad percibida" corresponde a un concepto abstracto, de múltiples dimensiones, que puede ser definido de distintas formas a partir de diversos enfoques. Los planteamientos concuerdan en que la calidad en los servicios se evalúa al momento de recibir el servicio, comparando las expectativas con las percepciones de los usuarios.
- Entre los planteamientos conceptuales propuestos en las últimas 4 décadas para medir calidad en los servicios, sobresale la propuesta de la escuela norteamericana, la cual se encuentra representada por el modelo Gaps y por la escala ServQual, planteados por Parasuraman et al. (1985, 1988). Este esquema conceptual logra explicar el concepto de calidad percibida en servicios de forma exhaustiva.

- Bajo el esquema conceptual utilizado (escala ServQual), la calidad percibida en los servicios se compone de cinco dimensiones, las cuales son producto de una serie de variables medidas en escala tipo Likert, que pretenden medir expectativas y percepciones, antes y después de recibir el servicio.
- El modelo de regresión lineal múltiple expone la calidad percibida como una combinación lineal de las dimensiones que la componen, mediante una suma ponderada. El índice calculado con este modelo carece de bondad de ajuste, pero brinda evidencias de fiabilidad y evidencias de contenido y de criterio. No se propone el uso de este modelo para medir la calidad percibida en los servicios públicos.
- El modelo del análisis de factores confirmatorio representa a la calidad percibida por medio de relaciones lineales directas entre las dimensiones del concepto y el concepto, así como entre las variables observadas y las dimensiones, e indirectas entre las variables y el concepto. Este modelo se ajustó al esquema conceptual planteado, lo que permitió medir el concepto de forma exhaustiva.
- El modelo AFC logró cumplir con las condiciones para su estimación y los datos observados indican un buen ajuste. Se logró recolectar evidencias de fiabilidad y validez a favor del puntaje del índice de calidad percibida, calculado a partir del modelo AFC. El ICP es consistente, confiable y logra medir el concepto de la calidad percibida tal y como se propuso en el planteamiento teórico.
- El modelo más adecuado para medir la calidad percibida en los servicios públicos, desde la perspectiva externa de los usuarios de los servicios, es el modelo de análisis de factores confirmatorio, por medio del índice de calidad percibida.
- El modelo AFC ajustado presentó ciertas limitaciones que se centran en las variables utilizadas para su ajuste y en la forma en que fueron medidas (escala de medición). Las limitaciones son producto de utilizar datos recolectados en otra investigación con un instrumento de recolección de datos diseñado con otros objetivos.
- Los servicios públicos de las instituciones evaluadas fueron percibidos como de buena calidad por los usuarios, reflejada por puntajes intermedios (entre 81 y 83 puntos) en el

índice de calidad percibida. Se obtuvieron puntajes de moderados a altos (de 73 a 92 puntos) en las dimensiones del modelo, medidas por los subíndices.

- La dimensión mejor calificada corresponde a *confiabilidad*, que evalúa los temas sobre información solicitada y atrasos en el servicio. Por el contrario, la dimensión con la calificación más baja fue *tangibles*, que analiza el estado físico de las instalaciones y del mobiliario. Los usuarios perciben las 5 dimensiones con una calidad aceptable.
- Los servicios brindados por la SETENA son los que poseen la menor calidad percibida por los usuarios, mientras que los servicios ofrecidos por el Registro Nacional son los que se perciben con mayor calidad. Los restantes servicios recibieron calificaciones moderadas de calidad (entre 78 y 83 puntos).
- El puntaje del índice de calidad percibida está inversamente relacionado con el nivel educativo y la condición laboral y directamente relacionado con la edad: a menor edad, mayor nivel educativo y más trabajo del usuario, menor es el puntaje en el ICP.
- Los puntajes de los subíndices de las dimensiones del modelo mantienen el mismo comportamiento en los usuarios que el puntaje del índice de calidad percibida.
- Se identificaron tres tipos de usuarios: los que poseen una alta percepción de calidad, asociados con niveles bajos de escolaridad y de mayor edad; los que califican con más exigencia la calidad percibida, asociados con edades bajas y a niveles altos de educación, y los que asignan calificaciones intermedias en el índice de calidad percibida, de mediana edad y niveles intermedios de educación.

6.4 Recomendaciones para futuras investigaciones

Con el fin de subsanar las limitaciones identificadas en la realización del presente trabajo, se proponen las siguientes recomendaciones, a ser consideradas en futuras investigaciones sobre el tema de calidad en servicios:

- ✓ Diseñar un nuevo instrumento de recolección de datos partiendo del esquema conceptual de 5 dimensiones. Se recomienda medir cada descriptor de cada

dimensión del esquema conceptual, por medio de una o varias variables. El fin es medir las 5 dimensiones del concepto en su totalidad.

Se considera, además, que el instrumento debe medir expectativas (evaluación antes de recibir el servicio) y percepciones (evaluación posterior al uso del servicio) sobre las características de las dimensiones.

El instrumento se debe componer de ítems en escala de razón para medir todos los aspectos que componen cada una de las dimensiones. Se recomienda utilizar calificaciones de 1 a 10 para medir cada aspecto a evaluar.

En el anexo H se incluye una propuesta de aspectos a evaluar con el fin de medir las 5 dimensiones por completo.

- ✓ Recolectar datos nuevos, siguiendo un diseño metodológico ajustado al esquema conceptual de 5 dimensiones. Se propone aplicar nuevo instrumento a muestras probabilísticas de usuarios de distintos servicios públicos, evaluando cada servicio al menos en dos ocasiones.
- ✓ Realizar un análisis comparativo de modelos factoriales diseñados sobre la base del mismo modelo conceptual, descartando el análisis del modelo de regresión lineal múltiple de previo. Se recomienda ajustar diferentes modelos AFC con los mismos datos recolectados y aplicar pruebas de invarianza entre grupos de la muestra.

Específicamente se propone ajustar un modelo de segundo orden (variables observadas relacionadas con las dimensiones y las dimensiones relacionadas con un factor global), un modelo de primer orden (variables observadas relacionadas a un factor único) y un modelo que incluya correlaciones entre las variables latentes, entre otros.

Adicionalmente se recomienda ajustar el modelo en diferentes grupos de la misma muestra para evaluar diferencias de puntajes por género, nivel educativo, edad y por servicio evaluado.

VII. Referencias bibliográficas

- Abad, F., Garrido, J., Olea, J & Ponsada, V. (2006). *Introducción a la Psicometría. Teoría Clásica de los Tests y Teoría de la Respuesta al Ítem*. Madrid, España: Facultad de Psicología, Universidad Autónoma de Madrid.
- Aiken, L.R. (2003). *Tests psicológicos y evaluación*. México Distrito Federal, México: Pearson Educación, undécima edición (85-106).
- Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster Analysis Series: Quantitative Applications in the Social Science* (44). London, U.K: SAGE Publications Ltd.
- Alvaro, M. (1997). Principios psicométricos de la evaluación psicológica. En G. Buela-Casal & J.C. Sierra (Eds.), *Manual de evaluación psicológica* (pp. 173-192). Madrid, España: Siglo Veintiuno.
- Arias Martínez, B. (2008). Desarrollo de un ejemplo de análisis factorial confirmatorio con LISREL, AMOS y SAS. En M.A. Verdugo, M. Crespo, M. Badía & B. Arias (Eds.) *Metodología en la investigación sobre discapacidad. Introducción al uso de ecuaciones estructurales*. VI Simposio Científico SAID 2008 (pp. 75-120). Salamanca, España: INICO.
- Baker, F. (2001). *The Basics of Item Response Theory*. Wisconsin, USA: ERIC Clearinghouse on Assessment and Evaluation.
- Bartholomew, D., Steele, F., Moustaki, I., & Galbraith, J. (2002). *The Analysis and Interpretation of Multivariate Data for Social Scientists. Texts in Statistical Science*. New York, USA: Chapman & Hall/CRC.
- Barrera, R., Navarro, A. & Peris, M. (2015). El papel de la invarianza factorial en la validación del constructo calidad de servicio electrónico. *Revista Europea de Dirección y Economía de la Empresa*, 24, 13-24.
- Berry, L., Parasuraman, A. & Zeithaml, V. (1988). The Service Quality Puzzle. *Business Horizons*, (September-October), 35-43.
- Berry, L., Parasuraman, A. & Zeithaml, V. (1994). Improving Service Quality in America: Lessons Learned. *Academy of Management Executive*, (May), 35-52.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, USA: John Wiley & Sons.
- Boomsma, A. (2008). Structural Equation Modeling. The SIMPLIS Command Language. Recuperado de: http://www.gmw.rug.nl/~boomsma/csadata/boomsma_2008b.pdf

- Brenes, H. & Gutiérrez, E. (2007). Propuesta de un índice para la medición de la calidad de vida en Costa Rica. *Revista Ciencias Sociales*, 116(2), 113-132.
- Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, 30 (1), 8-32.
- Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walking through the process. *Psicothema*, 20(4), 872-882.
- Cabanellas de Torres, G. (2011). *Diccionario Jurídico Elemental*. Guatemala: Editorial Heliasta.
- Catena, A., Ramos, M., & Trujillo, H. (2003). *Análisis multivariado: Un manual para investigadores*. Madrid, España: Biblioteca Nueva.
- Cea D'Ancona, M. A. (2002). *Análisis Multivariable: Teoría y Práctica en la Investigación Social*. Madrid, España: Editorial Síntesis S.A.
- CEA. (2010). Plan de trabajo de la CEA 2011. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica. Tomado de: <http://www.mideplan.go.cr/>
- CEA. (2011). Informe Noviembre 2010-Noviembre 2011 Comisión de Eficiencia Administrativa (CEA). San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- CEARE. (2008). Propuesta de Acciones de reforma del Estado. I Etapa: Acciones potestad del Poder Ejecutivo. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- CLAD. (2008). Carta Iberoamericana de Calidad en la Gestión Pública. San Salvador, El Salvador: Centro Latinoamericano de Administración para el Desarrollo.
- CLAD. (2009). Carta Iberoamericana de Participación Ciudadana en la Gestión Pública. *XI Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado*. Lisboa, Portugal: Centro Latinoamericano de Administración para el Desarrollo.
- Colmenares, O.A. & Saavedra, J.L. (2007). Aproximación teórica a los modelos conceptuales de la calidad del servicio. *Ciencia y Técnica Administrativa*, 6 (4), 138-175.
- Coto Gamboa, M. (1995). Control Estadístico de Calidad en los Servicios (Tesis de Licenciatura). Universidad de Costa Rica, San José, Costa Rica.
- Cronin, J. J. & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. *Journal of Marketing*, 58 (January), 125-131.

- Crosby, P.B. (1979). *Quality is free: The Art of Making Quality Certain*. New York, USA: New American Library.
- Delvó Gutiérrez, P. & García Santamaría, C. (1986). Control Estadístico de la Calidad (Tesis de Licenciatura). Universidad de Costa Rica, San José, Costa Rica.
- DFOE. (2011a). Análisis y Opinión sobre la Gestión de los Gobiernos Locales en el período 2010. Sexto Informe. San José, Costa Rica: Contraloría General de la República.
- DFOE. (2011b). Informe sobre el proyecto denominado Índice de Gestión Institucional 2010. *Informe N°DEFOE-AHP-1-2011*. San José, Costa Rica: Contraloría General de la República.
- DFOE. (2012a). Informe de los resultados del proyecto denominado Índice de Gestión Institucional. *Informe N°DEFOE-ST-AHP-1-2012*. San José, Costa Rica: Contraloría General de la República.
- DFOE. (2012b). Resultados del Índice de Gestión Municipal del periodo 2011. San José, Costa Rica: Contraloría General de la República.
- Diamantopoulos, A. & Sigua, J. A. (2000). *Introducing LISREL*. (1era ed.). California, Estados Unidos: SAGE Publications Inc.
- Díaz González, Y. & Pons García, R. (2009). Conceptualización y modelización de la calidad de servicio percibida: análisis crítico. *Contribuciones a la Economía*, (Octubre). Recuperado de: <http://www.eumed.net/ce/2009b/dpg3.htm>
- Duarte Castillo, J. L. (s.f.) Calidad de servicio (Revisión de corrientes principales y Propuesta para investigación futura). *Keisen Consultores*. México Distrito Federal, México. Recuperado de: <http://www.keisen.com/portal/wp-content/uploads/2009/12/Corrientes-de-Calidad-en-Servicio.pdf>
- Duque Oliva, E. J. (2005). Revisión del concepto de calidad del servicio y sus modelos de medición. *Innovar*, (Enero a Junio), 64-80.
- Elosua, P. (2005). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto. *Psicothema*, 17 (2), 356-362.
- Elosua, P. & Zumbo, B. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20 (4), 896-901.
- Fernández Pulido, R. (2008). Modelos de medida y análisis factorial confirmatorio. En M.A. Verdugo, M. Crespo, M. Badía & B. Arias (Eds.) *Metodología en la investigación sobre discapacidad. Introducción al uso de ecuaciones estructurales*. VI Simposio Científico SAID 2008 (pp. 29-42). Salamanca, España: INICO.

- Ferrando, P & Anguiano, C. (2010). Análisis Factorial como técnica investigación en psicología. *Papeles del Psicólogo*, 31 (1), 18-33.
- Freiberg, A., Stover, J., de la Iglesia, G. & Fernández, M. (2013). Correlaciones policóricas y tetracóricas en estudios factoriales exploratorios y confirmatorios. *Ciencias Psicológicas*, VII (2), 151-164.
- García López, R. & García Moreno, M. (2010). *La Gestión para Resultados en el Desarrollo: Avances y Desafíos en América Latina y el Caribe*. Washington DC, Estados Unidos: Banco Interamericano para el Desarrollo.
- García Moreno, M. (2011). *Manual SEP: Sistema de Evaluación PRODEV. Gobiernos Subnacionales*. Washington DC, Estados Unidos: Banco Interamericano para el Desarrollo.
- Garvin, D.A. (1983). Quality on the line. *Harvard Business Review*, 61 (September-October), 65-73.
- Gröroos, C. (1990). *Marketing y Gestion de Servicios*. Madrid, España: Ediciones Díaz de Santos S.A.
- Gujarati, D. (2004). *Econometría (4ta ed.)*. México Distrito Federal, México: McGraw-Hill Interamericana.
- Holbrook, M. & Corfman, K. (1985) "Quality and value in the consumption experience: Phaldrus rides again" in *Perceived Quality* Jacoby, J. & Olson, J. Massachusetts, USA: Lexington Books, 31-57.
- Jöreskog, K. G. (1993). *Modelado de Ecuaciones Estructurales con LISREL*. País Vasco, España: EUSTAT.
- Jöreskog, K. G. and Sörbom, D. (1993). *LISREL® 8: User's Reference Guide*. Chicago, USA: Scientific Software International.
- Jöreskog, K. G. (2002). *Structural Equation Modeling with Ordinal Variables using LISREL*. Recuperado de <http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf>
- Kang, G. & James, J. (2004). Service quality dimensions: an examination of Grönroos's service quality model. *Emerald*, 14 (4), 266-277.
- Kaplan, D. (2009). *Structural Equation Modeling: foundations and extensions*. (2nd ed.). California, USA: SAGE Publications Inc.
- Lee, H., Lee, Y. & Yoo, D. (2000). The determinants of perceive service quality ans its relationship with satisfaction. *Journal of Services Marketing*, 14 (3), 217-231.
- Lohr, S. (2000). *Muestreo: Diseño y Análisis*. México Distrito Federal, México: International Thomson Editores.

- Madrigal Pana, J. (2009). Percepción de la Calidad en los Servicios Públicos (noviembre 2008 a febrero 2009). Informe Resultados Generales. San José, Costa Rica: Escuela de Estadística, Universidad de Costa Rica.
- Madrigal Pana, J. (2010). Percepción de la Calidad en los Servicios Públicos. Informe de la segunda ronda con resultados generales para 2009 y 2010. San José, Costa Rica: Escuela de Estadística, Universidad de Costa Rica.
- Martínez, J. A. & Martínez, L. (2008). La medición de la calidad percibida en servicios deportivos; un enfoque en primera persona. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 8 (31) 244-255.
- Messick, S. (1989). Validity. The specification and development of tests of achievement and ability. In R. Lino (Ed.), *Educational Measurement* (3th ed.). Washington D.C., USA: American Council on Education.
- MIDEPLAN. (2007a). Manual Explicativo de los Organigramas del Sector Público Costarricense. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2007b). Plan Nacional de Desarrollo Jorge Manuel Dengo Obregón 2006-2010. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2009). Lineamientos de Gestión para las Contralorías de Servicios. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2010a). Diagnóstico del Sistema Nacional de Contralorías de Servicios 2009. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2010b). Plan Nacional de Desarrollo María Teresa Obregón Zamora 2011-2014. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2010c). Programa de Modernización del Estado 2010-2014. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica. Recuperado de: <http://www.mideplan.go.cr/>
- MIDEPLAN (2011a). Consideraciones sobre el Balance del cumplimiento de metas del PND 2006-2010, según la Memoria Anual de la Contraloría General de la República (2010). San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.

- MIDEPLAN (2011b). Diagnóstico del Sistema Nacional de Contralorías de Servicios 2010: Una visión actual de las Contralorías de Servicios. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2011c). Guía Metodológica para elaboración del Informe Anual de Labores de las Contralorías de Servicios. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN. (2012). Diagnóstico del Sistema Nacional de Contralorías de Servicios Año 2011. Principales retos de la Administración Pública para direccionar los servicios públicos hacia la calidad. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica.
- MIDEPLAN & DGSC. (2010). Premio Nacional a la Calidad y Reconocimiento a Prácticas Promisorias en la Gestión Pública: Glosario de términos. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica y Dirección General de Servicio Civil.
- MIDEPLAN & MH (2013). Lineamientos técnicos y metodológicos para la planeación y seguimiento a metas del Plan Nacional de Desarrollo, para la programación presupuestaria y para la evaluación estratégica en el sector público en Costa Rica. San José, Costa Rica: Ministerio de Planificación Nacional y Política Económica y Ministerio de Hacienda.
- Montero, E. (2012). Los modelos de ecuaciones estructurales como herramienta para explorar posibles relaciones causales en investigación educativa: Una ilustración con datos del PISA 2009 en Costa Rica (Cuarto Informe del Estado de la Educación). Recuperado de: http://www.estadonacion.or.cr/files/biblioteca_virtual/educacion/004/montero_modelos_ecuaciones_estructurales.pdf
- Morales Vallejo, P. (2007). Estadística aplicada a las Ciencias Sociales. La fiabilidad de los test y escalas. Madrid, España: Universidad Pontificia Comillas.
- Morales Vallejo, P. (2011). El Análisis Factorial en la construcción e interpretación de tests, escalas y cuestionarios. Madrid, España: Facultad de Ciencias Humanas y Sociales, Universidad Pontificia Comillas.
- Mulaik, S. A. (2009). Linear causal modeling with structural equations. Florida, USA: Chapman & Hall/CRC Statistics in the social and behavioral sciences series.
- Muñiz, J. (2005). La validez desde una óptica psicométrica. Acta Comportamentalia, 13 (1), 9-20.

- Muñiz, J. (2010). Las teorías de los tests: Teoría Clásica y Teoría de Respuesta a los Ítems. *Papeles del Psicólogo*, 31 (1), 57-66.
- Oh, H. (1999). Service quality, customer satisfaction, and customer value: A holistic perspective. *International Journal of Hospitality Management*, 18, 67-82.
- Oliden, P.E. & Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20 (4), 896-901.
- Oliver, R. (1981). Measurement and Evaluation of Satisfaction Process in Retail Settings. *Journal of Retailing*, 57 (Fall), 25-48.
- Orgaz Baz, M. B. (2008). Introducción a la Metodología SEM: conceptos y propósitos fundamentales. En M.A. Verdugo, M. Crespo, M. Badía & B. Arias (Eds.) *Metodología en la investigación sobre discapacidad. Introducción al uso de ecuaciones estructurales. VI Simposio Científico SAID 2008* (pp. 13-28). Salamanca, España: INICO.
- Ortiz Moya, R. & Montero Rojas, E. (2007). Propuesta para la medición de la calidad de los servicios públicos de salud reproductiva, México 2003. *Población y Salud en Mesoamérica*, 4 (2), 1-26.
- Ossorio, M (2007). *Diccionario de Ciencias Jurídicas, Políticas y Sociales* (1era ed.). Guatemala: Datascan.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49 (Fall), 41-50.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing*, 64 (Spring), 12-40.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1991). Refinement and Reassessment of the SERVQUAL Scale. *Journal of Retailing*, 67 (4), 420-450.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1993). More on Improving Service Quality Measurement. *Journal of Retailing*, 69 (1), 140-147.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1994a). Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria. *Journal of Retailing*, 70 (3), 193-199.
- Parasuraman, A., Zeithaml, V. & Berry, L. (1994b). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research. *Journal of Marketing*, 58 (January), 111-124.

- Pérez-Gil, J., Chacón Moscoso, S. & Moreno Rodríguez, R. (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*, 12 (2), 442-446.
- PNUD & UCR. (2011). Atlas del Desarrollo Humano Cantonal de Costa Rica 2011. San José, Costa Rica: Programa de las Naciones Unidas para el Desarrollo y Universidad de Costa Rica.
- Presidencia de la República y Ministerio de Planificación Nacional y Política Económica. (2009). Expediente N°17.473. *La Gaceta*. (204).10-15.
- Presidencia de la República y Ministerio de Planificación Nacional y Política Económica. (2010a). Decreto Ejecutivo N°35755-PLAN. *La Gaceta*. (36).5-6.
- Presidencia de la República y Ministerio de Planificación Nacional y Política Económica. (2010b). Decreto Ejecutivo N°35925-PLAN. *La Gaceta*. (80).2-4.
- Prieto, G. & Delgado, A. (2010). Fiabilidad y Validez. *Papeles del Psicólogo*, 31 (1), 67-74.
- RAE (2011). *Diccionario de la Lengua Española* (23era ed.). Madrid, España: Real Academia Española.
- Ruiz Díaz, M. (2008). Introducción a los Modelos de Ecuaciones Estructurales. En M.A. Verdugo, M. Crespo, M. Badía & B. Arias (Eds.) *Metodología en la investigación sobre discapacidad. Introducción al uso de ecuaciones estructurales. VI Simposio Científico SAID 2008* (pp. 43-74). Salamanca, España: INICO.
- Ruiz, M., Pardo, A. & San Martín, R. (2010). Modelos de Ecuaciones Estructurales. *Papeles del Psicólogo*, 31 (1), 35-45.
- Sabbir Rahaman, M., Highe Khan, A. & Mahmudul Haque, M. (2012). A Conceptual Study on the Relationship between Service Quality towards Customer Satisfaction: Servqual and Gronroos's Service Quality Model Perspective. *Asian Social Science*, 8 (13), 201-210.
- Sachdev, S. B. & Verma, H. V. (2004). Relative importance of service quality dimensions: a multisectoral study. *Journal of Services Research*, 4 (1), 93-116.
- Sancho Royo, D. (1999) *Gestión de Servicios Públicos: estrategias de marketing y calidad* (1era ed.). Madrid, España: Editorial Tecnos.
- Satorra, A. & Rivera, P. (2008). *Modelos Estadísticos LISREL: Teoría e Ilustraciones*. Barcelona, España: Universidad Pompeu Fabra.
- Serrano Bedia, A.M. & López Fernández, M.C. (2007). Modelos de gestión de la calidad del servicio: revisión y propuesta de integración con la estrategia empresarial. En AEDEM, *El comportamiento de la empresa ante entornos dinámicos. Simposio*

- llevado a cabo en el XIX Congreso anual y XV Congreso Hispano Francés de AEDEM, Vol.2 (Comunicaciones), España.
- Serrano Bedía, A.M., López Fernández, M.C. & García Piqueres, G. (2007). Gestión de la calidad en servicios: una revisión desde la perspectiva del management. *Cuadernos de Gestión*, 7 (1), 33-49.
- Timmos, A.C. (2010). Establishing Factorial Invariance for Multiple-Group. Confirmatory Factor Analysis in KUant Guides. Recuperado de: <http://crmda.ku.edu/guides>
- Toit, S., Toit, M., Mels, G. & Cheng, Y. (2002). LISREL for Windows: SIMPLIS Syntax Files. Recuperado de <http://www.ssicentral.com/lisrel/techdocs/SIMPLISSyntax.pdf>
- Tomás, J., Oliver, A., & Hontangas, P. (2000). Análisis factorial confirmatorio de segundo orden y matrices multirrasgo-multimétodo. *Psicothema*, 12 (2), 534-539.
- Viera, A.L. (2011). Interactive LISREL in practice. Getting started with a SIMPLIS approach. Aveiro, Portugal: Springer.
- Zeithaml, V., Berry, L. & Parasuraman, A. (1993). The Nature and Determinants of Customer Expectations of Service. *Journal of the Academy of Marketing Science*, 21 (1), 1-12.

VIII. Anexos

Anexo A. Variables de estudio

Cuadro A1. Preguntas comunes en los cuestionarios

Tema	Nombre de la pregunta	Escala de respuesta	Selección 1	Dimensión asignada	Selección 2	No Respuesta	Selección 3
	Institución (servicio)		√	---	√	0,0%	√
Características sociodemográficas	Género		√	---	√	0,0%	√
	Edad actual		√	---	√	0,7%	√
	Condición laboral		√	---	√	0,0%	√
	Nivel educativo formal		√	---	√	0,4%	√
	Nacionalidad		√	---	X	---	X
Calificación del servicio recibido	Calificación del servicio recibido	De pésimo (1) a excelente (7)	√	---	√	0,0%	√
	Calificación del servicio del examen teórico		√	---	√	0,3%	√
	Calificación del servicio de prueba manejo		√	---	√	0,0%	√
	Calificación servicio entrega de licencias		√	---	√	1,2%	√
Acceso a los servicios	Medio transporte utilizó para llegar a este lugar	Lista de opciones	√	---	X	---	X
	Dificultad para llegar al lugar	De muy difícil (1) a muy fácil (5)	√	---	X	---	X
Pago por servicios	Pago en colones por el servicio recibido	Monto pagado	√	---	X	---	X
Calificación de las instalaciones físicas	Calificación de la ventilación en el lugar	De pésimo (1) a excelente (7)	√	T	√	1,1%	√
	Calificación de la calma en el lugar por la ausencia de ruido		√	T	√	0,7%	√
	Calificación de la iluminación		√	T	√	1,5%	√
	Calificación del estado de los mobiliario (muebles de espera y de atención)		√	T	√	0,4%	√

Cuadro A1. Preguntas comunes en los cuestionarios

Tema	Nombre de la pregunta	Escala de respuesta	Selección 1	Dimensión asignada	Selección 2	No Respuesta	Selección 3
Calificación de las instalaciones físicas	Calificación de la limpieza en el lugar	De pésimo (1) a excelente (7)	✓	T	✓	1,1%	✓
	Calificación de la limpieza en los servicios sanitarios		✓	T	✓	48,7%	X
	Calificación del tamaño del lugar en comparación con la afluencia de personas		✓	T	✓	0,2%	✓
	Calificación de la seguridad dentro de las instalaciones		✓	T	✓	3,0%	✓
	Calificación de la seguridad en las cercanías de las instalaciones		✓	T	✓	6,6%	X
	Calificación de condiciones existentes para personas con discapacidad		✓	T	✓	11,6%	X
Trato recibido por funcionarios	Hasta qué punto los funcionarios fueron respetuosos	De nada (1) a mucho (4)	✓	E	✓	0,6%	✓
	Hasta qué punto los funcionarios estaban comprometidos con el trabajo		✓	E	✓	1,3%	✓
	Hasta qué punto los funcionarios dieron soluciones rápidas		✓	E	✓	1,3%	✓
	Hasta qué punto los funcionarios tenían habilidades para dar un buen servicio		✓	E	✓	0,9%	✓
	Hasta qué punto los funcionarios mostraron interés por su trámite		✓	E	✓	1,0%	✓
	Hasta qué punto los funcionarios trataron a todos por igual, sin preferencias		✓	E	✓	4,4%	X
Capacidad técnica de funcionarios	Y cuánto diría que sabían lo que hacían	De nada (1) a mucho (4)	✓	A	✓	1,2%	✓
	Y cuánto diría que estaban bien preparados		✓	A	✓	1,3%	✓
	Y cuánto diría que explicaron con instrucciones claras		✓	A	✓	1,2%	✓
	Y cuánto diría que inspiraron confianza		✓	A	✓	1,4%	✓
	Y cuanto diría que tenían experiencia		✓	A	✓	3,2%	✓
Instrucciones durante el servicio	Las instrucciones o explicaciones que le dieron fueron fáciles o difíciles de entender	Sí (1) / No (2)	✓	R	✓	2,3%	✓
	Las instrucciones o explicaciones fueron completas o incompletas		✓	R	✓	2,2%	✓
	Tuvo que completar algún formulario antes o durante el servicio		✓	---	X	---	X
	Algún funcionario completo por Usted algún formulario		✓	---	X	---	X

Cuadro A1. Preguntas comunes en los cuestionarios

Tema	Nombre de la pregunta	Escala de respuesta	Selección 1	Dimensión asignada	Selección 2	No Respuesta	Selección 3
Recorrido dentro de la institución	El recorrido dentro de la institución fue fácil o enredado	De fácil (1) a enredado (3)	√	---	X	---	X
	Se guio por rótulos o señalización	Sí (1) / No (2)	√	---	X	---	X
	Fue guiado por algún funcionario		√	---	X	---	X
	Se guio preguntando		√	---	X	---	X
	Se guio siguiendo las personas de la fila		√	---	X	---	X
	Otro forma de guiarse		√	---	X	---	X
Medios para sugerencias	Enterado(a) de la existencia de algún medio para poner sugerencias, quejas o denuncias	Sí (1) / No (2)	√	---	X	---	X
Ley 8220- Requisitos	Durante la última visita le pidieron alguna información que Usted ya había entregado en otro momento	Sí (1) / No (2)	√	C	√	4,7%	√
	Le pidieron alguna información que haya generado o tenga en su poder alguna otra institución pública		√	C	√	4,3%	√
	Ha oído hablar de la Ley No 8220, la Ley de Protección al Ciudadano del Exceso de Requisitos y Trámites Administrativos		√	---	X	---	X
Tiempo de espera	Tiempo en horas entre el momento que entró a la institución y el momento que empezaron a atenderlo	Cantidad de horas	√	---	X	---	X
	Tiempo en minutos entre el momento que entró a la institución y el momento que empezaron a atenderlo	Cantidad de minutos	√	---	X	---	X
	Ese tiempo fue breve o tardaron mucho	De breve (1) a tardaron mucho (3)	√	R	√	2,4%	√
	Mientras esperaba a ser atendido(a) algún funcionario de los que atendían retrasó el servicio sin justificación	Sí (1) / No (2)	√	C	√	3,4%	√

Cuadro A1. Preguntas comunes en los cuestionarios

Tema	Nombre de la pregunta	Escala de respuesta	Selección 1	Dimensión asignada	Selección 2	No Respuesta	Selección 3
Duración del servicio	Tiempo en horas entre el momento que llegó al puesto de atención y el momento que terminaron el servicio	Cantidad de horas	√	---	X	---	X
	Tiempo en minutos entre el momento que llegó al puesto de atención y el momento que terminaron el servicio	Cantidad de minutos	√	---	X	---	X
	Ese tiempo fue breve o tardó mucho	De breve (1) a tardaron mucho (3)	√	R	√	3,2%	√
	Mientras era atendido(a) el funcionario realizó otras actividades no relacionadas con su trámite	Sí (1) / No (2)	√	C	√	2,3%	√
Conteo de variables	Total de variables		56		38		34
	Total de variables de trabajo (representantes de las dimensiones)		NA		29		25

*Notas: -Selección 1: variables presentes en todos los cuestionarios

-Dimensiones: T = tangibles, E = empatía, A = aseguramiento, R = responsabilidad, C = confiabilidad

-Selección 2: variables seleccionadas según criterio de validez aparente (representantes de las dimensiones)

-Selección 3: variables seleccionadas según criterio de no respuesta (menor al 5%)

Fuente: elaboración propia

Cuadro A2. Variables modeladas por tema y dimensión.

Dimensión	Tema	Nombre de la variable	Número de la variable
Variables de control	Características Socio-demográficas	Género	V ₁
		Edad actual	V ₂
		Condición laboral	V ₃
		Nivel educativo formal	V ₄
	Calificación de calidad	Calificación del servicio recibido (RN, DGME, DGSC, SETENA, MOPT)	V ₅
D1 Confiabilidad	Información solicitada	Solicitud de información que ya había entregado	y ₁
		Solicitud de información que tenga otra institución	y ₂
	Atrasos en la atención	Atraso durante la espera	y ₃
		Atraso durante la atención	y ₄
D2 Responsabilidad	Instrucciones recibidas	Instrucciones fáciles o difíciles	y ₅
		Instrucciones completas o incompletas	y ₆
	Tiempos de espera y de atención	Tiempo de espera	y ₇
		Tiempo de atención	y ₈
D3 Aseguramiento	Capacitación técnica de los funcionarios	Conocimiento de los funcionarios	y ₉
		Preparación de los funcionarios	y ₁₀
		Explicaciones de los funcionarios	y ₁₁
		Confianza en los funcionarios	y ₁₂
		Experiencia de los funcionarios	y ₁₃
D4 Empatía	Compromiso de los funcionarios	Respeto de los funcionarios	y ₁₄
		Compromiso con el trabajo	y ₁₅
		Soluciones rápidas	y ₁₆
		Habilidades para dar buen servicio	y ₁₇
		Interés mostrado en su trámite	y ₁₈
D5 Tangibles	Estado de instalaciones físicas	Ventilación del lugar	y ₁₉
		Ausencia de ruido	y ₂₀
		Iluminación	y ₂₁
		Estado del mobiliario	y ₂₂
		Limpieza del lugar	y ₂₃
		Tamaño del lugar	y ₂₄
		Seguridad dentro de instalaciones	y ₂₅

Fuente: Elaboración propia.

Anexo B. Simbología Modelos de Ecuaciones Estructurales

Tabla B1. Nomenclatura. Modelos de Ecuaciones Estructurales.

ζ : error asociado a variables latentes endógenas	ε : error asociado a variables observadas endógenas
λ : cargas factoriales de primer orden	γ : cargas factoriales de segundo orden
η : variables latentes endógenas	ξ : variables latentes exógenas
y : variables observadas endógenas	

Fuente: elaboración propia

**Tabla B2. Variables latentes endógenas y exógenas.
Modelo AFC de segundo orden.**

ξ_1 : Calidad percibida (exógena)	η_3 : Aseguramiento (endógena)
η_1 : Confiabilidad (endógena)	η_4 : Empatía (endógena)
η_2 : Responsabilidad (endógena)	η_5 : Tangibles (endógena)

Fuente: Elaboración propia

Anexo C. Análisis de no respuesta

Cuadro C1. Distribución porcentual de no respuesta por género, edad y educación. Muestra de calibración.

Variable	Con respuesta	Género		Edad			Educación			n
		Mujer	Hombre	18 a 29	30 a 49	50 o más	Primaria	Secundaria	Universitaria	
Ventilación	No	20%	80%	47%	40%	13%	40%	27%	33%	15
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1344
Ausencia de ruido	No	33%	67%	22%	67%	11%	56%	33%	11%	9
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1350
Iluminación	No	52%	48%	38%	48%	14%	38%	24%	38%	21
	Sí	42%	58%	43%	40%	17%	10%	30%	60%	1338
Mobiliario	No	80%	20%	80%	0%	20%	0%	40%	60%	5
	Sí	42%	58%	43%	41%	17%	11%	30%	60%	1354
Limpieza	No	53%	47%	33%	47%	20%	27%	7%	67%	15
	Sí	42%	58%	43%	40%	17%	11%	30%	60%	1344
Tamaño	No	0%	100%	0%	100%	0%	33%	67%	0%	3
	Sí	43%	57%	43%	40%	17%	11%	30%	60%	1356
Seguridad	No	63%	37%	32%	46%	22%	12%	20%	68%	41
	Sí	42%	58%	43%	40%	16%	11%	30%	59%	1318
Respetuosos	No	13%	88%	38%	50%	13%	50%	25%	25%	8
	Sí	43%	57%	43%	40%	17%	11%	30%	60%	1351
Preparados	No	22%	78%	39%	44%	17%	44%	28%	28%	18
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1341
Soluciones rápidas	No	24%	76%	53%	41%	6%	41%	35%	24%	17
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1342
Buen servicio	No	50%	50%	42%	42%	17%	33%	8%	58%	12
	Sí	42%	58%	43%	40%	17%	11%	30%	60%	1347
Interés	No	29%	71%	29%	50%	21%	50%	21%	29%	14
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1345
Sabían	No	19%	81%	38%	44%	19%	44%	25%	31%	16
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1343
Preparados	No	22%	78%	44%	39%	17%	39%	39%	22%	18
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1341
Instrucciones claras	No	25%	75%	44%	50%	6%	31%	25%	44%	16
	Sí	43%	57%	43%	40%	17%	10%	30%	60%	1343
Confianza	No	37%	63%	37%	42%	21%	37%	21%	42%	19
	Sí	43%	57%	43%	40%	16%	10%	30%	60%	1340
Experiencia	No	55%	45%	50%	41%	9%	23%	14%	64%	44
	Sí	42%	58%	43%	40%	17%	10%	30%	59%	1315

Cuadro C1. Distribución porcentual de no respuesta por género, edad y educación. Muestra de calibración.

Variable	Respuesta	Género		Edad			Educación			n
		Mujer	Hombre	18 a 29	30 a 49	50 o más	Primaria	Secundaria	Universitaria	
Instrucciones fáciles	No	23%	77%	45%	35%	19%	6%	32%	61%	31
	Sí	43%	57%	43%	41%	16%	11%	30%	60%	1328
Instrucciones completas	No	30%	70%	47%	37%	17%	7%	33%	60%	30
	Sí	43%	57%	43%	41%	17%	11%	30%	60%	1329
Información entregada	No	45%	55%	41%	38%	22%	11%	30%	59%	64
	Sí	42%	58%	43%	41%	16%	11%	30%	60%	1295
Información otra institución	No	48%	52%	43%	34%	22%	7%	36%	57%	58
	Sí	42%	58%	43%	41%	16%	11%	29%	60%	1301
Tiempo de espera	No	30%	70%	43%	39%	17%	7%	33%	61%	46
	Sí	43%	57%	43%	41%	17%	11%	30%	60%	1313
Retrasó el servicio	No	39%	61%	39%	29%	32%	19%	39%	42%	31
	Sí	43%	57%	43%	41%	16%	11%	29%	60%	1328
Tiempo de atención	No	47%	53%	44%	44%	13%	9%	28%	63%	32
	Sí	42%	58%	43%	40%	17%	11%	30%	60%	1327
Realizó otras actividades	No	61%	39%	59%	36%	5%	7%	16%	77%	44
	Sí	42%	58%	42%	41%	17%	11%	30%	59%	1315

Fuente: elaboración propia

Anexo D. Distribución de la no respuesta**Cuadro D1. Distribución porcentual de variables control y calificación según imputación. Ambas muestras.**

Variable	Muestra de calibración		Muestra de validación	
	Sin imputación	Con Imputación	Sin imputación	Con Imputación
n	1.359	1.359	1.408	1.408
Género				
Hombre	57,5%	57,5%	56,1%	56,1%
Mujer	42,5%	42,5%	43,9%	43,9%
NS/NR	0,0%		0,0%	
Edad				
18 a 29	42,8%	43,0%	39,9%	12,4%
30 a 49	40,0%	40,5%	41,4%	30,4%
50 o más	16,6%	16,6%	17,8%	57,2%
NS/NR	0,7%		0,9%	

Cuadro D1. Distribución porcentual de variables control y calificación según imputación. Ambas muestras.

Variable	Muestra de calibración		Muestra de validación	
	Sin imputación	Con Imputación	Sin imputación	Con Imputación
n	1.359	1.359	1.408	1.408
Condición laboral				
Trabaja	58,6%	58,6%	55,7%	55,7%
Estudia	6,0%	6,0%	6,6%	6,6%
Trabaja y estudia	20,9%	20,9%	24,1%	24,3%
Labores de hogar	7,7%	7,7%	7,6%	7,6%
Otro	6,8%	6,8%	5,8%	5,8%
NS/NR	0,0%		0,2%	
Educación				
Primaria	10,4%	10,7%	11,7%	12,4%
Secundaria	29,7%	29,7%	30,4%	30,4%
Universitaria	59,6%	59,6%	57,2%	57,2%
NS/NR	0,4%		0,6%	
Calificación				
Pésimo	2,0%	2,0%	2,3%	2,3%
Muy malo	1,7%	1,7%	1,5%	1,5%
Malo	7,4%	7,4%	4,9%	4,9%
Regular	10,2%	10,2%	11,2%	11,5%
Bueno	31,3%	31,3%	35,5%	35,5%
Muy bueno	23,4%	23,4%	20,0%	20,0%
Excelente	24,1%	24,1%	24,2%	24,2%
NS/NR	0,0%		0,4%	

Fuente: elaboración propia

Cuadro D2. Distribución de tamaños de muestra por servicio, con y sin eliminación de casos. Ambas muestras.

Institución	Sin eliminación*			Con eliminación*		
	Total	Calibración	Validación	Total	Calibración	Validación
RN	708	26%	26%	550	26%	27%
MOPT	608	22%	22%	432	21%	21%
DGME	606	23%	21%	489	23%	24%
DGSC	642	22%	24%	468	22%	23%
SETENA	203	8%	7%	137	9%	4%
Total	2.767	1.359	1.408	2.076	1.041	1.035

*Eliminación de casos con datos faltantes

Fuente: elaboración propia

Anexo E. Análisis de regresión lineal múltiple

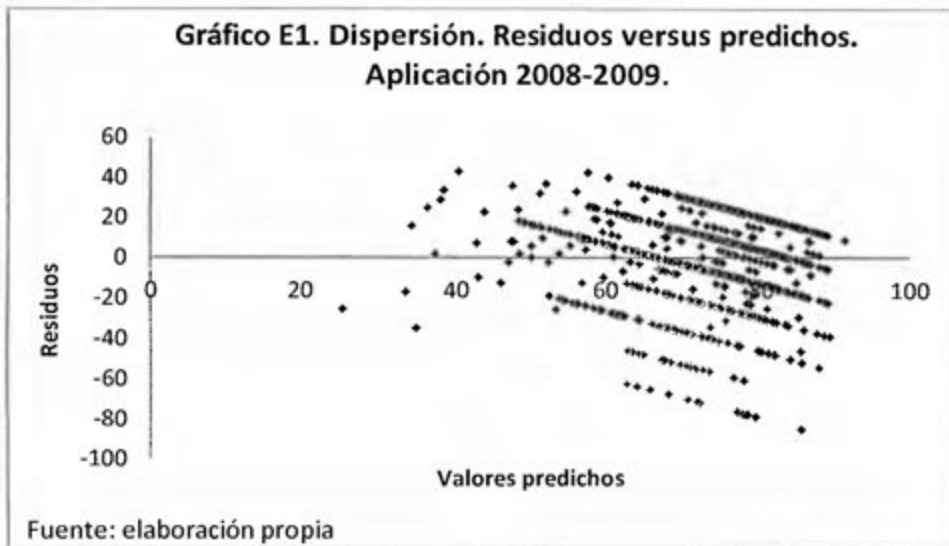
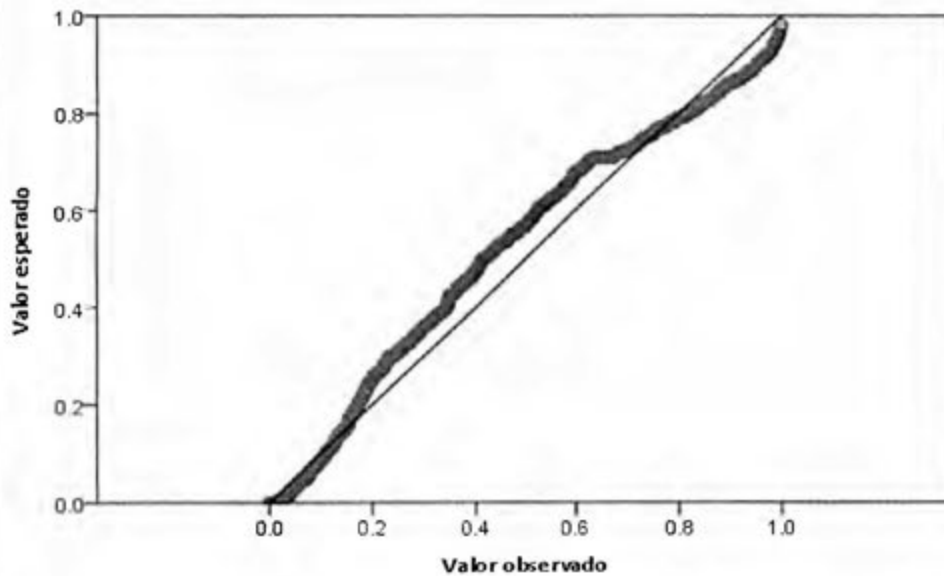
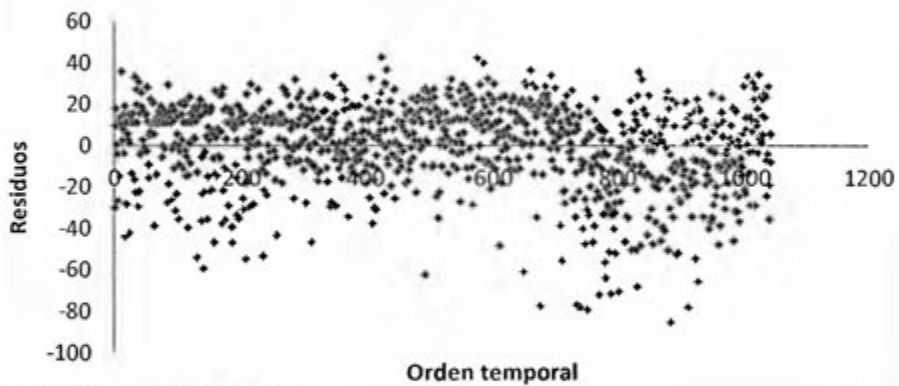


Gráfico E2. P-P normal de residuos estandarizados. Aplicación 2008-2009



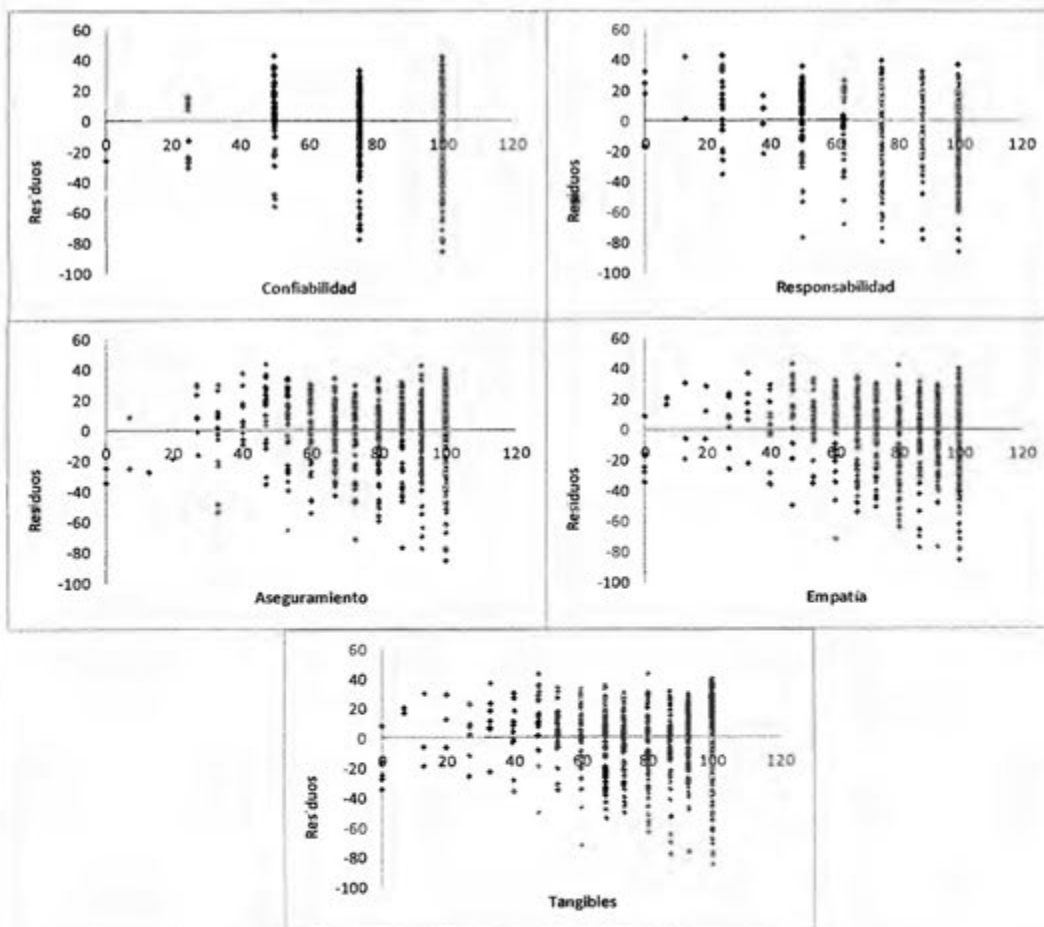
Fuente: elaboración propia

Gráfico E3. Dispersión. Residuos y orden temporal. Aplicación 2008-2009.



Fuente: elaboración propia

Gráfico E4. Dispersión. Residuos versus variables independientes. Aplicación 2008-2009



Fuente: elaboración propia

Cuadro E1. Estadísticos descriptivos. Índice de calificación de calidad (ICC) observado y pronosticado, por servicio. Muestra de validación.

Servicio	ICC	Media	Desviación estándar	Máximo	Mínimo	Mediana	Moda	Asimetría	Curtosis	Li	Ls
Total	Observado	74,24	22,88	100,00	0,00	77,78	100,00	-0,90	0,62	72,85	75,63
	Pronosticado	74,24	9,82	91,44	25,44	75,83	88,68	-1,02	1,69	73,64	74,84
RN	Observado	81,61	21,78	100,00	16,67	83,33	100,00	-1,10	0,37	79,02	84,20
	Pronosticado	79,76	7,45	91,44	49,00	80,77	88,68	-0,76	0,41	78,87	80,65
MOPT	Observado	74,88	15,30	100,00	27,78	77,78	66,67	-0,35	-0,06	72,84	76,93
	Pronosticado	71,73	10,72	87,89	36,29	74,88	80,77	-1,11	0,99	70,30	73,16
DGME	Observado	82,01	20,08	100,00	0,00	83,33	100,00	-1,43	2,88	79,46	84,55
	Pronosticado	74,54	8,78	89,06	34,92	76,54	76,81	-0,88	1,39	73,42	75,65
DGSC	Observado	59,11	25,73	100,00	0,00	66,67	66,67	-0,56	-0,22	55,75	62,47
	Pronosticado	71,62	9,82	87,89	25,44	73,33	85,52	-1,25	3,12	70,34	72,90
SETENA	Observado	67,77	20,46	100,00	16,67	66,67	66,67	-0,20	-0,36	63,56	71,97
	Pronosticado	69,43	9,14	86,12	37,94	70,58	74,47	-0,80	0,64	67,56	71,31

*Siglas: RN = Registro Nacional, MOPT = Ministerio de Obras Públicas y Transportes, DGME = Dirección General de Migración y Extranjería, DGSC = Dirección General de Servicio Civil, SETENA = Secretaría Técnica Nacional Ambiental

Li = límite inferior, Ls = límite superior (límites de confianza al 95%)

Fuente: elaboración propia

Cuadro E2. Correlaciones bivariadas entre variables y subíndices de dimensión. Ambas muestras.

Dimensión	Variable	Muestra calibración		Muestra validación	
		Contingencia	Spearman	Contingencia	Spearman
Confiabilidad	y ₁	0,45	---	0,52	---
	y ₂	0,53	---	0,54	---
	y ₃	0,53	---	0,50	---
	y ₄	0,56	---	0,55	---
Responsabilidad	y ₅	0,47	---	0,53	---
	y ₆	0,54	---	0,55	---
	y ₇	0,67	---	0,69	---
	y ₈	0,66	---	0,67	---
Aseguramiento	y ₉	---	0,66	---	0,71
	y ₁₀	---	0,72	---	0,76
	y ₁₁	---	0,70	---	0,76
	y ₁₂	---	0,76	---	0,82
	y ₁₃	---	0,71	---	0,72
Empatía	y ₁₄	---	0,58	---	0,66
	y ₁₅	---	0,65	---	0,80
	y ₁₆	---	0,74	---	0,79
	y ₁₇	---	0,70	---	0,79
	y ₁₈	---	0,76	---	0,85
Tangibles	y ₁₉	---	0,67	---	0,81
	y ₂₀	---	0,61	---	0,73
	y ₂₁	---	0,67	---	0,77
	y ₂₂	---	0,69	---	0,77
	y ₂₃	---	0,66	---	0,73
	y ₂₄	---	0,66	---	0,74
	y ₂₅	---	0,65	---	0,70

Fuente: elaboración propia

Cuadro E3. Matriz de correlaciones entre índice y subíndices. Ambas muestras.

Variable	Muestra de calibración					Muestra de validación				
	ICC	C	R	A	E	ICC	C	R	A	E
Índice de Calificación	1,00					1,00				
Confiabilidad	0,43	1,00				0,45	1,00			
Responsabilidad	0,68	0,21	1,00			0,70	0,24	1,00		
Aseguramiento	0,47	0,22	0,33	1,00		0,51	0,27	0,37	1,00	
Empatía	0,56	0,25	0,31	0,73	1,00	0,59	0,20	0,34	0,70	1,00
Tangibles	0,82	0,13	0,35	0,29	0,29	0,82	0,45	0,34	0,33	0,37

*ICC= índice de calificación, C= confiabilidad, R= responsabilidad, A= aseguramiento, E= empatía

Fuente: elaboración propia

Anexo F. Análisis de factores confirmatorio

Cuadro F1. Parámetros estimados. Modelo AFC de segundo orden.
Muestra de calibración.

Indicador	Dimensión	Coficiente	Estimador	EE	Estadístico t	CCE	CMC
Carga factorial de primer orden	Confiabilidad	$\lambda 1$	1,00	---	---	0,64	0,41
		$\lambda 2$	1,08*	1,09	0,99	0,47	0,22
		$\lambda 3$	1,11*	1,13	0,99	0,57	0,32
		$\lambda 4$	1,38*	1,27	1,08	0,61	0,37
	Responsabilidad	$\lambda 5$	0,37*	0,27	1,35	0,65	0,43
		$\lambda 6$	0,50*	0,28	1,76	0,75	0,57
		$\lambda 7$	1,63	0,74	2,21	0,85	0,72
		$\lambda 8$	1,00	---	---	0,61	0,37
	Aseguramiento	$\lambda 9$	0,96	0,06	14,98	0,79	0,63
		$\lambda 10$	1,00	---	---	0,86	0,74
		$\lambda 11$	1,04	0,08	12,40	0,79	0,63
		$\lambda 12$	1,21	0,08	15,06	0,83	0,69
		$\lambda 13$	0,93	0,07	13,81	0,78	0,61
	Empatía	$\lambda 14$	0,74	0,08	9,11	0,80	0,64
		$\lambda 15$	0,76	0,08	9,94	0,78	0,61
		$\lambda 16$	0,92	0,08	11,69	0,78	0,60
		$\lambda 17$	0,87	0,07	13,16	0,87	0,76
		$\lambda 18$	1,00	---	---	0,83	0,69
	Tangibles	$\lambda 19$	1,07	0,11	9,67	0,83	0,70
		$\lambda 20$	0,82	0,08	10,90	0,66	0,43
		$\lambda 21$	0,93	0,14	6,56	0,92	0,84
		$\lambda 22$	1,00	---	---	0,82	0,68
		$\lambda 23$	0,90	0,08	11,36	0,87	0,76
		$\lambda 24$	0,97	0,09	10,43	0,73	0,54
		$\lambda 25$	0,94	0,09	10,18	0,82	0,67
Carga factorial de segundo orden	Confiabilidad	$\gamma 1$	0,09*	0,07	1,26	0,65	0,42
	Responsabilidad	$\gamma 2$	0,26	0,11	2,34	0,65	0,43
	Aseguramiento	$\gamma 3$	0,50	0,04	12,27	0,91	0,83
	Empatía	$\gamma 4$	0,69	0,05	13,42	1,00	1,00
	Tangibles	$\gamma 5$	0,45	0,09	4,76	0,43	0,18
Errores asociados a variables observadas	Confiabilidad	$\epsilon 1$	0,02*	0,04	0,59	0,59	---
		$\epsilon 2$	0,07*	0,04	1,58	0,78	---
		$\epsilon 3$	0,05*	0,04	1,06	0,68	---
		$\epsilon 4$	0,06*	0,05	1,10	0,63	---
	Responsabilidad	$\epsilon 5$	0,03*	0,04	0,72	0,57	---
		$\epsilon 6$	0,03*	0,05	0,66	0,43	---
		$\epsilon 7$	0,17*	0,34	0,51	0,28	---

**Cuadro F1. Parámetros estimados. Modelo AFC segundo orden.
Muestra de calibración.**

Indicador	Dimensión	Coefficiente	Estimador	EE	Estadístico t	CCE	CMC
Errores asociados a variables observadas	Responsabilidad	ε8	0,28*	0,18	1,53	0,63	----
		ε9	0,16	0,06	2,58	0,37	----
		ε10	0,11	0,05	2,16	0,26	----
	Aseguramiento	ε11	0,19*	0,10	1,95	0,37	----
		ε12	0,20	0,09	2,18	0,31	----
		ε13	0,17	0,06	2,72	0,39	----
	Empatía	ε14	0,15*	0,12	1,24	0,36	----
		ε15	0,17*	0,10	1,72	0,39	----
		ε16	0,27	0,10	2,65	0,4	----
		ε17	0,11*	0,12	0,94	0,24	----
		ε18	0,21	0,09	2,30	0,31	----
	Tangibles	ε19	0,55*	0,62	0,88	0,30	----
		ε20	0,98	0,19	5,21	0,57	----
		ε21	0,18*	0,99	0,18	0,16	----
		ε22	0,52*	0,30	1,74	0,32	----
ε23		0,29*	0,42	0,69	0,24	----	
ε24		0,88	0,15	5,89	0,46	----	
ε25		0,48*	0,40	1,20	0,33	----	
Errores asociados a variables latentes	Confiability	ζ1	0,01*	0,02	0,49	0,58	----
	Responsabilidad	ζ2	0,09*	0,10	0,96	0,57	----
	Aseguramiento	ζ3	0,05*	0,04	1,44	0,17	----
	Empatía	ζ4	0,00*	0,06	-0,01	0,00	----
	Tangibles	ζ5	0,90	0,14	6,51	0,82	----

Siglas: EE: error estándar

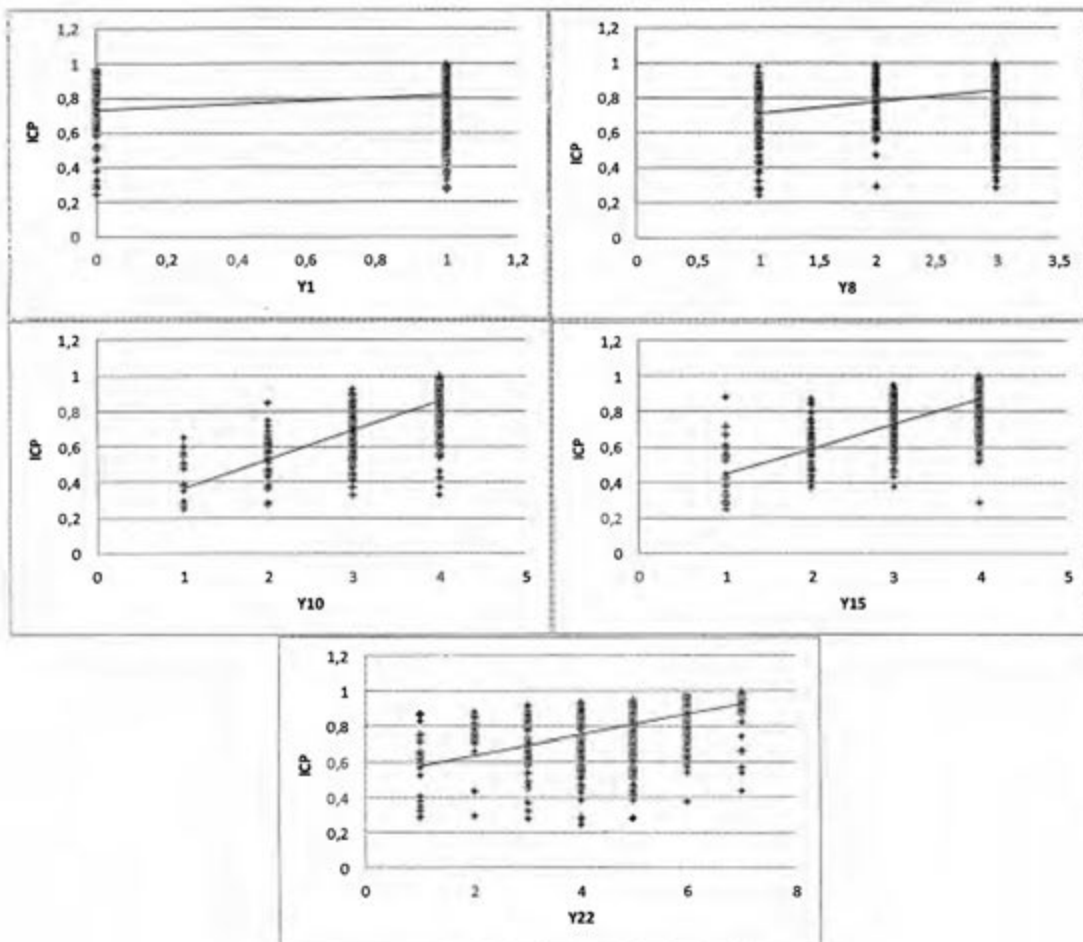
CCE: coeficiente completamente estandarizado

CMC: correlación múltiple cuadrada (R^2)

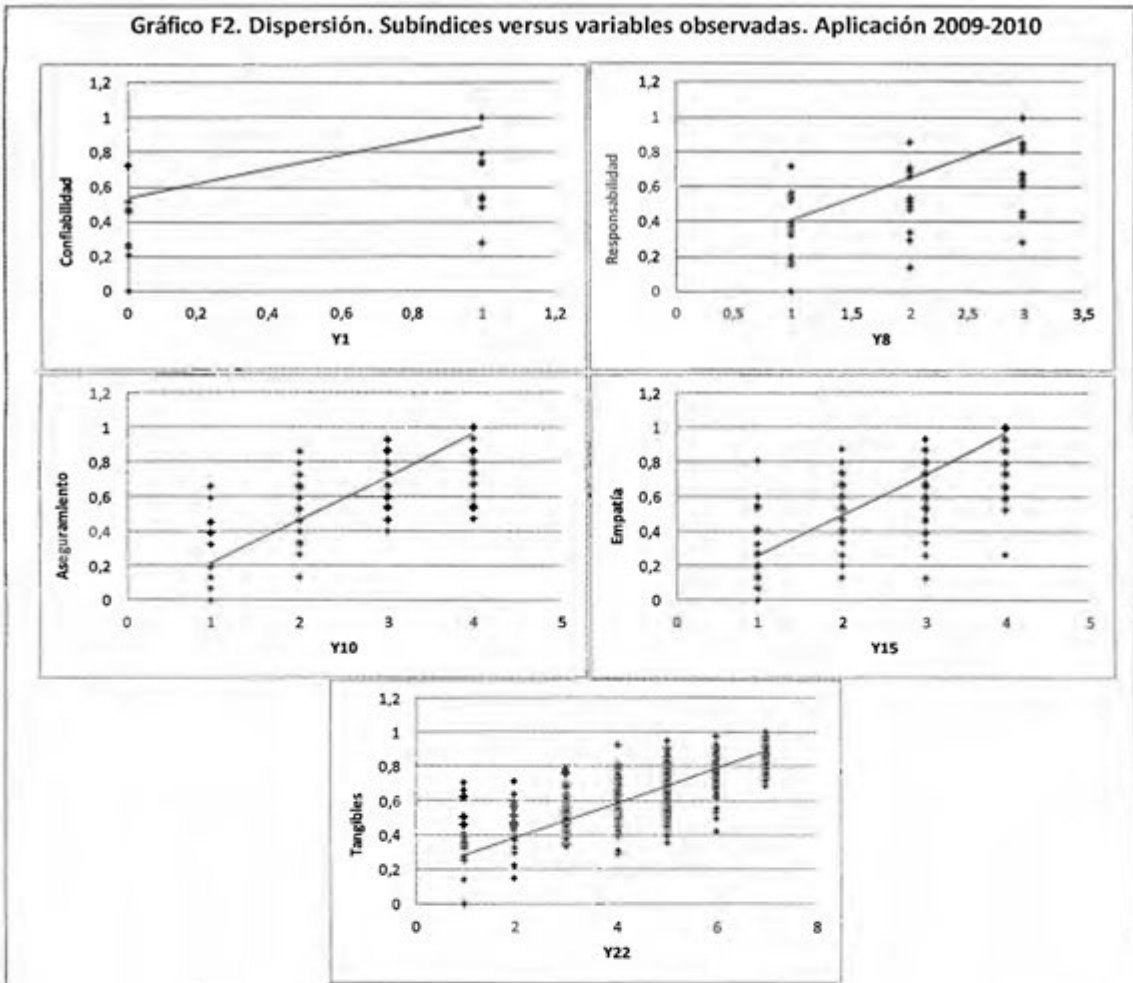
* = coeficiente no significativo con $\alpha = 5\%$

Fuente: elaboración propia

Gráfico F1. Dispersión. ICP versus variables observadas. Aplicación 2009-2010

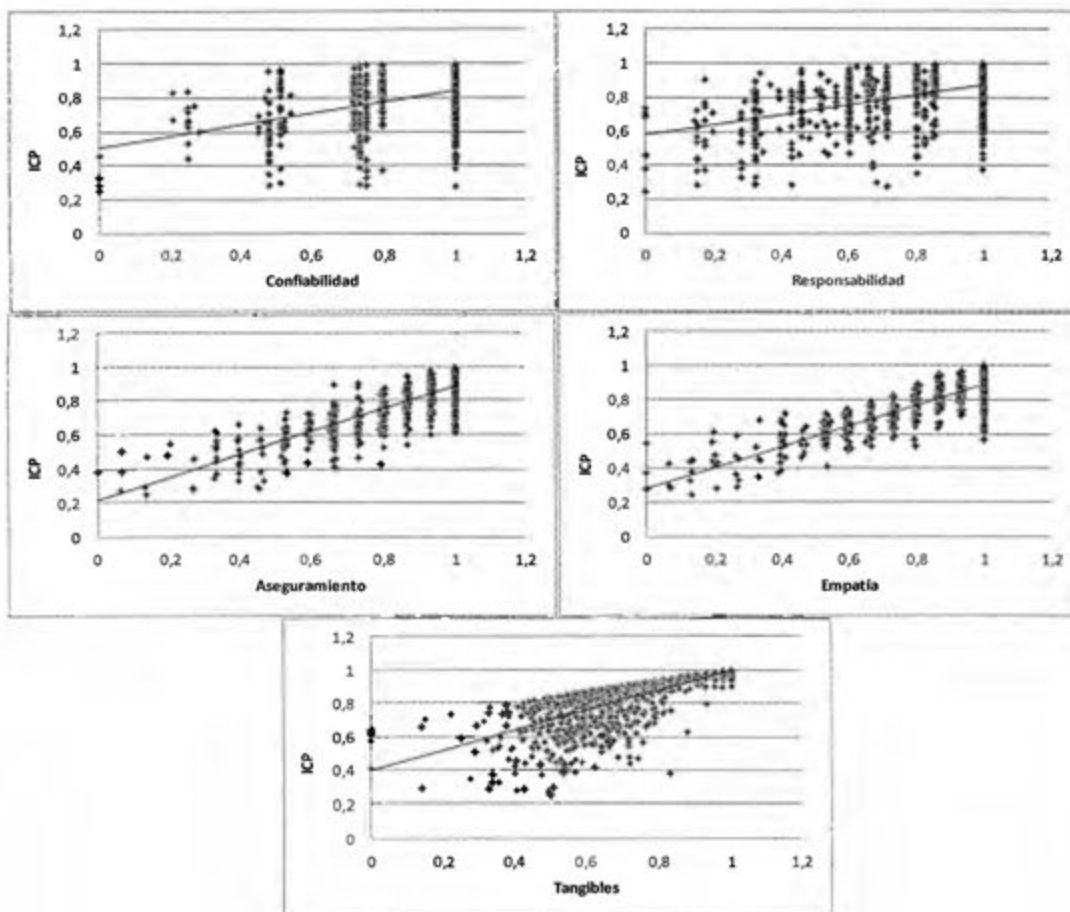


Fuente: elaboración propia



Fuente: elaboración propia

Gráfico F3. Dispersión. ICP versus subíndices. Aplicación 2009-2010



Fuente: elaboración propia

Cuadro F2. Matriz de correlaciones policóricas. Modelo AFC de segundo orden. Muestra de calibración.

Variable	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10	y11	y12	y13	y14	y15	y16	y17	y18	y19	y20	y21	y22	y23	y24	y25
y1	1	1,00	1,00	1,00	1,00	1,00	1,00	0,30	0,84	0,82	0,41	0,40	0,78	0,88	0,25	0,03	0,48	0,62	0,02	0,98	0,90	0,21	0,56	0,14	0,05
y2	0,49	1	1,00	1,00	1,00	1,00	0,35	0,46	0,91	0,46	0,33	0,68	0,20	0,44	0,81	0,35	0,96	0,41	0,19	0,11	0,55	0,00	0,02	0,00	0,04
y3	0,05	0,15	1	1,00	1,00	1,00	0,00	0,10	0,70	0,53	0,32	0,22	0,47	0,78	0,90	0,68	0,90	0,42	0,03	0,02	0,29	0,21	0,99	0,44	0,06
y4	0,32	0,13	0,64	1	1,00	1,00	0,03	0,00	0,31	0,06	0,22	0,69	0,69	0,15	0,58	0,16	0,14	0,03	0,12	0,15	0,14	0,87	0,54	0,54	0,55
y5	0,04	0,21	0,31	0,28	1	1,00	0,06	0,47	0,21	0,54	0,31	0,64	0,57	0,51	0,59	0,47	0,49	0,01	0,25	0,22	0,00	0,49	0,38	0,03	0,87
y6	0,28	0,27	0,41	0,41	0,67	1	0,27	0,01	0,14	0,76	0,57	0,61	0,78	0,71	0,41	0,86	0,64	0,50	0,11	0,65	0,52	0,21	0,07	0,08	0,72
y7	0,26	0,06	0,21	0,18	0,04	0,28	1	0,00	0,98	0,92	0,47	0,05	0,77	0,64	0,85	0,04	0,27	0,04	0,11	0,86	0,29	0,12	0,14	0,06	0,16
y8	0,13	0,20	0,19	0,20	0,12	0,27	0,46	1	0,97	0,48	0,79	0,81	0,73	0,02	0,38	0,46	0,24	0,12	0,01	0,18	0,11	0,57	0,16	0,11	0,01
y9	0,16	0,07	0,29	0,21	0,28	0,35	0,23	0,24	1	0,00	0,00	0,00	0,00	0,42	0,00	0,08	0,07	0,00	0,31	0,39	0,21	0,35	0,20	0,00	0,35
y10	0,26	0,12	0,27	0,29	0,37	0,34	0,27	0,26	0,81	1	0,00	0,00	0,00	0,10	0,00	0,03	0,00	0,00	0,02	0,00	0,01	0,61	0,66	0,00	0,26
y11	0,26	0,10	0,34	0,36	0,41	0,43	0,26	0,23	0,65	0,75	1	0,00	0,00	0,10	0,02	0,00	0,02	0,00	0,25	0,08	0,29	0,15	0,05	0,01	0,18
y12	0,29	0,21	0,33	0,33	0,44	0,45	0,34	0,28	0,66	0,72	0,77	1	0,00	0,00	0,00	0,00	0,00	0,01	0,15	0,01	0,10	0,70	0,01	0,71	
y13	0,18	0,19	0,34	0,30	0,37	0,36	0,26	0,21	0,72	0,78	0,71	0,74	1	0,33	0,01	0,00	0,02	0,00	0,11	0,02	0,17	0,21	0,06	0,01	0,21
y14	0,24	0,05	0,34	0,31	0,37	0,37	0,33	0,26	0,59	0,63	0,66	0,69	0,60	1	0,00	0,00	0,34	0,00	0,04	0,04	0,24	0,05	0,11	0,91	0,11
y15	0,18	0,12	0,36	0,41	0,36	0,36	0,25	0,24	0,62	0,67	0,64	0,63	0,60	0,73	1	0,00	0,00	0,00	0,04	0,18	0,27	0,77	0,29	0,38	0,10
y16	0,23	0,22	0,42	0,40	0,31	0,42	0,29	0,20	0,69	0,67	0,62	0,63	0,64	0,67	0,66	1	0,00	0,00	0,09	0,14	0,07	0,95	0,76	0,64	0,08
y17	0,26	0,17	0,39	0,34	0,25	0,38	0,29	0,26	0,60	0,72	0,63	0,67	0,64	0,68	0,71	0,74	1	0,00	0,04	0,40	0,11	0,67	0,78	0,36	0,09
y18	0,15	0,09	0,41	0,36	0,35	0,31	0,35	0,26	0,66	0,72	0,68	0,71	0,64	0,70	0,70	0,73	0,74	1	0,02	0,16	0,06	0,33	0,33	0,69	0,04
y19	0,17	0,23	0,07	0,11	0,20	0,31	0,26	0,27	0,13	0,21	0,16	0,29	0,21	0,21	0,17	0,26	0,24	0,25	1	0,00	0,00	0,00	0,00	0,00	0,00
y20	0,21	0,09	0,01	0,07	0,16	0,24	0,30	0,20	0,21	0,27	0,16	0,25	0,26	0,24	0,21	0,26	0,22	0,25	0,58	1	0,00	0,00	0,00	0,00	0,00
y21	0,22	0,16	0,07	0,11	0,27	0,33	0,26	0,25	0,34	0,39	0,29	0,38	0,34	0,26	0,28	0,37	0,32	0,33	0,65	0,61	1	0,00	0,00	0,00	0,00
y22	0,18	0,20	0,11	0,13	0,24	0,33	0,27	0,24	0,18	0,24	0,17	0,29	0,19	0,16	0,22	0,23	0,21	0,23	0,66	0,53	0,64	1	0,00	0,00	0,00
y23	0,15	0,11	0,16	0,19	0,24	0,25	0,33	0,28	0,40	0,37	0,28	0,38	0,33	0,30	0,30	0,39	0,33	0,35	0,58	0,54	0,67	0,71	1	0,00	0,00
y24	0,19	0,19	0,14	0,22	0,22	0,20	0,33	0,26	0,15	0,23	0,15	0,26	0,21	0,23	0,24	0,22	0,28	0,24	0,61	0,52	0,59	0,64	0,60	1	0,00
y25	0,16	0,13	0,15	0,19	0,22	0,25	0,29	0,25	0,31	0,36	0,26	0,34	0,34	0,27	0,27	0,29	0,28	0,29	0,58	0,51	0,62	0,62	0,66	0,63	1

*Abajo de la diagonal: los coeficientes de correlación / Arriba de la diagonal: valores p (significancia)

Fuente: elaboración propia

Cuadro F3. Estadísticos descriptivos. Índice y subíndices por servicio. Muestra de calibración.

Servicio	Media	Desviación estándar	Mínimo	Máximo	Mediana	Moda	Asimetría	Curtosis	Li	Ls
ICP	81,87	13,71	13,39	100,00	85,18	100,00	-1,18	1,99	81,04	82,70
RN	87,20	11,91	54,89	100,00	91,03	100,00	-0,88	-0,24	85,79	88,62
MOPT	78,39	15,05	27,16	99,28	83,99	88,47	-1,00	0,26	76,38	80,41
DGME	83,21	12,06	13,50	100,00	86,48	88,47	-1,32	3,99	81,68	84,74
DGSC	80,02	13,77	13,39	99,34	82,87	13,39	-1,68	4,26	78,22	81,82
SETENA	75,23	13,32	26,14	98,45	77,75	88,47	-1,05	1,76	72,50	77,97
IC	92,27	15,71	0,00	100,00	100,00	100,00	-2,22	4,95	91,31	93,22
RN	96,25	10,25	48,47	100,00	100,00	100,00	-2,78	7,33	95,03	97,47
MOPT	91,01	16,80	24,89	100,00	100,00	100,00	-1,74	1,95	88,76	93,25
DGME	94,73	13,05	20,52	100,00	100,00	100,00	-2,92	9,64	93,07	96,38
DGSC	90,70	16,78	0,00	100,00	100,00	100,00	-2,28	6,28	88,51	92,89
SETENA	80,78	22,50	24,89	100,00	79,48	100,00	-0,90	-0,28	76,15	85,40
IR	83,69	22,14	0,00	100,00	100,00	100,00	-1,32	1,05	82,35	85,04
RN	88,74	18,27	0,00	100,00	100,00	100,00	-1,75	3,04	86,56	90,91
MOPT	80,24	26,24	0,00	100,00	100,00	100,00	-1,13	0,19	76,73	83,74
DGME	86,40	20,27	14,12	100,00	100,00	100,00	-1,33	0,91	83,83	88,97
DGSC	77,43	22,88	0,00	100,00	85,88	100,00	-0,93	0,23	74,44	80,42
SETENA	85,21	20,25	15,05	100,00	100,00	100,00	-1,54	2,12	81,05	89,37
IA	86,49	18,84	0,00	100,00	93,58	100,00	-1,53	1,97	85,35	87,63
RN	87,92	17,15	26,26	100,00	100,00	100,00	-1,40	1,00	85,87	89,96
MOPT	83,92	20,18	6,83	100,00	93,50	100,00	-1,18	0,62	81,22	86,61
DGME	90,82	16,29	0,00	100,00	100,00	100,00	-2,28	5,95	88,75	92,88
DGSC	86,13	19,53	0,00	100,00	93,50	100,00	-1,76	3,01	83,58	88,69
SETENA	77,83	21,27	6,50	100,00	80,16	100,00	-0,89	0,40	73,46	82,20
IE	86,03	19,42	0,00	100,00	93,60	100,00	-1,68	2,98	84,85	87,21
RN	88,68	16,51	13,55	100,00	100,00	100,00	-1,59	2,19	86,71	90,64
MOPT	81,76	22,80	0,00	100,00	93,19	100,00	-1,27	0,90	78,71	84,80
DGME	90,31	16,82	0,00	100,00	100,00	100,00	-2,23	6,23	88,18	92,44
DGSC	85,32	19,51	0,00	100,00	93,19	100,00	-1,88	4,59	82,77	87,87
SETENA	78,70	20,94	0,00	100,00	80,79	100,00	-1,22	1,76	74,40	83,00
IT	73,45	15,85	0,00	100,00	72,30	100,00	-0,28	0,43	72,49	74,41
RN	84,19	13,32	41,30	100,00	83,83	100,00	-0,49	-0,49	82,60	85,77
MOPT	69,73	13,53	32,09	100,00	69,23	66,67	-0,03	0,12	67,92	71,54
DGME	69,69	17,98	0,00	100,00	69,09	66,67	-0,50	0,75	67,41	71,97
DGSC	70,34	13,61	18,88	100,00	69,12	66,67	-0,04	0,59	68,56	72,12
SETENA	67,85	12,10	17,11	100,00	66,67	66,67	-0,40	3,59	65,36	70,34

*Siglas: ICP = índice de calidad percibida, IC = índice de confiabilidad, IR = índice de responsabilidad, IA = índice de aseguramiento, IE = índice de empatía, IT = índice de tangibles, Li = límite inferior, Ls = límite superior (límites de confianza al 95%)

Fuente: elaboración propia

Tabla F1. Comparación de variables por modelo.

Modelo conceptual	Modelo ajustado
<p>Confiabilidad Brindado a tiempo De la misma forma y sin errores Servicio confiable y preciso Servicio prometido</p>	<p>Confiabilidad Retraso en el servicio Solicitud de información que ya había entregado o posee otra institución Funcionario realizó otras actividades</p>
<p>Responsabilidad Ayuda a usuarios No dejar usuarios esperando Servicio ágil y rápido</p>	<p>Responsabilidad Instrucciones o explicaciones fáciles y completas Tiempos de espera y de atención</p>
<p>Aseguramiento Conocimiento y competencia Cortesía, amabilidad y respeto Inspirar confianza y seguridad Comunicación</p>	<p>Aseguramiento Funcionarios: sabían lo que hacían, estaban preparados, dieron instrucciones claras, inspiraron confianza, tenían experiencia</p>
<p>Empatía Interés y cuidado Necesidades de usuarios Atención esmerada</p>	<p>Empatía Funcionarios: fueron respetuosos, estaban comprometidos, dieron soluciones, tenían habilidades, mostraron interés</p>
<p>Tangibles Facilidades físicas, apariencia física y ambiente Aspecto del personal Materiales y equipo</p>	<p>Tangibles Calificación de: ventilación, ruido, mobiliario, iluminación, limpieza, tamaño, seguridad del lugar</p>

Fuente: elaboración propia

**Cuadro F4. Análisis de Factores Exploratorio. Solución rotada.
Modelo AFC. Muestra de validación.**

Dimensión	Variable	Selección libre de factores						Selección restringida a 5 factores				
		F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5
Confiabilidad	y ₁	0,07	0,06	0,03	0,01	0,58	0,06	0,08	0,06	0,03	0,02	0,59
	y ₂	0,06	0,03	0,03	0,06	0,64	-0,02	0,05	0,03	0,07	0,04	0,62
	y ₃	0,30	0,03	0,35	0,08	0,00	-0,15	0,18	0,03	0,47	0,05	-0,01
	y ₄	0,33	0,08	0,23	0,12	0,15	-0,10	0,25	0,08	0,36	0,07	0,14
Responsabilidad	y ₅	0,10	0,10	0,09	0,56	0,03	0,05	0,11	0,10	0,12	0,56	0,02
	y ₆	0,27	0,08	0,06	0,67	0,07	0,04	0,27	0,08	0,16	0,56	0,07
	y ₇	0,18	0,15	0,56	0,02	0,03	0,02	0,14	0,17	0,45	0,06	0,03
	y ₈	0,11	0,15	0,53	0,10	0,03	0,07	0,09	0,17	0,39	0,15	0,03
Aseguramiento	y ₉	0,58	0,09	-0,01	0,12	0,06	0,54	0,73	0,10	-0,10	0,19	0,07
	y ₁₀	0,67	0,15	0,02	0,15	0,08	0,53	0,80	0,15	-0,04	0,21	0,08
	y ₁₁	0,62	0,11	0,17	0,27	0,02	0,17	0,63	0,10	0,24	0,26	0,02
	y ₁₂	0,72	0,11	0,23	0,22	0,05	0,09	0,69	0,11	0,36	0,19	0,04
	y ₁₃	0,59	0,10	0,07	0,18	0,03	0,42	0,69	0,10	0,02	0,23	0,04
Empatía	y ₁₄	0,69	0,13	0,11	0,02	-0,04	-0,04	0,61	0,13	0,29	-0,06	-0,05
	y ₁₅	0,81	0,18	0,04	0,01	0,04	-0,09	0,71	0,17	0,30	-0,09	0,03
	y ₁₆	0,68	0,17	0,18	0,08	0,09	0,13	0,67	0,17	0,27	0,05	0,09
	y ₁₇	0,76	0,14	0,16	0,05	0,06	0,12	0,74	0,14	0,28	0,00	0,05
	y ₁₈	0,76	0,19	0,14	0,10	0,12	0,04	0,71	0,18	0,31	0,03	0,12
Tangibles	y ₁₉	0,12	0,71	0,16	0,08	0,11	0,06	0,12	0,72	0,14	0,10	0,11
	y ₂₀	0,15	0,68	0,10	0,03	-0,03	0,00	0,13	0,68	0,12	0,01	-0,03
	y ₂₁	0,12	0,79	-0,01	0,01	0,00	0,02	0,13	0,79	0,01	-0,01	0,00
	y ₂₂	0,10	0,75	0,05	0,10	0,10	0,02	0,10	0,74	0,07	0,10	0,09
	y ₂₃	0,09	0,77	0,01	0,09	0,00	-0,02	0,08	0,76	0,05	0,06	0,00
	y ₂₄	0,19	0,56	0,31	0,06	0,05	0,07	0,17	0,57	0,28	0,09	0,05
	y ₂₅	0,14	0,64	0,12	0,02	0,02	0,06	0,14	0,65	0,11	0,04	0,02

Fuente: elaboración propia

**Cuadro F5. Correlaciones bivariadas entre variables y subíndices de dimensión.
Ambas muestras.**

Dimensión	Variable	Muestra calibración		Muestra validación	
		Contingencia	Spearman	Contingencia	Spearman
Confiabilidad	y ₁	0,71	----	0,71	----
	y ₂	0,71	----	0,71	----
	y ₃	0,71	----	0,71	----
	y ₄	0,71	----	0,71	----
Responsabilidad	y ₅	0,71	----	0,71	----
	y ₆	0,71	----	0,71	----
	y ₇	0,82	----	0,82	----
	y ₈	0,82	----	0,82	----
Aseguramiento	y ₉	----	0,72	----	0,70
	y ₁₀	----	0,81	----	0,77
	y ₁₁	----	0,77	----	0,75
	y ₁₂	----	0,84	----	0,83
	y ₁₃	----	0,76	----	0,70
Empatía	y ₁₄	----	0,63	----	0,66
	y ₁₅	----	0,71	----	0,79
	y ₁₆	----	0,80	----	0,78
	y ₁₇	----	0,79	----	0,81
	y ₁₈	----	0,85	----	0,86
Tangibles	y ₁₉	----	0,79	----	0,81
	y ₂₀	----	0,71	----	0,71
	y ₂₁	----	0,80	----	0,78
	y ₂₂	----	0,82	----	0,77
	y ₂₃	----	0,80	----	0,73
	y ₂₄	----	0,77	----	0,72
	y ₂₅	----	0,77	----	0,71

Fuente: elaboración propia

**Cuadro F6. Correlaciones bivariadas entre variables e índice de calidad percibida.
Ambas muestras.**

Dimensión	Variable	Muestra calibración		Muestra validación	
		Contingencia	Spearman	Contingencia	Spearman
Confiabilidad	y1	0,71	----	0,71	----
	y2	0,71	----	0,71	----
	y3	0,71	----	0,70	----
	y4	0,71	----	0,70	----
Responsabilidad	y5	0,71	----	0,71	----
	y6	0,71	----	0,71	----
	y7	0,82	----	0,82	----
	y8	0,82	----	0,82	----
Aseguramiento	y9	----	0,54	----	0,54
	y10	----	0,61	----	0,61
	y11	----	0,57	----	0,57
	y12	----	0,62	----	0,62
	y13	----	0,52	----	0,52
Empatía	y14	----	0,53	----	0,53
	y15	----	0,62	----	0,62
	y16	----	0,65	----	0,65
	y17	----	0,65	----	0,65
	y18	----	0,67	----	0,67
Tangibles	y19	----	0,67	----	0,67
	y20	----	0,60	----	0,60
	y21	----	0,62	----	0,62
	y22	----	0,63	----	0,63
	y23	----	0,58	----	0,58
	y24	----	0,65	----	0,65
	y25	----	0,59	----	0,59

Fuente: elaboración propia

**Cuadro F7. Parámetros estimados. Validación cruzada.
Modelo AFC segundo orden. Ambas muestras.**

Coeficiente	Muestra calibración					Muestra validación				
	Estimador	E.E.	Estadístico t	Li	Ls	Estimador	E.E.	Estadístico t	Li	Ls
λ_1	1,00	---	---	---	---	1,00	---	---	---	---
λ_2	0,94*	0,68	1,38	-0,39	2,27	1,03*	0,80	1,29	-0,54	2,60
λ_3	0,78*	0,65	1,21	-0,49	2,05	0,95*	0,78	1,22	-0,58	2,48
λ_4	1,15*	0,74	1,55	-0,30	2,60	1,38*	0,98	1,41	-0,54	3,30
λ_5	0,38*	0,26	1,46	-0,13	0,89	0,50	0,24	2,05	0,03	0,97
λ_6	0,51*	0,26	1,95	0,00	1,02	0,59	0,25	2,36	0,10	1,08
λ_7	1,62	0,67	2,41	0,31	2,93	1,57	0,60	2,60	0,39	2,75
λ_8	1,00	---	---	---	---	1,00	---	---	---	---
λ_9	0,97	0,06	16,14	0,85	1,09	0,94	0,05	18,89	0,84	1,04
λ_{10}	1,00	---	---	---	---	1,00	---	---	---	---
λ_{11}	1,03	0,08	13,03	0,87	1,19	1,03	0,06	16,59	0,91	1,15
λ_{12}	1,20	0,08	15,69	1,04	1,36	1,22	0,07	18,78	1,08	1,36
λ_{13}	0,94	0,06	15,02	0,82	1,06	0,90	0,05	17,42	0,80	1,00
λ_{14}	0,74	0,08	9,81	0,58	0,90	0,72	0,06	12,99	0,60	0,84
λ_{15}	0,77	0,07	10,80	0,63	0,91	0,85	0,05	16,05	0,75	0,95
λ_{16}	0,94	0,07	12,62	0,80	1,08	0,94	0,06	15,90	0,82	1,06
λ_{17}	0,87	0,06	13,98	0,75	0,99	0,89	0,05	18,12	0,79	0,99
λ_{18}	1,00	---	---	---	---	1,00	---	---	---	---
λ_{19}	1,07	0,11	10,14	0,85	1,29	1,11	0,08	13,48	0,95	1,27
λ_{20}	0,84	0,07	11,98	0,70	0,98	0,84	0,06	13,56	0,72	0,96
λ_{21}	0,95	0,13	7,11	0,70	1,20	0,94	0,08	11,29	0,78	1,10
λ_{22}	1,00	---	---	---	---	1,00	---	---	---	---
λ_{23}	0,91	0,08	11,63	0,75	1,07	0,90	0,07	12,27	0,76	1,04
λ_{24}	0,94	0,08	11,55	0,78	1,10	0,97	0,08	12,77	0,81	1,13
λ_{25}	0,94	0,09	10,57	0,76	1,12	0,92	0,08	11,68	0,76	1,08
γ_1	0,10*	0,06	1,68	-0,02	0,22	0,08*	0,05	1,61	-0,02	0,18
γ_2	0,25	0,11	2,37	0,03	0,47	0,25	0,09	2,78	0,07	0,43
γ_3	0,51	0,03	16,01	0,45	0,57	0,50	0,03	17,68	0,44	0,56
γ_4	0,68	0,04	15,93	0,60	0,76	0,68	0,04	16,78	0,60	0,76
γ_5	0,45	0,09	4,96	0,27	0,63	0,48	0,07	6,60	0,34	0,62

E.E. = error estándar

Li = límite inferior – Ls = límite superior (límites de confianza al 95%)

* = coeficiente no significativo al 5 %

Fuente: elaboración propia

Anexo G. Análisis de usuarios

Cuadro G1. Distribución porcentual de los niveles del puntaje por índice. Muestra de calibración.

Índice	Nivel del puntaje		
	0 a 70	71 a 85	86 a 100
Calidad percibida	18,0%	31,3%	50,7%
Confiabilidad	6,6%	17,1%	76,3%
Responsabilidad	25,8%	13,8%	60,3%
Aseguramiento	20,0%	10,9%	69,2%
Empatía	19,7%	13,5%	66,8%
Tangibles	45,8%	29,8%	24,4%

Fuente: elaboración propia

Cuadro G2. Distribución de las variables control. Muestra de calibración.

Variable	Frecuencia	Porcentaje
Género		
Mujer	422	40,5%
Hombre	619	59,5%
Edad		
15 a 29	431	41,4%
30 a 49	431	41,4%
50 o más	179	17,2%
Condición laboral		
Trabaja	827	79,4%
Otro	214	20,6%
Educación		
Primaria o menos	113	10,9%
Secundaria	310	29,8%
Universitaria	618	59,4%

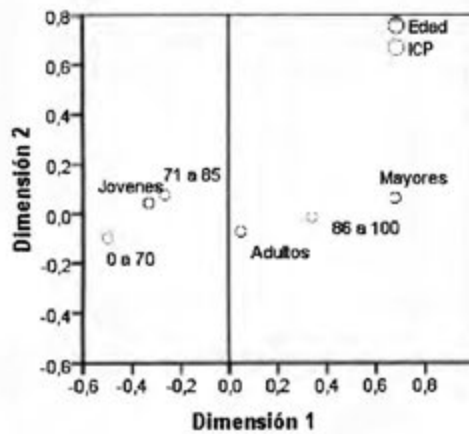
Fuente: elaboración propia

Cuadro G3. Análisis de correspondencias. Muestra de calibración.

Variable	Índice	Dimensión	Valor propio	Inercia	χ^2	Valor p	Proporción de inercia explicada
Edad	Calidad Percibida	1	0,12	0,02	16,24	0,00	1,00
		2	0,00	0,00			0,00
		Total		0,02			
	Confiabilidad	1	0,08	0,01	5,92	0,21	1,00
		2	0,01	0,00			0,00
		Total		0,01			
	Responsabilidad	1	0,12	0,01	14,56	0,01	1,00
		2	0,00	0,00			0,00
		Total		0,01			
	Aseguramiento	1	0,05	0,00	2,56	0,63	0,84
		2	0,02	0,00			0,16
		Total		0,00			
	Empatía	1	0,08	0,01	6,23	0,18	1,00
		2	0,00	0,00			0,00
		Total		0,01			
	Tangibles	1	0,16	0,02	25,55	0,00	0,99
		2	0,02	0,00			0,01
		Total		0,02			
Educación	Calidad Percibida	1	0,15	0,02	25,61	0,00	0,97
		2	0,03	0,00			0,03
		Total		0,02			
	Confiabilidad	1	0,13	0,02	19,49	0,00	0,92
		2	0,04	0,00			0,08
		Total		0,02			
	Responsabilidad	1	0,12	0,01	16,85	0,00	0,90
		2	0,04	0,00			0,10
		Total		0,02			
	Aseguramiento	1	0,08	0,01	7,03	0,13	0,99
		2	0,01	0,00			0,01
		Total		0,01			
	Empatía	1	0,11	0,01	11,95	0,02	0,99
		2	0,01	0,00			0,01
		Total		0,01			
	Tangibles	1	0,13	0,02	19,99	0,00	0,84
		2	0,06	0,00			0,16
		Total		0,02			

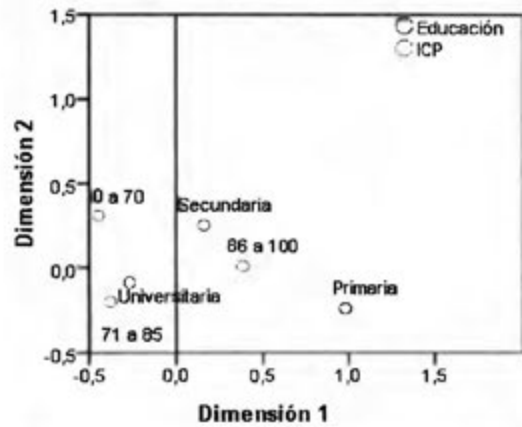
Fuente: elaboración propia

**Gráfico G1. Diagrama dimensional.
ICP vs edad. Muestra calibración.**



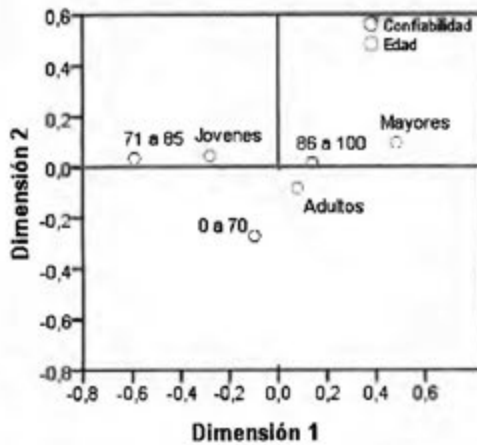
Fuente: elaboración propia

**Gráfico G2. Diagrama dimensional.
ICP vs educación. Muestra calibración.**



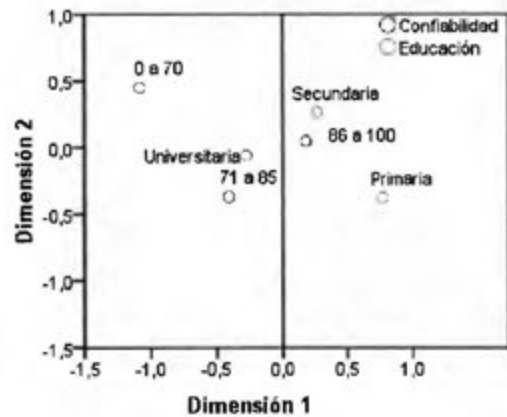
Fuente: elaboración propia

**Gráfico G3. Diagrama dimensional.
IC vs edad. Muestra calibración.**



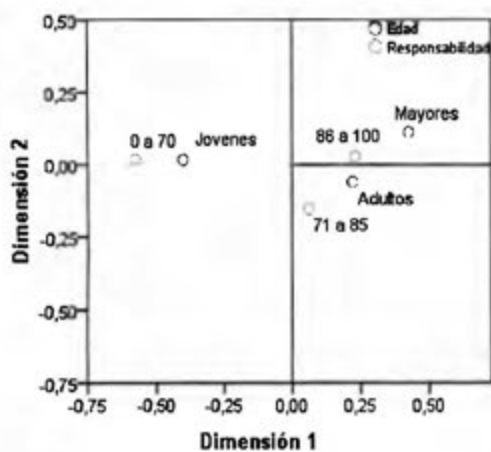
Fuente: elaboración propia

**Gráfico G4. Diagrama dimensional.
IC vs educación. Muestra calibración.**



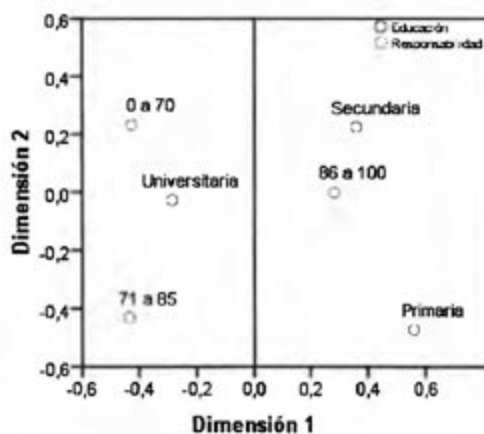
Fuente: elaboración propia

**Gráfico G5. Diagrama dimensional.
IR vs edad. Muestra calibración.**



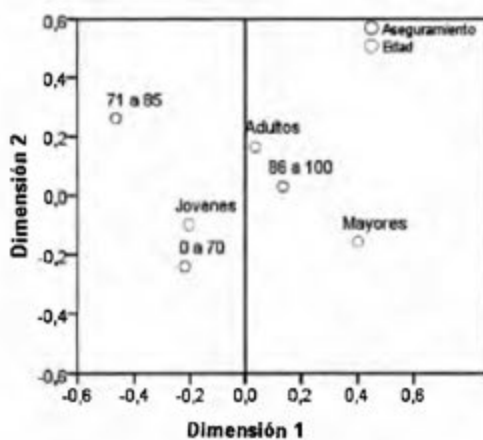
Fuente: elaboración propia

**Gráfico G6. Diagrama dimensional.
IR vs educación. Muestra calibración.**



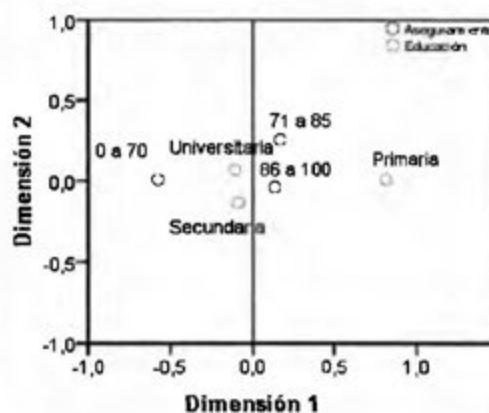
Fuente: elaboración propia

**Gráfico G7. Diagrama dimensional.
IA vs edad. Muestra calibración.**



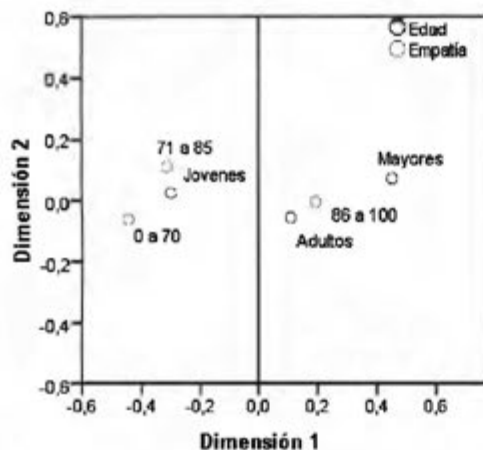
Fuente: elaboración propia

**Gráfico G8. Diagrama dimensional.
IA vs educación. Muestra calibración.**



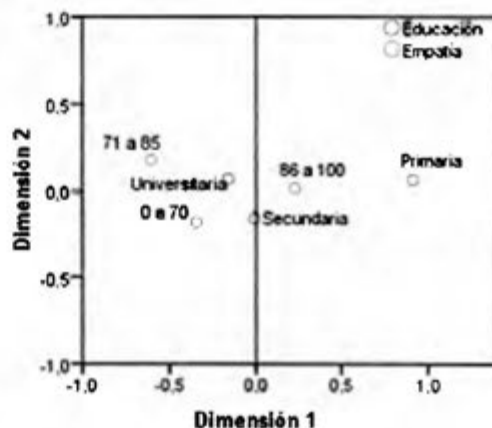
Fuente: elaboración propia

**Gráfico G9. Diagrama dimensional.
IE vs edad. Muestra calibración.**



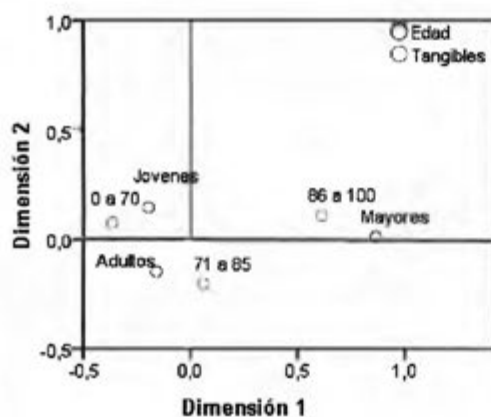
Fuente: elaboración propia

**Gráfico G10. Diagrama dimensional.
IE vs educación. Muestra calibración.**



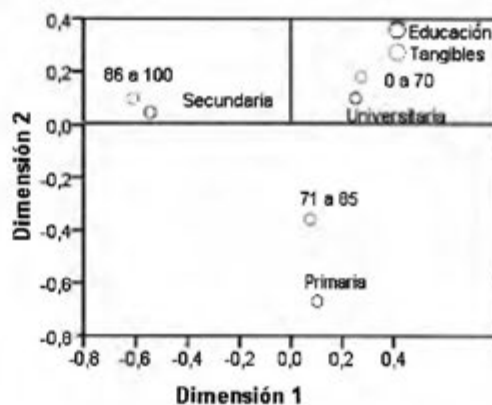
Fuente: elaboración propia

**Gráfico G11. Diagrama dimensional.
IT vs edad. Muestra calibración.**



Fuente: elaboración propia

**Gráfico G12. Diagrama dimensional.
IT vs educación. Muestra calibración.**



Fuente: elaboración propia

Anexo H. Propuesta aspectos a evaluar

Lista de aspectos a evaluar por dimensión

Dimensión	Tema	Aspecto evaluado
Variables de control	Características Socio-demográficas	Género
		Edad
		Ocupación laboral
		Nivel educativo
Calidad	Calificación de calidad	Calificación del servicio recibido
Confiabilidad	Servicio prometido	Solución de la necesidad del usuario
	Servicio confiable y preciso	Cantidad de requisitos solicitados
		Idoneidad de los requisitos solicitados
	Brindado a tiempo	Respeto del horario de atención
Atención a la hora pactada en cita previa		
Responsabilidad	Ayudar al usuario	Medios para solicitar ayuda con el servicio
		Instrucciones/explicaciones brindadas
	No dejar esperando al usuario	Atención continua de usuarios
		Atrasos durante la prestación del servicio
	Servicio ágil y rápido	Tiempo de espera para ser atendido
		Tiempo de atención
Aseguramiento	Cortesía y respeto	Respeto de los funcionarios
		Cortesía de los funcionarios
	Conocimiento y competencia	Conocimiento de los funcionarios
		Preparación de los funcionarios
		Experiencia de los funcionarios
	Confianza y seguridad	Funcionarios inspiran confianza
		Funcionarios inspiran seguridad
	Comunicación	Medios de comunicación disponibles
Empatía	Atención individualizada	Atención individualizada
	Cuidado y comprensión de necesidades	Compromiso con el trabajo
		Preocupación por resolver la necesidad
		Habilidades para dar buen servicio
		Interés mostrado en el trámite
Accesibilidad	Medios de acceso al servicio	

Tangibles	Facilidades físicas	Inmobiliario disponible
		Elevadores – Escaleras eléctricas
		Rampas de acceso
	Apariencia física y ambiente	Iluminación del lugar
		Tamaño del lugar
		Limpieza del lugar
		Ventilación del lugar
		Apariencia de las instalaciones
	Aspecto del personal	Apariencia de los funcionarios
	Materiales y equipo	Equipo disponible para utilizar el servicio
		Materiales necesarios para hacer uso del servicio