



La calidad del servicio percibida en entornos virtuales de formación superior

María Jesús Martínez Argüelles

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



LA CALIDAD DEL SERVICIO PERCIBIDA

EN ENTORNOS VIRTUALES DE

FORMACIÓN SUPERIOR

Tesis doctoral: María Jesús Martínez Argüelles

Director: Dr. José María Castán Farrero

Departamento de Economía y Organización de Empresas

Universidad de Barcelona

Barcelona, Enero de 2006

CAPÍTULO 7. DISEÑO DE LA INVESTIGACIÓN

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong."

Richard Feynman

7.1. Introducción

Una vez expuesto, a lo largo de los capítulos precedentes, el marco teórico en el que se insiere el análisis que nos proponemos realizar en este trabajo de investigación, presentamos ahora, relacionándolos de forma sucinta con la teoría relevante en la materia, los distintos objetivos e hipótesis en torno a los que se articula el eje central de nuestro estudio: la validación de un instrumento de medida que nos permita evaluar el grado de calidad del servicio percibida (CSP) por los estudiantes de formación superior en entornos virtuales.

Antes de acometer la elaboración del instrumento de medida cuantitativo –la encuesta–, que nos permitirá obtener la información primaria necesaria para el contraste de las hipótesis objeto de estudio, hemos realizado un análisis preliminar exploratorio de naturaleza cualitativa con el propósito de identificar, empleando el método del incidente crítico, qué atributos de los servicios formativos en entornos virtuales son relevantes desde el punto de vista de sus usuarios –los estudiantes–. Tras justificar la idoneidad del procedimiento elegido, comenzamos este capítulo describiendo las líneas generales del proceso de implementación del método del incidente crítico y resaltando los aspectos más significativos de la información obtenida.

Finalmente, en la última parte del capítulo, partiendo de los atributos generados en el análisis cualitativo previo y tomando como referencia la escala *e-SERVQUAL*, se explica el procedimiento seguido para la elaboración del cuestionario y la implementación de la encuesta, y, finalmente, se analizan las características de los datos muestrales obtenidos y su grado de adecuación para el contraste de las hipótesis previamente establecidas, cuyos resultados se presentan *in extenso* a lo largo del capítulo 8.

7.2. Objetivos e hipótesis de la investigación

7.2.1. Introducción

El objetivo de cualquier sistema de evaluación de la calidad del servicio percibida es, en última instancia, determinar cuál o cuáles son los factores o dimensiones de un servicio que los consumidores toman como referencia a la hora de enjuiciar la calidad del mismo y, además, averiguar cuál es el peso relativo que otorgan a cada una de esas dimensiones en el conjunto de sus evaluaciones. Al tratarse de un proceso psicológico complejo, de naturaleza elusiva y abstracta, la determinación de los elementos que integran esos juicios de naturaleza subjetiva obliga al investigador a acotar previamente los perfiles del área objeto de estudio por medio de un conjunto de objetivos e hipótesis verificables empíricamente que sirvan de punto de partida a la hora de definir un sistema de evaluación válido y consistente del constructo que se pretende evaluar. Esos interrogantes, que animan los cinco objetivos de la investigación que pretendemos llevar a cabo, se enuncian a continuación como sigue:

Objetivo 1: Estudio de la dimensionalidad del constructo. ¿Qué dimensiones integran el constructo calidad del servicio percibida en la formación superior online? ¿Es posible definir una escala que permita realizar evaluaciones fiables y válidas del constructo calidad del servicio percibida en la formación superior en entornos virtuales? ¿Cuál es su capacidad explicativa?

Objetivo 2: Comparación de los modelos basados en las percepciones con los modelos disconfirmatorios. ¿Tienen mayor validez predictiva y explicatoria las evaluaciones de la calidad del servicio percibida basadas en las percepciones – esto es, en el rendimiento o resultado del servicio– o aquellas que incorporan medidas de las expectativas?

Objetivo 3: Análisis de la relación entre expectativas e importancia. ¿Distinguen los estudiantes en sus evaluaciones entre las expectativas y la importancia atribuidas a los diferentes atributos del servicio? ¿Tienden a emplear ambos conceptos de forma indistinta e intercambiable como defiende parte de la literatura en la materia?

Objetivo 4: Identificación de factores contextuales relevantes. ¿Varían las evaluaciones de la calidad del servicio en función de las características personales de los estudiantes? ¿Influyen el grado de motivación y el rendimiento académico en el nivel de calidad del servicio percibida? ¿Y otros

factores como el tipo y lugar de conexión a Internet, el precio de los estudios o el área de conocimiento en que se integra el programa cursado?

Objetivo 5: Análisis de las relaciones entre los constructos calidad del servicio percibida, satisfacción, fidelidad y disposición a recomendar el servicio. A tenor de la escala que hemos definido, ¿en qué medida el nivel de calidad del servicio percibida o, en su caso, el nivel de satisfacción, tienen un impacto apreciable y significativo sobre la fidelidad y/o la disposición a recomendar el servicio? ¿Cómo se relacionan entre sí estos constructos?

En los siguientes apartados de este capítulo iremos desglosando cada uno de los objetivos de la investigación. La presentación de dichos objetivos, que serán objeto de ulterior contraste empírico en la presente tesis doctoral, se acompaña de una somera contextualización teórica que remite, en última instancia, a la teoría en materia de calidad del servicio percibida que hemos tenido ocasión de sistematizar en los anteriores capítulos de este trabajo.

7.2.2. Dimensionalidad del constructo, fiabilidad y validez de la escala

En la literatura en materia de marketing de servicios el constructo CSP ha sido definido como una evaluación de naturaleza cognitiva (Oliver, 1997) o como una forma de actitud que resulta de comparar las expectativas del consumidor con el resultado del servicio (Parasuraman *et al.*, 1985, 1988; Bolton y Drew, 1991). Esa evaluación, en tanto que juicio general acerca de la excelencia o superioridad de un servicio (Zeithaml, 1987), constituye un constructo de segundo orden, integrado o definido por referencia a un conjunto de variables de primer orden más elementales cuya elucidación constituye el primer paso a la hora de diseñar de un instrumento que permita evaluar la calidad de un servicio como el que nos ocupa.

Aunque existe un relativo consenso entorno al carácter multidimensional del constructo calidad del servicio percibida (Grönroos, 1990; Cronin y Taylor, 1992; Parasuraman *et al.*, 1985), no existe acuerdo, en cambio, a la hora de definir las dimensiones que lo integran (Brady y Cronin, 2001). Como tuvimos ocasión de exponer en el capítulo 3, al sistematizar y resumir la vasta literatura académica en la materia, existen múltiples aproximaciones –con frecuencia, divergentes entre sí– que intentan aprehender cuál es la estructura dimensional del constructo CSP.³³¹ De hecho, incluso en el caso del modelo

³³¹ LeBlanc y Nguyen, 1985; Gumerson y Grönroos, 1987; Edvarsson, 1989; Schwartz y Brown, 1989; Eiglier y Langeard, 1989; Kelly *et al.*, 1990; Higgins y Ferguson, 1991; Lewis, 1993; Baker y

de evaluación de la calidad del servicio más conocido (*SERVQUAL*), diversos estudios demuestran que las cinco dimensiones que lo integran—tangibles, fiabilidad, capacidad de respuesta, seguridad y empatía— no tienen validez universal sino que cambian en función de las características específicas del servicio objeto de estudio (Carman, 1990; Babakus y Boller, 1992; Asubonteng *et al.*, 1996), por lo que cualquier intento de evaluación pasa por adaptar el sistema o modelo de evaluación a esas características.³³²

Esta necesidad de adaptación contextual de las evaluaciones de la calidad del servicio percibida es más intensa, si cabe, en el caso de los servicios *online*, puesto que no existe una evidencia científica consistente que explique de qué forma son percibidos y evaluados estos servicios por parte de los consumidores (Parasuraman *et al.*, 2004). La paradoja estriba en que, si bien el análisis de los modelos diseñados para evaluar la calidad del servicio en los entornos *online* ha recibido una notable cobertura por parte de la literatura académica (LoCascio, 2000; Ahmad, 2002; Lennon y Harris, 2002; Cox, 2002; Gaudin, 2003), sin embargo, “sólo un limitado número de artículos académicos se ocupa directamente de la forma en la que los consumidores evalúan la calidad del servicio en entornos *online* (*e-SQ*) y sus antecedentes y consecuencias” (Parasuraman *et al.*, 2005; pág. 214).

Esta fragilidad conceptual se explica mejor si tenemos en cuenta que la conceptualización de la calidad del servicio percibida ha estado centrada tradicionalmente en capturar la naturaleza interpersonal —física y directa— de los encuentros de servicio entre una organización y sus clientes (Meuter *et al.*, 2000). Esa interacción, se produce en los entornos virtuales en un contexto diferente, virtual y tecnológico, mediado por una interfaz electrónica. Ello, a su vez, ha generado un intenso debate doctrinal que cuestiona la aplicación de la teoría general en materia de calidad del servicio a los entornos *online* y que propugna que un entorno nuevo como el que nos

Limb, 1993; Richard y Allaway, 1993; McDougall y Levesque, 1994; Rust y Oliver, 1994; Humphreys y Williams, 1996; Dabholkar *et al.*, 1996; Brady, 1997; Grönroos, 2000; Brady y Cronin, 2001; por citar sólo algunos ejemplos. En general, se reconoce que en sus evaluaciones de la calidad del servicio los consumidores no tienen en cuenta únicamente el resultado sino también “el proceso de provisión del servicio y los aspectos asociados al mismo” (Zeithaml y Parasuraman, 2004; pág. 2). Se trata, en última instancia, de dos dimensiones primarias: el resultado o servicio básico (*core quality*) y los aspectos relacionales que definen la forma en que el servicio es prestado (*relational aspects*), conceptualizadas originalmente por Grönroos (1982) en su ya clásica distinción entre calidad técnica y calidad funcional. Estas dos dimensiones aparecen, con una u otra denominación, en la mayor parte de los estudios en la materia (*vid.*, por ejemplo, Parasuraman *et al.* 1985; McDougall y Levesque, 1992; Dabholkar *et al.*, 1996; Brady, 1997).

³³² Así lo reconocen incluso los propios creadores de *SERVQUAL* que, en un artículo posterior (Parasuraman *et al.*, 1991b), al revisar la literatura en la materia, observan que el número de dimensiones obtenidas varía entre una (Cronin y Taylor, 1992), dos (Babakus y Boller, 1992), tres (Bouman y van der Wiele, 1994), cuatro (Gagliano y Hathcote, 1994), cinco (Dabholkar *et al.*, 1996), y alcanza cifras incluso más dispares como ocho (Carman, 1990) o diecinueve (Robinson y Pidd, 1998).

ocupa ha de estar asociado con dimensiones específicas –distintas a las tradicionales– de la calidad del servicio que den cuenta de sus particularidades (Cox y Dale, 2001).

En este sentido, el que definimos como **Objetivo nº 1**:

“Determinar las dimensiones que conforman la calidad del servicio percibida en el ámbito de la formación superior online”

recoge esta visión contingente de la calidad del servicio con el propósito de determinar, por medio del análisis factorial –exploratorio y confirmatorio– qué dimensiones integran el constructo calidad del servicio percibida en la formación superior *online*. Por tanto, la primera hipótesis a contrastar es la siguiente:

Hipótesis H1a: *La escala obtenida a partir de las percepciones de los estudiantes de programas de grado virtuales (e-SERVPERF) presenta un carácter multidimensional.*

El instrumento de medida obtenido debe proporcionar resultados consistentes y, además, debe ser capaz de medir realmente aquel constructo del que pretende dar cuenta. Para acreditarlo enunciamos las siguientes hipótesis:

Hipótesis H1b: *La escala e-SERVPERF es fiable.*

Un instrumento de medida es válido si cumple correctamente el propósito para el que ha sido diseñado. Desde el punto de vista teórico se considera que esa validez se expresa en tres sentidos diferentes: validez de contenido, validez de construcción y validez predictiva. Dada la limitada utilidad que la validez de contenido presenta para la estimación de la validez de las medidas empíricas de los conceptos teóricos propios del ámbito del Marketing y la Dirección de Empresas, nuestro análisis de la validez se centra en el análisis de la validez de construcción (Sánchez y Sarabia, 1999; pág. 383).

Hipótesis H1c: *La escala e-SERVPERF presenta validez de construcción.*

Para asegurar desde un punto de vista operativo la validez de construcción es preciso contrastar la validez convergente, la validez discriminante y la validez nomológica del instrumento de medida empleado (Sánchez y Sarabia, 1999; pág. 385). Se afirma que existe validez convergente cuando es posible corroborar por procedimientos independientes el concepto objeto de estudio (Campbell y Fiske, 1959). Por procedimientos independientes se entiende tanto un tipo diferente de escala como valoraciones procedentes de diferentes tipos de sujetos en relación con el mismo concepto (Churchill, 1979). Así, en el caso que nos ocupa, comprobaremos la validez convergente aplicando un tipo diferente de escala y realizando un análisis de la varianza (Bigné *et*

al., 1997; Capelleras y Veciana, 2001) y, en segundo lugar, aplicando la escala a los estudiantes de programas de postgrado *online*, para comprobar si es igualmente válida cuando se aplica a otro grupo de individuos.

Proponemos, por tanto, las siguientes subhipótesis relativas a la comprobación de la validez convergente de la escala:

Subhipótesis H1c1: *Las valoraciones de la escala son significativamente diferentes entre los estudiantes con valoraciones superiores, intermedias e inferiores de la calidad global del servicio.*

Subhipótesis H1c2: *Las valoraciones de la escala son significativamente diferentes entre los estudiantes con valoraciones superiores, intermedias e inferiores del grado de satisfacción con el servicio.*

Subhipótesis H1c3: *La escala e.SERVPERF es aplicable a otros entornos de formación online: los programas de postgrado online.*

Para acabar de contrastar la validez de construcción de la escala es preciso comprobar su validez discriminante –el grado en que la medida no se correlaciona con otras de las que se supone debe diferir– y nomológica –la escala se correlaciona según la forma teóricamente prevista con las medidas de conceptos diferentes, pero teóricamente relacionados–, lo que se expresa a través de las siguientes subhipótesis:

Subhipótesis H1c4: *La escala presenta validez discriminante*

Subhipótesis H1c5: *La escala es válida desde un punto de vista nomológico*

La validez predictiva se refiere al grado de eficacia con el que se puede pronosticar una variable a partir de la medición efectuada (Muñiz, 1996). Esta validez se contrasta mediante el análisis de la correlación entre la escala y la variable criterio. En este caso se habla de validez concurrente, ya que se contrastan simultáneamente los datos relativos a los atributos de la escala y la variable criterio (Sánchez y Sarabia, 1999; pág. 387).

Hipótesis H1d: *La escala e-SERVPERF presenta validez predictiva (concurrente).*

Como variable criterio se utiliza tanto la calidad percibida del servicio como la satisfacción, al tratarse de dos constructos que, como tuvimos ocasión de exponer en el capítulo sexto de esta tesis, parecen exhibir cierto grado de correlación. Por tanto:

Subhipótesis H1d₁: *La correlación entre la calidad global y la escala es positiva y significativa.*

Subhipótesis H1d₂: *La correlación entre la satisfacción y la escala es positiva y significativa.*

Además de examinar la validez concurrente de la escala, resulta imprescindible conocer cuál es su capacidad explicativa. Para ello, en primer lugar, contrastamos la hipótesis de que todas y cada una de las dimensiones detectadas con anterioridad tienen capacidad para explicar parcialmente y de manera significativa tanto la calidad percibida como la satisfacción de los estudiantes. Además, en segundo lugar, suponemos que aunque la escala obtenida es capaz de explicar ambos constructos, tiene mayor capacidad explicativa de la calidad del servicio percibida que de la satisfacción. Por tanto:

Hipótesis H1e: *Cada una de las dimensiones de la escala tiene un efecto positivo y significativo sobre la calidad del servicio percibida.*

Hipótesis H1f: *Cada una de las dimensiones de la escala tiene un efecto positivo y significativo en el grado de satisfacción del estudiante con el servicio.*

Hipótesis H1g: *La escala presenta una mayor capacidad explicativa de la calidad del servicio percibida que de la satisfacción.*

7.2.3. Percepciones *versus* expectativas en las evaluaciones de la calidad del servicio

El modelo *SERVQUAL* (Parasuraman *et al.*, 1985, 1988) constituye el procedimiento más utilizado a la hora de desarrollar escalas que permitan evaluar la calidad del servicio percibida. Inspirado en el paradigma disconfirmatorio de la literatura en materia de satisfacción del consumidor (*vid.*, por ejemplo, Oliver, 1980), se basa en la idea de que la calidad de un servicio, definida como el juicio del consumidor acerca de la superioridad o excelencia del mismo (Zeithaml, 1988), es el resultado de un proceso intelectual en el que los consumidores comparan sus expectativas previas con sus percepciones acerca del servicio recibido.

Al sistematizar las críticas al modelo *SERVQUAL* comprobamos que la más sustantiva, en tanto que afecta uno de los pilares del modelo, es la que alude a la insuficiente evidencia –teórica y empírica– que justifique la necesidad de comparar las expectativas con el servicio recibido para evaluar la calidad del servicio (Carman, 1990). En palabras de Cronin y Taylor (1992, pág. 56), “resulta escasa, de existir alguna, la evidencia empírica que apoya la relevancia del modelo como base para medir la calidad del servicio”.³³³ De ahí que, pese a la amplia difusión del modelo *SERVQUAL*,³³⁴ exista una corriente doctrinal que postule la superioridad de las evaluaciones de los servicios basadas únicamente en las percepciones o resultados del servicio (Carman, 1990; Bolton y Drew, 1991a y b; Cronin y Taylor, 1992; Babakus y Boller, 1992; Boulding *et al.*, 1993).

Estos autores han desarrollado un instrumento estadístico específico denominado *SERVPERF* (Bolton y Drew, 1991; Cronin y Taylor 1992, 1994) que ha recibido el apoyo en los últimos años de diversos trabajos que ponen de manifiesto que las evaluaciones de la calidad del servicio percibida basadas únicamente en las percepciones –esto es, en el rendimiento o resultado del servicio– gozan de mayor validez predictiva y explicatoria que las que incorporan expectativas o preferencias y que, por otra parte, constituyen mejores predictores de las intenciones de compra de los consumidores (Dabholkar *et al.*, 2000; Lee *et al.*, 2000; Brady *et al.* 2002; Page y Spreng, 2002).

³³³ Algunos autores consideran, además, que la inclusión de las expectativas no sólo no aporta información adicional (Bolton y Drew, 1990a, 1990b) sino que, además, reduce el valor discriminante de los resultados (Cronin y Taylor, 1992; McAlexander *et al.*, 1994) por cuanto que esas expectativas: a) tienden a elevarse de forma arbitraria (Brown *et al.*, 1993), b) pueden no existir o no ser lo bastante claras como para ser evaluadas (Iacobucci *et al.*, 1994) y, por último, c) existen múltiples y contradictorias definiciones del concepto (Teas, 1993).

³³⁴ Los partidarios de *SERVQUAL* consideran que su modelo ha sido objeto de un intenso contraste empírico en múltiples ámbitos (Parasuraman *et al.*, 1994) que ratifica la validez de la inclusión de las expectativas en las evaluaciones de la calidad del servicio.

La superioridad del modelo *SERVPERF* justifica que éste haya sido el procedimiento empleado en el *Objetivo 1* a la hora de definir las dimensiones que integran el constructo calidad del servicio percibida en la formación superior *online*. Con todo, como quiera que la cuestión dista de ser pacífica, puesto que, aun siendo minoritarios, existen otros estudios sostienen que *SERVQUAL* y *SERVPERF* presentan una dimensionalidad y una validez predictiva similares (*vid.*, por ejemplo, Zhou, 2004) resulta pertinente comparar ambos procedimientos en el contexto de la formación superior en entornos virtuales de formación; tarea ésta que acometemos en el marco de la *Hipótesis H2* integrada en el que definimos como **Objetivo nº 2**:

“Análisis comparativo de las escalas de medida directas, basadas en las percepciones, y de las escalas inspiradas en el paradigma disconfirmatorio”.

Hipótesis H2: Las escalas directas, basadas en las percepciones, exhiben una mayor capacidad explicativa que las escalas inspiradas en el paradigma disconfirmatorio.

7.2.4. Importancia versus expectativas

Si en el *Objetivo nº 1* empleábamos un instrumento directo (*SERVPERF*) para evaluar la calidad del servicio y, en el *Objetivo nº 2* analizábamos y comparábamos el valor predictivo de ambos tipos de escalas –las directas, basadas en las percepciones o resultados del servicio y las indirectas o disconfirmatorias–, en el ámbito de la formación superior *online*, nos acercamos ahora al que constituye el tercer modelo de evaluación de la calidad del servicio percibida más conocido: el modelo importancia-percepciones (*IPA*).

El origen de este modelo remite, de nuevo, a la literatura en materia de satisfacción del consumidor, que acuñó la idea de que no todas las dimensiones de un servicio tienen un impacto equivalente sobre el nivel de satisfacción (Johnston, 1995; Oliver, 1997). Si, como afirman los autores que defienden este modelo, sólo algunas de las dimensiones que los consumidores emplean para evaluar la calidad del servicio son realmente relevantes a la hora de definir el constructo, resulta imprescindible obtener información sobre el nivel de importancia que los estudiantes, en nuestro caso, atribuyen a cada una de esas dimensiones (Joseph y Joseph, 1997, Ford *et al.* 1999, O’Neill y Palmer, 2004).³³⁵

³³⁵ Los trabajos que acabamos de citar aplican el modelo *IPA* en el ámbito de la formación superior. Otros autores que emplean el modelo *IPA* en ámbitos diversos son: Guadagnolo, 1985;

El principal problema del modelo *IPA* estriba en que no existe una clara definición de lo que se considera “importancia”, por lo que suele emplearse el término “simplemente como el nivel de relevancia de un atributo que parece ser un determinante de una decisión de compra” (Oh, 2001; pág. 619). En este sentido, algunos autores (*vid.*, por ejemplo, Oh y Parks, 1997; Oh, 2001) destacan que los conceptos *importancia* y *expectativas* tienden a ser empleados de forma indistinta e intercambiable, como si se tratara de sinónimos (*vid.*, por ejemplo, en este sentido, Chon *et al.*, 1991; Hollenshorst *et al.*, 1992; Hill, 1995; Tzung-Cheng y O’Leary, 2002).³³⁶ Sin embargo, otros autores como Ryan (1999), distinguen entre importancia y expectativas partiendo de la base de que éstas constituyen la expresión del “resultado tolerado” del servicio, mientras que la importancia es un reflejo del “resultado esperado”. De igual forma, los propios creadores del modelo *IPA* (Martilla y James, 1977) proponen una sutil distinción al considerar que la satisfacción del consumidor es una función tanto de sus expectativas en relación con ciertos atributos importantes, como de sus evaluaciones del resultado del servicio en relación con dichos atributos.

La polémica que acabamos de exponer, unida a la ausencia de una clara definición del concepto *importancia*, hace que resulte relevante determinar, conforme se plantea en la hipótesis que integra este **Objetivo nº 3**, si los estudiantes en un entorno de formación superior *online* conceptualizan de forma diferenciada las expectativas y la importancia en el contexto de sus evaluaciones de la calidad del servicio.

“Análisis comparativo de la conceptualización que los estudiantes realizan de los constructos *expectativas* e *importancia* en relación con los diferentes atributos del servicio”.

Hipótesis H3: *Los estudiantes conceptualizan de forma diferente las expectativas y la importancia en relación con cada uno de los atributos del servicio.*

7.2.5. El modelo en su contexto: factores relevantes

Aunque los instrumentos de medida de la CSP del tipo *SERVQUAL* y *SERVPERF* fueron diseñados con vocación de universalidad, esto es, como escalas genéricas

Ortinau *et al.*, 1989; Carman, 1990; McDougall y Levesque, 1992; Ennew *et al.*, 1993; Oh y Parks, 1998; Wade y Eagles, 2003; Hudson *et al.*, 2004; Aigbedo y Parameswaran, 2004.

³³⁶ De hecho, algunos autores (*vid.*, por ejemplo, Li y Kaye, 1998) utilizan la importancia percibida de cada uno de los atributos de la docencia como indicador de las expectativas de los estudiantes.

aplicables en principio a todo tipo de servicios (Parasuraman *et al.*, 1985; 1988; 1991a), cada servicio posee un conjunto de factores críticos singulares o específicos. Es por eso que la modelización de la CSP requiere una previa adaptación de la escala que se pretende construir a las particularidades del servicio objeto de estudio en cada caso (Cox y Dale, 2001; Chen, 2004).

En el caso de la formación superior *online* uno de sus principales rasgos distintivos es la presencia de una interfaz tecnológica que refuerza la necesidad de participación por parte de los destinatarios de los servicios formativos –los estudiantes– en el proceso de provisión del mismo (Bitner, 1992). Conviene recordar, en este sentido, que en un entorno tecnológico y virtual como el que nos ocupa, la existencia de una interacción mediada por la tecnología entre el usuario y la organización prestadora del servicio, exige que el usuario disponga de cierto dominio de las herramientas tecnológicas asociadas a los entornos *online* y, específicamente, a los entornos virtuales de formación, ya que en otro caso su percepción del servicio será presumiblemente peor (Semeijn *et al.* 2005).

Esta necesidad de intervención directa y activa por parte del estudiante hace que resulte interesante comprobar si se observan diferencias apreciables en las percepciones acerca de la calidad del servicio en función de factores tales como el sexo, la edad, la experiencia previa en formación universitaria y, en particular, en formación *online*, el grado de motivación o el nivel de rendimiento académico de los estudiantes, ya que ésta ha sido la conclusión general alcanzada en otros estudios en la materia (Gagliano y Hathcote, 1994; Zeithaml *et al.*, 2002, Dabholkar y Bagozzi, 2002).³³⁷ Para descubrir si dichos factores influyen efectivamente sobre las percepciones de los consumidores en el contexto de la formación superior en entornos virtuales nos proponemos, en el marco de lo que hemos definido como **Objetivo n° 4**:

“Examinar cómo varía el nivel de calidad del servicio percibida en la formación superior virtual en función de diversos factores contextuales”

Y con ese objetivo enunciamos las siguientes hipótesis:

³³⁷ Los factores demográficos continúan siendo uno de las bases más aceptadas para la segmentación de mercados y consumidores (*vid.*, por ejemplo, Kotler y Armstrong, 1991), entre otras razones porque resultan más sencillos de evaluar que otros factores de tipo conductual o psicográfico. La edad, por ejemplo, es un importante factor predictor de los cambios en actitudes y comportamientos de los usuarios (Hansman y Schutjens, 1993). Lo mismo ocurre con el sexo (Kotler y Armstrong, 1991).

Hipótesis H4a: *El sexo del estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4b: *La edad del estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4c: *La experiencia del estudiante en formación universitaria afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4d: *La posesión de una titulación universitaria afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4e: *La experiencia del estudiante en formación online afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4f: *La experiencia del estudiante en la propia institución universitaria afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4g: *La motivación del estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4h: *El rendimiento académico del estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4i: *El área de conocimiento a la que pertenece el programa cursado afecta de forma directa e indirecta a la calidad del servicio percibida.*

En los entornos *online* los usuarios interactúan con otros individuos a través de una interfaz electrónica y, por ello, la tecnología juega un papel fundamental a la hora de facilitar el proceso de intercambio de información entre los usuarios –en este caso, los estudiantes– y entre éstos y el proveedor del servicio (Parasuraman y Grewal, 2000). Si las percepciones sobre la calidad en los entornos tradicionales son el resultado, en buena medida, de los encuentros de servicio entre el proveedor del servicio y los consumidores, en los servicios *online* las percepciones sobre la calidad están basadas generalmente en evaluaciones de la interacción entre el usuario y la interfaz del usuario (van Riel *et al.*, 2004). Uno de los factores esenciales de dicha interfaz es, junto al diseño y las consideraciones estéticas, la facilidad de acceso al servicio (Dabholkar, 1996; Kaynama y Black, 2000) que viene determinada por factores tales como el tipo y la velocidad de la conexión a Internet (Gummerus *et al.*, 2004; van Riel *et al.*, 2004). Para evaluar su impacto sobre el nivel de CSP formulamos las Hipótesis H4j y H4k:

Hipótesis H4j: *El tipo de conexión a Internet que utiliza el estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Hipótesis H4k: *El lugar de conexión del estudiante afecta de forma directa e indirecta a la calidad del servicio percibida.*

Son múltiples los estudios (Jarvenpaa y Todd, 1997; Voss *et al.*, 1998; Mittal *et al.*, 1998; Bolton y Lemon, 1999) que correlacionan el precio con las percepciones de los consumidores en relación con la calidad del servicio. En el caso de los servicios adquiridos *online*, el hecho de que los consumidores no puedan examinar previamente de forma directa el servicio que van a recibir, ni inferirlo por referencia a otras claves físicas –como las instalaciones o la apariencia del personal– que integran el denominado “*servicescape*” o entorno de servicio, hace que los consumidores *online* “estén obligados a utilizar el precio como indicador de la calidad del servicio” (Jiang y Rosenbloom, 2005, pág. 156), lo que convierte a éste en un factor determinante de la CSP (Jarvenpaa y Todd, 1997; Liu y Arnett, 2000). De ahí que nos propongamos verificar la siguiente relación:

Hipótesis H4l: *El precio de los estudios afecta de forma directa e indirecta a la calidad del servicio percibida.*

7.2.6. Calidad del servicio, satisfacción, fidelidad y disposición a recomendar el servicio

La investigación académica en materia de calidad del servicio percibida ha estudiado la relación entre este constructo y otros tales como la satisfacción del consumidor (Bolton y Drew, 1991; Boulding *et al.*, 1993; Anderson *et al.*, 1994; Spreng *et al.*, 1996), las intenciones de compra (Reichheld y Sasser, 1990; Bolton y Drew, 1991; Cronin y Taylor, 1992; Boulding *et al.*, 1993; Bitner y Hubbert, 1994; Taylor y Baker, 1994), la disposición a recomendar el servicio (Buchanan y Gillies, 1990; Dent, 1990; Christopher *et al.*, 1991) o la rentabilidad empresarial (Rust y Zahorik, 1993; Zahorik y Rust, 1994; Greising, 1994; Bolton y Drew, 1994; Rust *et al.*, 1995 y Zeithaml *et al.*, 1996).

Verificar el sentido y la intensidad de esas relaciones conceptuales en el ámbito de la formación superior *online* resulta fundamental a la hora de evaluar las implicaciones estratégicas y operativas del constructo calidad del servicio percibida. De ahí que, en consonancia con lo que acabamos de exponer, propongamos el **Objetivo nº 5** que se enuncia como sigue:

“Analizar la relación existente entre las dimensiones del servicio, la calidad del servicio percibida, la satisfacción, la disposición a repetir la compra (rematricularse) y la disposición a recomendar los estudios”.

Como expresión del mismo definimos las siguientes hipótesis:

Hipótesis H5a: *Las dimensiones del servicio afectan positiva y significativamente a la calidad del servicio percibida.*

Hipótesis H5b: *Las dimensiones del servicio afectan positiva y significativamente a la satisfacción de los estudiantes.*

Por lo que se refiere, en particular, a la satisfacción, hay que decir que, si bien existe un consenso relativo en torno al hecho de que se trata de un constructo conceptualmente distinto de la CSP, tal y como ésta aparece definida en la literatura académica en la materia (Bitner, 1990; Bolton y Drew, 1991; Oliver, 1993; Spreng y Mackoy, 1996), la relación entre ambos constructos ha sido objeto de un intenso debate que intenta elucidar si la calidad del servicio constituye un antecedente de la satisfacción del consumidor o viceversa (Teas, 1993).

La respuesta clásica toma como punto de partida la idea de que el grado de satisfacción del consumidor con cada experiencia de servicio singular genera, a lo largo del tiempo, una actitud o evaluación general acerca de la calidad del servicio percibida (Oliver, 1980, 1981; Parasuraman *et al.*, 1988; Bitner, 1990). Desde este punto de vista, en la mente del consumidor la satisfacción es un precursor de la CSP, entendiendo ésta como una actitud, esto es, como una forma de evaluación general de la bondad de un producto o servicio, mientras que la satisfacción sería similar a la actitud, pero a corto plazo, ya que resultaría de la evaluación de una determinada experiencia de consumo (Bigné *et al.*, 2003). En los últimos años, empero, se ha asentado el punto de vista contrario. Oliver (1993) fue el primero en sugerir que la calidad del servicio podía ser un antecedente de la satisfacción del consumidor, con independencia de si este constructo era evaluado en relación con una experiencia particular de servicio o a lo largo del tiempo. Diversos estudios empíricos han confirmado con posterioridad este argumento (Anderson y

Sullivan, 1993; Spreng y Mackoy, 1996; Dabholkar *et al.*, 2000), consolidando así la idea de que la CSP es un antecedente de la satisfacción del consumidor.

Nos proponemos validar esta relación en el contexto de la educación superior en entornos virtuales de formación por medio de la siguiente hipótesis:

Hipótesis H5c: *La calidad del servicio percibida influye positiva y significativamente sobre el grado de satisfacción de los estudiantes.*

Existe una fuerte evidencia teórica y empírica que indica que los clientes habituales de una empresa no sólo realizan adquisiciones con mayor frecuencia y de importe superior, sino que, además, también son más resistentes a las promociones de los competidores y difunden más activamente una imagen positiva de aquella (Dick y Basu, 1994; Bolton, 1998; Rust *et al.*, 1995). Por otra parte, la CSP influye sobre las intenciones de compra de los consumidores (Reichheld y Sasser, 1990; Anderson y Sullivan, 1993; Brady y Robertson, 2001; Zeithaml *et al.*, 2004). No obstante, la mayor parte de los estudios evidencian que la relación de causalidad entre calidad del servicio percibida e intenciones de compra es indirecta, ya que parece estar mediada por otras variables como la satisfacción (Oliver, 1980; Dabholkar y Thorpe, 1994; Rust y Williams, 1994).³³⁸

En el caso de los servicios *online* la fidelidad de los consumidores resulta fundamental puesto que captar nuevos usuarios a través de Internet es un proceso costoso (Reichheld y Scheffer, 2000) y, por otra parte, los competidores se encuentran “a un sólo *click* de distancia”, lo que dificulta aún más un proceso de fidelización ya comprometido, de hecho, por la ausencia de interacción física directa el usuario y el prestador del servicio (Semeijn *et al.*, 2005, pág. 182). De ahí que resulte fundamental verificar si la relación entre calidad del servicio percibida e intenciones de compra observada en las transacciones en entornos tradicionales se produce también en los entornos *online* (Reichheld y Scheffer, 2000; Zeithaml *et al.*, 2000).

En general, los estudios que evalúan las repercusiones de la calidad del servicio y la satisfacción sobre la fidelidad y la disposición a recomendar el servicio de los consumidores *online* han sido escasos (Dabholkar, 1996; Bitner *et al.*, 2000; Meuter *et al.*, 2000) y sólo en los últimos años ha comenzado a ratificarse esta relación (Baker y

³³⁸ Boulding *et al.* (1993) encuentran una correlación positiva entre la calidad del servicio y la intención de recompra y el deseo de recomendar los servicios a otros consumidores potenciales. En una encuesta realizada a estudiantes universitarios se puso de manifiesto una fuerte correlación entre calidad del servicio y otros factores, tales como el deseo de referirse de forma positiva a la universidad, la voluntad de apoyar económicamente a la institución tras la graduación o el propósito de recomendar la institución a potenciales empleadores como un lugar idóneo para el reclutamiento de personal.

Crompton, 2000; Sivadas y Prewitt, 2000 y Zhu *et al.*, 2002).³³⁹ De ahí la relevancia de las siguientes hipótesis, que intentan validar esta relación en el ámbito de la formación superior en entornos virtuales:

Hipótesis H5d: *El grado de satisfacción de los estudiantes afecta positivamente a su fidelidad (disposición a re matricularse).*

Hipótesis H5e: *La fidelidad de los estudiantes influye positiva y significativamente sobre su disposición a recomendar los estudios.*

Hipótesis H5f: *El grado de satisfacción de los estudiantes ejerce una influencia positiva sobre su disposición a recomendar los estudios.*

7.3. Población y unidad de análisis

Teniendo en cuenta que el objetivo de esta investigación es identificar los elementos – atributos o dimensiones– determinantes del nivel de calidad del servicio percibida en el ámbito de la educación superior *online*, la población de referencia está formada por estudiantes de programas formativos universitarios cuyo proceso formativo se desarrolla en un entorno virtual de formación (EVF). La participación de los estudiantes de alguna de las instituciones seleccionadas inicialmente resultó imposible por razones relacionadas con la normativa en materia de protección de datos personales y/o con el deseo de evitar la saturación de los estudiantes, que suelen mostrar cierta resistencia a responder a las múltiples encuestas que se les remiten con propósitos diversos.³⁴⁰ Finalmente se consiguió la participación de seis universidades que ofrecen todos o algunos de sus programas en entornos virtuales de formación: la Universitat Oberta de Catalunya (UOC), la Universidad de Alicante, la Institución Universitaria CEIPA (Colombia), la Universidad Interamericana de Puerto Rico (Recinto Ponce), la Universidad Autónoma de San Luis de Potosí (Méjico) y la Universidad de Magdalena (Colombia).³⁴¹ No obstante, como veremos, por razones de representatividad estadística y

³³⁹ En esta línea, en sendos trabajos de reciente publicación, a cargo de Lee y Lin (2005) y de Semeijn *et al.* (2005) encuentran una relación positiva significativa entre calidad del servicio percibida, satisfacción del consumidor, intenciones de compra y entre aquellas y la disposición a recomendar un servicio *online*.

³⁴⁰ Es el caso, por ejemplo, de la Universidad de La Rioja, que sin embargo, cuenta con dos programas virtuales, Musicología y Ciencias del Trabajo.

³⁴¹ La participación de cada una de estas instituciones y el proceso para contactar con sus respectivos estudiantes fueron diversos, y serán indicados con posterioridad.

de congruencia del análisis que nos proponemos desarrollar, las hipótesis esenciales serán contrastadas considerando únicamente las opiniones de los estudiantes de la UOC.

7.4. Análisis cualitativo

7.4.1. Introducción

En esta fase de la investigación resulta esencial, en primer lugar, garantizar la validez de contenido del cuestionario, de forma que éste sea un reflejo adecuado y suficientemente representativo de los contenidos o dimensiones del concepto o fenómeno objeto de estudio (Sánchez y Sarabia, 1999; pág. 382). Además, en segundo lugar, antes de contrastar las hipótesis que acabamos de presentar es necesario determinar qué atributos, factores o dimensiones del servicio van a ser sometidos a la consideración de los estudiantes (Joseph y Joseph, 1997).

Por lo que se refiera a la primera cuestión, para obtener medidas con validez de contenido es preciso llevar a cabo una serie de actividades interrelacionadas que consisten en: a) especificar todo el dominio de contenido que es relevante para una determinada situación; b) especificar la muestra de características representativas del dominio especificado para la redacción de las preguntas y, por último, c) una vez que las características han sido seleccionadas, expresarlas de una forma tal que permitan ser comprendidas, medidas y contrastadas (Carmines y Zeller, 1994).

Para garantizar la validez de contenido resulta fundamental la revisión de la literatura académica relevante en la materia (Sánchez y Sarabia, 1999; pág. 382). Como se ha podido observar, dicha revisión –que llevamos a cabo en los capítulos precedentes– se realizó distinguiendo tres ámbitos distintos pero conexos: a) la teoría general en materia de calidad del servicio, b) la teoría en materia de calidad del servicio en entornos *online* y, por último, c) la calidad del servicio en la formación superior y, específicamente, en la formación superior *online*. En el curso de esa revisión teórica hemos comprobado que la mayor parte de los modelos diseñados para evaluar la CSP se basan, en buena medida, en el componente humano de la interacción del servicio –el denominado “encuentro de servicio”– (Sureschandar *et al.*, 2001). Sin embargo, es necesario descubrir qué sucede cuando los consumidores interactúan por medio de una interfaz tecnológica en lugar de hacerlo directamente (Parasuraman y Grewal, 2000a), como ocurre en los entornos virtuales de formación, y realizar las adaptaciones que resulten precisas en el modelo de evaluación de la CSP que se pretende diseñar.

En relación con la segunda cuestión resulta imprescindible desarrollar un análisis cualitativo previo que revele qué atributos, ítems o características son representativos de la calidad del servicio percibida desde el punto de vista de los estudiantes de programas de formación superior virtuales. Con este objetivo, se ha llevado a cabo un análisis exploratorio de carácter cualitativo empleando la técnica del incidente crítico. Esta técnica, utilizada con gran profusión en la evaluación de las fuentes subyacentes de satisfacción e insatisfacción de los consumidores de servicios (Bitner *et al.*, 1990; Edvarsson, 1988; 1992; 1998), apenas ha sido aplicada en el contexto de los e-servicios (Sweeney y Lapp, 2004), ámbito en el consideramos, no obstante, que, dadas sus características, presenta notables ventajas frente a otras técnicas cualitativas alternativas.

Así, hay que tener en cuenta, en primer lugar, que los estudiantes matriculados en programas virtuales no acostumbran a coincidir física y temporalmente, lo que complica la realización de grupos de discusión presenciales. Por otra parte, se trata de personas que suelen compatibilizar su actividad académica con otras actividades profesionales, por lo que resulta difícil conseguir que participen en grupos de discusión o entrevistas presenciales. En esas condiciones, además, puede aparecer un sesgo en el estudio, en la medida en que resulta previsible una mayor participación por parte de aquellos estudiantes que gozan de mayor disponibilidad y/o flexibilidad horaria –sin que sus preferencias constituyan necesariamente un reflejo adecuado de las del conjunto de los estudiantes–.

Por otra parte hay que tener en cuenta que las tecnologías de la información y la comunicación permiten aplicar procedimientos de encuesta por medios electrónicos que gozan de un nivel de eficacia similar, a un coste más reducido –para encuestador y encuestado– y que, además, permiten superar las clásicas limitaciones espacio-temporales asociadas a otros procedimientos análogos. De hecho, como veremos, al dirigir un mensaje al buzón personal de los estudiantes se hace posible la participación de alumnos de distintas universidades cuyos periodos vacacionales y lectivos no coinciden necesariamente y que pueden así enviar sus respuestas de forma asíncrona.³⁴²

³⁴² Inicialmente se pensó en realizar grupos de discusión virtuales, al menos con los estudiantes de la UOC, si los resultados obtenidos con el enfoque de los incidentes críticos no eran suficientemente clarificadores. Esta posibilidad fue desechada con posterioridad dados los resultados obtenidos con el anterior procedimiento.

7.4.2. El enfoque del incidente crítico

El método del incidente crítico, como técnica que agrupa un conjunto de procedimientos que permiten recoger, analizar y clasificar observaciones sobre el comportamiento humano fue introducido en el ámbito de las ciencias sociales hace ya más de cincuenta años por Flanagan (1954) y ha sido utilizado desde entonces en un amplio abanico de disciplinas. Stauss (1993) y Chell (1998) ofrecen dos de las revisiones más completas del concepto. En particular Chell (*op. cit.*, pág. 56) describe esta técnica como un método de naturaleza cualitativa que facilita la investigación acerca de acontecimientos –sucesos, incidentes, procesos o asuntos– identificados por el encuestado, acerca de cómo son gestionados y acerca de sus resultados o efectos percibidos con el objetivo de mejorar la comprensión del incidente desde la óptica del individuo.

Por incidente crítico se entiende, en este sentido, cualquier actividad humana observable que es lo bastante compleja como para permitir inferencias y predicciones por parte de la persona que la realiza o involucrada en la misma (Bitner *et al.*, 1990). Lo que la técnica del incidente crítico persigue es, en última instancia, conseguir una mejor comprensión de las dimensiones cognitiva, afectiva y conductual del fenómeno objeto de estudio (Bitner *et al.*, 1990; Grove y Fisk, 1997). Para conseguirlo se sigue, en lo sustancial, el procedimiento fijado originalmente por Flanagan (1954): se trata de recoger –mediante encuesta o un procedimiento análogo– un catálogo de incidentes críticos enunciados por sus propios protagonistas y, a continuación, analizar su contenido. Ese análisis involucra, en esencia, dos actividades distintas: elegir un marco de referencia en el que insertar los incidentes recogidos, y, en segundo lugar, desarrollar por mecanismos inductivos un conjunto de categorías y subcategorías que permitan clasificar dichos incidentes.

A la hora de llevar a cabo ese proceso de análisis y clasificación el investigador toma como punto de partida el propósito general del estudio y los esquemas y clasificaciones previamente establecidos en el área en cuestión (Neuhaus, 1996), procurando escrutar de forma precisa la información contenida en los incidentes para identificar categorías más generales que resulten relevantes y significativas –en el sentido de que den cuenta de la frecuencia y del tipo de factores que afectan al fenómeno objeto de estudio– (Grove y Fisk, 1997). Ello quiere decir, por tanto, que esas categorías pueden articularse, en función de las circunstancias de cada caso, recurriendo a los modelos teóricos preexistentes, por medio de un proceso de interpretación de corte inductivo o combinando ambos procedimientos (Gremier, 2004).

Entre las ventajas del método del incidente crítico acostumbra a mencionarse el hecho de que ofrece información recogida desde el punto de vista de los propios sujetos involucrados (Edvarsson, 1992). Ello permite que sean ellos mismos los que determinen libremente qué incidentes consideran más relevantes (Gabbott y Hogg, 1996), evitando así incurrir en el error de predeterminar de forma arbitraria qué es lo que resulta importante o no para el encuestado (Chell, 1998). Además, en la medida en que se les deja un amplio margen de libertad a la hora de exponer los incidentes, los encuestados lo hacen empleando su propio lenguaje y el nivel de detalle en la narración que consideran oportuno (Stauss y Weinlich, 1997), articulando así un discurso personal que constituye un reflejo de sus pensamientos y emociones (Stauss, 1993).

La naturaleza inductiva del procedimiento hace que resulte especialmente útil cuando el constructo objeto de investigación ha sido poco documentado (Grove y Fisk, 1997), como un método exploratorio que permite incrementar el conocimiento acerca de un determinado fenómeno o, por último, cuando se requiere un entendimiento omnicomprendivo y profundo de un fenómeno de cierta complejidad (Bitner *et al.*, 1990). Así, es especialmente apto cuando se trata –como en el caso que nos ocupa– de “desarrollar la estructura conceptual que luego va a ser empleada y evaluada en la investigación subsiguiente” (Gremler, 2004; pág. 67). De hecho, la naturaleza flexible del modelo, que no obliga a seguir un conjunto rígido de principios, permite su adaptación en función de las contingencias del objeto específico de estudio (Neuhaus, 1996) y le confiere particular interés cuando se trata de estudiar fenómenos para los cuales es difícil o imposible determinar *ex ante* todas las variables que es preciso tomar en consideración (Ruyter *et al.*, 1995).

Por otra parte, el hecho de que este procedimiento facilite información obtenida directamente de los relatos de las personas involucradas ayuda al investigador a comprender de forma más íntegra el conjunto de percepciones –cognitivas y emocionales– que el informante asocia con un determinado encuentro de servicio. Ello tiene un valor fundamental en el contexto de los servicios, por cuanto que en este ámbito los conceptos de satisfacción y calidad involucran un proceso psicológico complejo, imbricado de factores emocionales y cognitivos, por medio del cual el consumidor transforma sus percepciones en relación con los múltiples atributos de un servicio con un nivel determinado de satisfacción (Zeithaml y Bitner, 2003). En este sentido, los incidentes críticos permiten objetivar, de alguna forma, ese proceso psicológico –sin determinar *a priori* qué es importante o significativo– facilitando así a los gestores

información concreta y relevante (Stauss, 1993) que tiene importantes implicaciones desde el punto de vista de la gestión empresarial (Chell y Pittaway, 1998).³⁴³

Resumimos a continuación el procedimiento seguido para la generación y recolección de incidentes críticos y su ulterior clasificación:

a) Generación y recogida de incidentes críticos

El proceso de recopilación de la información relativa a los incidentes críticos consta de dos fases. En la primera se entrevista a los consumidores para obtener la información específica sobre el servicio estudiado –los incidentes críticos–, mientras que en la segunda esa información es clasificada en categorías que se pretende sean expresión de diferentes dimensiones de la calidad del servicio. En el caso que nos ocupa la primera fase fue realizada enviando un correo electrónico a un grupo de estudiantes virtuales. Desde el punto de vista teórico se recomienda que el número de consumidores entrevistados oscile entre 10 y 20. Ello permite, en el caso de que alguno de ellos proporcione una información falsa o errónea, contrastar su punto de vista con la información procedente del resto. Por lo que se refiere al contenido del cuestionario, el entrevistador debe solicitar que cada entrevistado describa entre 5 y 10 ejemplos positivos y otros tantos negativos de experiencias personales concretas y específicas relacionadas con el servicio (Hayes, 1999).

En nuestro caso se optó por enviar un mensaje solicitando la remisión de incidentes críticos (*vid.* Anexo II) a un número sustancialmente más elevado de estudiantes del teóricamente requerido, solicitando que indicasen unos 5 incidentes críticos positivos –y otros tantos negativos–. El procedimiento elegido se justifica por dos razones: en primer lugar, porque la tasa de respuesta en encuestas virtuales suele ser reducida –habitualmente se sitúa en torno a un 10%– y, en segundo lugar, porque resultaba más sencillo encuestar a un número superior de individuos que solicitar a cada estudiante que indique un número mayor de incidentes, ya que una encuesta demasiado exigente podía desincentivar la participación.

La población objeto de estudio está formada por estudiantes que cursan programas de formación *online* y el objetivo final era obtener un mínimo de 200 incidentes críticos, por

³⁴³ El empleo del método del incidente crítico no ha estado exento, empero, de críticas que cuestionan su nivel de validez y fiabilidad, en tanto que se entiende que los relatos obtenidos pueden ser mal entendidos o mal interpretados (Gabott y Hogg, 1996). Se trata, además, de un método de naturaleza retrospectiva, lo que puede generar sesgos indebidos, tanto en los propios relatos –si, por ejemplo, los encuestados reelaboran en exceso o reinterpretan el incidente al relatarlo–, como en el diseño de unas encuestas cuya extensión, por otra parte, hace que suela asociarse este método con una baja tasa de respuesta (Michel, 2001).

cuanto este es el nivel que se considera adecuado desde el punto de vista teórico (Hayes, 1999; pág. 32). Considerando que uno de los objetivos preliminares del estudio era contrastar si se observaban diferencias apreciables entre las percepciones de los estudiantes en función de factores tales como la institución en la que cursan sus estudios, el tipo de estudios cursados –grado o postgrado– y la cultura en la que se enmarcaba la institución, el mensaje fue remitido a estudiantes de cuatro instituciones diferentes: la Universitat Oberta de Catalunya (estudiantes de grado y de postgrado), la Universidad de Alicante (estudiantes de Criminología), la Institución Universitaria CEIPA (estudiantes de Administración y Dirección de Empresas) y la Universidad Interamericana de Puerto Rico, Recinto Ponce (UIPR Ponce) (estudiantes de Gerencia). Como el número de estudiantes matriculados en cada caso es muy diverso fue necesario ponderar la participación relativa de los estudiantes de cada una de esas instituciones. Así, se seleccionaron aleatoriamente 200 estudiantes de programas de grado y 100 de postgrado de la UOC, 20 estudiantes de la Universidad de Alicante, 20 de CEIPA y los 15 estudiantes de la UIPR Ponce, obteniendo un total de 23, 12, 2, 3 y 1 respuestas, respectivamente.³⁴⁴

Finalmente, el número total de participantes fue de 41 (21 hombres y 20 mujeres) que aportaron un total de 392 incidentes críticos, de los que doce se descartaron por falta de claridad en su exposición, quedando por tanto una muestra compuesta por 380 incidentes críticos válidos.³⁴⁵ A todos ellos se les envió un mensaje (*vid.* Anexo II) en el que se incluyeron ejemplos de incidentes críticos estándar con la finalidad de facilitar la generación de incidentes por parte de los estudiantes. Para no “contaminar” o sesgar en exceso las respuestas se decidió tomar como referencia incidentes “modelo” relacionados con un servicio diferente –un centro de salud–. El proceso se desarrolló durante el periodo de marzo-abril del 2004 y las respuestas obtenidas no permitieron detectar diferencias significativas entre los estudiantes de las diversas instituciones estudiadas, por lo que se decidió tratar la información recogida de forma conjunta.

b) Clasificación de los incidentes críticos

A continuación se explica de forma sucinta el método de análisis seguido para clasificar los incidentes recogidos. En primer lugar, se procedió a agrupar los incidentes en función

³⁴⁴ Aunque sólo se ha logrado la participación de un estudiante de la Universidad Interamericana de Puerto Rico, se ha decidido incluirlo en el análisis en atención a su naturaleza exploratoria.

³⁴⁵ En el caso de los estudiantes de la UOC, que son el colectivo más numeroso, una vez obtenidas las respuestas se comprobó que hubiera participación de estudiantes de los programas de las diversas áreas de conocimiento (Economía y Empresa, Informática y Multimedia, Psicología y Ciencias de la Educación, Ciencias de la Información y la Comunicación, Derecho y Humanidades).

del grado de similitud de su contenido (con independencia de que fueran positivos o negativos). A continuación se elaboró una frase descriptiva del contenido de los incidentes que integran cada grupo. Una vez agrupados todos los incidentes críticos se repitió el proceso anterior, utilizando ahora como criterio la similitud entre las frases descriptivas de los incidentes. Como resultado de este proceso se obtuvo una relación jerárquica entre los incidentes críticos, las descripciones agregadas de los mismos y, por último, las dimensiones de la calidad o categorías generales.

La clave de este proceso radica en los sucesivos procesos de agrupamiento. Uno de los métodos que permite controlar el rigor de este proceso consiste en asignar la responsabilidad de realizar esta tarea a dos personas diferentes que actúan como evaluadores: la primera establece cuáles son las dimensiones de la calidad siguiendo el procedimiento que acabamos de describir y la segunda, en función de las dimensiones establecidas, asigna directamente cada uno de los incidentes críticos a una de esas dimensiones. La calidad del proceso de reparto depende del grado de acuerdo existente entre ambos evaluadores, esto es, del porcentaje de incidentes que ambos sitúan en la misma dimensión. A medida que ese índice se acerca a la unidad, se pone de manifiesto un mayor grado de acuerdo. Un índice de aproximadamente 0,8 debería utilizarse como límite para determinar si las dimensiones son aceptables (Latham *et al.*, 1979), habiéndose obtenido en nuestro caso un índice del 0,94 por ciento de concordancia.

Se comprobó, por último, la complitud de las dimensiones consideradas a fin de confirmar que permiten definir en su totalidad el constructo calidad del servicio. Para ello se extrajeron, seleccionándolos al azar, aproximadamente el 10% de los incidentes críticos (35) y, a continuación, se repitió su agrupamiento y clasificación en dimensiones. Seguidamente se procedió a asignar los 35 incidentes críticos excluidos a estas dimensiones. El hecho de que fuera posible ubicar todos estos incidentes en las dimensiones predefinidas constituye una prueba de que éstas constituyen un buen reflejo del constructo que se pretende estudiar: la calidad del servicio en la formación superior *online*.

Tabla 7.1. Aplicación del método del incidente crítico

	Pasos	Puntos clave
1.	Generación y recopilación de los incidentes críticos	<ul style="list-style-type: none"> ✓ Envío de mensajes a los estudiantes solicitando incidentes críticos positivos y negativos. ✓ Recopilación de los incidentes críticos: ejemplos específicos que dan cuenta del nivel de calidad del servicio en relación con alguna dimensión concreta del mismo.
2.	Clasificación de los incidentes	<ul style="list-style-type: none"> ✓ La agrupación se debe basar en la similitud

	críticos en grupos	del contenido de los incidentes.
3.	Elaboración de definiciones significativas para cada grupo de incidentes críticos	✓ Cada definición ha de constituir una declaración asertiva y específica.
4.	Clasificación de las definiciones obtenidas en categorías	✓ La clasificación debe basarse en la similitud entre las definiciones. ✓ Las categorías deben ser descriptivas de deseos y/o necesidades de los clientes. De hecho, cada categoría representa una dimensión de la calidad.
5.	Validación de la calidad del proceso de clasificación	✓ Dos sujetos deben llevar a cabo el proceso de clasificación. ✓ El grado de acuerdo entre ellos debe alcanzar como mínimo el 80%.
6.	Comprobación de la complitud de las dimensiones consideradas	✓ Extracción del 10% de los incidentes críticos antes de establecer las dimensiones de la calidad. ✓ Verificación de que ese 10% puede ser colocado en las dimensiones preestablecidas.

Fuente: Adaptado de Hayes, 1999.

El proceso de categorización de los incidentes críticos (*vid.* Tabla 7.1) reveló, en primer lugar, la necesidad de excluir aquellos que no eran propiamente tales, ya que no contenían ejemplos o experiencias concretas del servicio recibido, sino que referían características que los estudiantes consideran propias de la formación *online* asíncrona.³⁴⁶ Concluido este proceso de refinamiento se obtuvieron 350 incidentes críticos *strictu sensu*, de los cuales 184 son positivos y 166 son negativos. Tras el procedimiento de clasificación realizado por los dos evaluadores y la comprobación de la validez y la complitud del proceso, estos incidentes críticos quedaron agrupados finalmente en 12 dimensiones que engloban un total de 33 definiciones (*vid.* tabla 7.2). A continuación se indican las características más importantes de las diferentes dimensiones:

Tabla 7.2. Clasificación de los incidentes críticos obtenidos

Dimensiones	Incidentes críticos					
	Positivo		Negativo		Total	
	N	%	N	%	N	%
Diseño y enfoque del programa	10	5,4	7	4,2	17	4,9
Itinerarios formativos	6	3,3	5	3,0	11	3,1
Enfoque del programa	4	2,2	2	1,2	6	1,7

³⁴⁶ Las afirmaciones excluidas tienen que ver, en general, con las ventajas de la formación *online* – ahorro de tiempo, disponibilidad del servicio, facilidad de acceso, posibilidad de compatibilizar estudios y trabajo– o con sus inconvenientes –sensación de aislamiento, necesidad de adaptación al entorno virtual–, pero enunciadas de forma genérica. En este sentido se observa que la posibilidad de acceder a los estudios y compatibilizarlos con la vida profesional y familiar es vista como muy positiva por una parte de la población que posiblemente no veía satisfecha esta necesidad formativa con anterioridad.

Materiales y recursos didácticos	26	14,1	21	12,7	47	13,4
Contenido	14	7,6	10	6,0	25	7,1
Formato	4	2,2	3	1,8	7	2,0
Biblioteca	8	4,3	8	4,8	16	4,6
Desarrollo del aprendizaje	34	18,5	35	21,1	69	19,7
Planificación del aprendizaje	4	2,2	3	1,8	7	2,0
Nivel de exigencia-dedicación requeridos	3	1,6	3	1,8	6	1,7
Orientación y seguimiento del proceso de aprendizaje por parte de los docentes	16	8,7	11	6,6	27	7,7
Contribución de las actividades al aprendizaje	10	5,4	3	1,8	13	3,7
Feedback de las actividades	1	0,5	15	9,0	16	4,6
Evaluación del aprendizaje	12	6,5	14	8,4	26	7,4
Sistema de evaluación	10	5,4	11	6,6	21	6,0
Gestión de la evaluación final presencial	2	1,1	3	1,8	5	1,4
Rapidez en las respuestas	23	12,5	13	7,8	36	10,3
Personal docente	13	7,1	4	2,4	17	4,9
Personal administrativo	8	4,3	8	4,8	16	4,6
Personal informático	2	1,1	1	0,6	3	0,9
Adecuación de las respuestas	10	5,4	16	9,6	26	7,4
Personal docente	5	2,7	1	0,6	6	1,7
Personal administrativo	4	2,2	13	7,8	17	4,9
Personal informático	1	0,5	2	1,2	3	0,9
Trato amable, personal accesible	17	9,2	4	2,4	21	6,0
Personal docente	7	3,8	2	1,2	9	2,6
Personal administrativo	10	5,4	1	0,6	11	3,1
Personal informático			1	0,6	1	0,3
Simplicidad trámites administrativos	5	2,7	11	6,6	16	4,6
Envío físico de la documentación	2	1,1	12	7,2	14	4,0
Docente	1	0,5	5	3,0	6	1,7
Administrativa	1	0,5	7	4,2	8	2,3
Relaciones con los compañeros	25	13,6	5	3,0	30	8,6
Interacción con otros compañeros	13	7,1	3	1,8	16	4,6
Ayuda en el aprendizaje	4	2,2			4	1,1
Disponibilidad de espacios de interacción	8	4,3	2	1,2	10	2,9
Interfaz del usuario	20	10,9	19	11,4	39	11,1
Apariencia estética	2	1,1	1	0,6	3	0,9
Conectividad	3	1,6	5	3,0	8	2,3
Navegabilidad	11	6,0	6	3,6	17	4,9
Solidez	3	1,6	6	3,6	9	2,6
Seguridad	1	0,5	1	0,6	2	0,6
Precio y compensaciones			9	5,4	9	2,6
Precio			7	4,2	7	2,0
Compensaciones			2	1,2	2	0,6
Total	184	100	166	100	350	100

La dimensión *diseño y enfoque del programa* representa aproximadamente el 5% del total de incidentes. Esta dimensión hace referencia a los itinerarios formativos ofrecidos por la institución y al enfoque del programa y de los diferentes cursos que lo conforman. En esta dimensión aparecen recogidos el 5,4% de los incidentes críticos positivos y el

4,2% de los negativos. Constituyen incidentes críticos positivos, por ejemplo, las siguientes afirmaciones: “los itinerarios formativos te permiten organizar tu aprendizaje según tus necesidades”, “los cursos se adaptan a la realidad empresarial actual”. Son incidentes negativos de este tipo los siguientes: “no existen itinerarios a seguir para completar un master tras los cursos de postgrado”, “el enfoque de los cursos es demasiado teórico”.

La dimensión *materiales y recursos didácticos* recoge el grado de satisfacción de los estudiantes en relación con los materiales y recursos didácticos. Los estudiantes han generado 47 incidentes críticos de este tipo (13,4% del total) de los cuales 26 son positivos y 21 negativos. Dentro de esta categoría se subsumen tres subcategorías. La primera de ellas está relacionada con el contenido de los materiales (“materiales actualizados”o “materiales superficiales”). La segunda alude al formato de los materiales (“disponibilidad de materiales en formato interactivo y en papel”, “el material se debería enviar en papel”). Por último, la tercera se refiere a la biblioteca, que parece tener notable relevancia para los estudiantes, puesto que han identificado ocho incidentes positivos y otros tantos negativos relacionados con la misma.

La dimensión *desarrollo del aprendizaje* engloba aquellas actividades que contribuyen directamente a facilitar el aprendizaje de los contenidos y el desarrollo de las competencias propias del programa. Dentro de esta dimensión se recogen aspectos relativos a la planificación del proceso de aprendizaje, al nivel de exigencia y de dedicación requeridos, a la contribución al aprendizaje de las diversas actividades realizadas a lo largo del periodo lectivo, al *feedback* obtenido tras la realización de dichas actividades y a la orientación, personalización, seguimiento y apoyo que el personal docente otorga a los estudiantes en su proceso de aprendizaje. Hay que destacar la importancia de esta dimensión, puesto que recoge casi el 20% de los incidentes críticos globales generados (el 18,5% de los positivos y el 21,1% de los negativos).

Tabla 7.3. Ejemplos de incidentes críticos relativos al desarrollo del aprendizaje

Desarrollo del aprendizaje	Incidentes positivos	Incidentes negativos
Planificación del aprendizaje	“Puedo saber la planificación de todo el curso desde el inicio, sin llevarme sustos de última hora”	“Cambios o incumplimientos en la planificación del curso (...)”
Nivel de exigencia-dedicación	“Requiere perseverancia y constancia”	“El ritmo de trabajo es elevado y cuesta seguir el curso cuando se trabaja”
Orientación y seguimiento	“Excelente orientación por parte de los profesores”	“Hay profesores que se limitan a reenviar las fechas del calendario, no a estimular o guiar a los estudiantes”
Contribución de las actividades	“Las actividades de debate son con las que más he aprendido”	“Algunas actividades suponen mucho trabajo de búsqueda en Internet (...), pasas más tiempo buscando información que estudiando”
Feedback de las actividades	“Feedback personalizado de las actividades”	“Raramente los consultores argumentan o comentan personalmente el contenido de las actividades que se presentan”

La dimensión *procedimiento de evaluación* agrupa el 7,4% del total de incidentes críticos. Los estudiantes valoran positivamente la existencia de un “sistema de evaluación continua” que “ayuda a estudiar y a aprender más fácilmente” y negativamente “el exceso de prácticas evaluables en algunas asignaturas”. Como en la mitad de las instituciones analizadas la evaluación final de las asignaturas es presencial, se considera como subcategoría la gestión de la evaluación final presencial (“posibilidad de elegir fecha de exámenes y pruebas de validación finales” o “la necesidad de desplazarse para realizar los exámenes finales”).

La dimensión *envío físico de la documentación* evalúa la calidad en la entrega de documentación a domicilio, ya se trate de materiales y otros recursos didácticos o de documentos administrativos como los títulos o certificados. Aunque el porcentaje total de incidentes relativos a esta dimensión no es muy elevado (14 en total), hay que destacar que la mayor parte de ellos (12) son, no obstante, negativos y aluden a retrasos en la entrega de la documentación.

Otros dos aspectos del servicio que se revelan como fundamentales son la *rapidez con la que los estudiantes obtienen respuestas ante las dudas, incidencias o problemas*, tanto las relacionadas con la actividad docente propiamente dicha, como las de tipo administrativo o informático. El 12,5% de los incidentes críticos positivos y el 7,8% de los negativos se refieren a esta dimensión. Además de la rapidez de respuesta también se evalúa la *correcta y adecuada resolución de las incidencias y dudas*. En este punto se observa un mayor porcentaje de incidentes negativos (9,6%) que positivos (5,4%), relativos, sobre todo, a las incidencias de carácter administrativo. También aparece un mayor

porcentaje de incidentes negativos (6,6% frente a 2,7%) en la categoría *simplicidad y personalización de los trámites administrativos* (por ejemplo, “problemas burocráticos con secretaría”).

Por contra, una de las dimensiones en las que hay un mayor porcentaje de incidentes positivos (9,2% frente a un 2,4%) es la relativa a la *amabilidad y facilidad de acceso al personal* de la institución, especialmente en relación con el personal administrativo (5,4%). Lo mismo ocurre (13,6% frente al 3%) con la dimensión relativa a las *relaciones con los compañeros*. Los estudiantes valoran muy positivamente la posibilidad de interactuar con otros compañeros “de otros lugares”, que son “profesionales”, con los que incluso surgen “nuevas amistades”, que además colaboran en el proceso de aprendizaje, generándose un “espíritu solidario”, así como la disponibilidad de espacios que faciliten esa interacción.

La categoría *interfaz del usuario* recoge aquellos aspectos del servicio relativos a la facilidad de uso y al funcionamiento técnico de la interfaz electrónica por medio de la cual se presta el servicio formativo. En concreto, incluye aspectos relativos a la apariencia estética de la interfaz, a la posibilidad de conectarse al *campus* con rapidez y en todo momento (conectividad), de cargar y descargar páginas y archivos sin errores (solidez) y a la seguridad de la interfaz. El mayor porcentaje de incidentes críticos positivos de esta dimensión hace referencia a la navegabilidad, es decir, a la sencillez e intuitividad de la navegación y a la facilidad para localizar la información que se busca dentro del *campus*.

Tabla 7.4. Ejemplos de incidentes relativos a la interfaz del usuario

Desarrollo del aprendizaje	Incidentes positivos	Incidentes negativos
Apariencia estética	“ <i>Campus</i> precioso”	“Espacio del aula poco gráfico, faltan colores”
Conectividad	“Acceso al <i>campus</i> perfecto, sin fallos”	“Problemas de conexión al <i>campus</i> , en determinados momentos y días”
Navegabilidad	“Es muy fácil navegar por el entorno virtual de aprendizaje”	“Es complejo navegar por la <i>web</i> ”
Solidez	“(…) funciona todo bien, el envío de ficheros, la descarga de materiales (...)”	“Errores constantes en el momento de descargar recursos, sobre todo, al inicio del curso y los domingos por la tarde”
Seguridad	“Perdí la clave de acceso y me pidieron muchos datos personales para garantizar mi identidad”	“El antivirus instalado dificulta la navegación”

En la categoría *precio y compensaciones* se recogen únicamente incidentes negativos (el 2,6% del total de incidentes). Así, los estudiantes perciben que “la formación es cara” y que no reciben ningún tipo de compensación si falla algún aspecto del servicio.

Remarcar finalmente, que, por orden de importancia –según el porcentaje de incidentes críticos generado en relación con cada una de ellas– las cuatro dimensiones básicas son las relativas al desarrollo del propio proceso de aprendizaje, a los materiales y recursos didácticos, a la interfaz del usuario y a la rapidez en las respuestas ante cualquier tipo de duda o incidencia.

7.5. Análisis cuantitativo

7.5.1. Introducción

Considerando los atributos generados en el proceso de análisis cualitativo que acabamos de exponer y tomando como referencia la estructura de la escala *e-SERVQUAL*³⁴⁷ se diseñó un cuestionario conformado por 33 ítems. Se realizó un pre-test del mismo enviándolo a diez estudiantes y, además, sometiéndolo a evaluación por parte de cinco profesores expertos en formación *online*. Tras el correspondiente proceso de refinamiento –en el que se valoró la claridad de la redacción de cada uno de los atributos, la complitud y extensión de la encuesta, la relevancia de los atributos y la posible redundancia entre el contenido de algunos de ellos, entre otros aspectos– se obtuvo finalmente un total de treinta ítems que se recogen en el Anexo IV.

7.5.2. Diseño del cuestionario

Como se puede observar en el Anexo IV, en la redacción de la encuesta se ha optado por enunciar de forma asertiva –evitando realizar afirmaciones “tajantes” con las que el encuestado haya de mostrar el acuerdo o el desacuerdo– los atributos obtenidos en la fase cualitativa de la investigación. En general se incluye un término y a continuación una pequeña aclaración sobre su significado. Si bien es cierto que este procedimiento puede condicionar de alguna forma las respuestas –guiándolas– se observó que este formato permitía a los estudiantes responder con mayor rapidez a la encuesta³⁴⁸ y, a la vez, que resultaba más fácil completarla, ya que se generaban menos dudas, aspectos

³⁴⁷ La relación de ítems que integra la escala fue amablemente remitida por el profesor Parasuraman vía correo electrónico antes de la publicación del artículo Parasuraman *et al.* (2004). “*e-SERVQUAL: A Multiple-Item Scale for Assessing Electronic Service Quality*”. *MSI Report*. Cambridge, Massachusetts: Marketing Science Institute.

³⁴⁸ Este formato favorece que cuando aparecen de nuevo los términos, en el segundo y el tercer bloque de la encuesta, los estudiantes ya saben a qué se hace referencia y se agiliza el proceso de respuesta.

ambos que, tratándose de una encuesta destinada a ser remitida por medio del correo electrónico, resultan fundamentales.

Teniendo en cuenta los objetivos de la investigación, y siguiendo el procedimiento ya utilizado, entre otros, por O'Neill *et al.* (2001) y Melián (2003) se solicitó a los estudiantes que indicasen, en relación con cada uno de los ítems: a) el nivel de calidad que esperan o consideran razonable obtener al matricularse en un curso de formación *online*, b) el nivel de calidad que consideran haber recibido en la institución en la que están matriculados y, por último, c) la importancia que atribuyen a cada uno de los atributos del servicio. Se utiliza a tal efecto (*vid.* Anexo IV) una escala Likert de cinco puntos, procedimiento ampliamente empleado en los estudios empíricos realizados en materia de calidad del servicio percibida en diversos contextos (*vid.* Tabla 5.2).

A continuación se incluyeron unas cuestiones de carácter global para averiguar la opinión de los estudiantes sobre: a) el nivel de calidad global de los servicios de la institución en la que están matriculados, b) su grado de satisfacción con estos servicios, c) su disposición a matricularse de nuevo en la institución y, por último, d) su disposición a recomendarla a un amigo o familiar. Además, se les pregunta si están satisfechos con la relación calidad-precio de la institución, se les pide que califiquen la evolución de la calidad global de los servicios y se solicita su opinión sobre cuál es, a su juicio, la reputación/imagen de la institución en la sociedad.

Por otro lado, y con el objetivo de caracterizar a los estudiantes se añadieron unas cuestiones generales de carácter clasificatorio. Así se les preguntaba si habían realizado previamente cursos de nivel universitario, si poseían alguna titulación universitaria, si el programa del que estaban matriculados era el primero que cursaban en esa institución y si habían cursado previamente formación *online* en otra institución. Además se les pedía información sobre el tipo de conexión que utilizan habitualmente y el lugar desde el que se conectan. Adicionalmente, se solicitaba que indicasen la motivación principal por la cual se matricularon en el programa formativo, así como su año de nacimiento, sexo y lugar de residencia habitual.³⁴⁹

Al final de la encuesta se añadió, por último, una cuestión abierta para que los estudiantes incluyeran aquellas observaciones o comentarios que estimasen oportunos (Aldridge y Rowley, 1998). Se trataba, en esencia, de ofrecerles la posibilidad de que

³⁴⁹ En el caso de la UOC no fue necesario formular estas cuestiones puesto que la encuesta se realizaba a través del número de identificación personal de los estudiantes, con lo que se podía conocer indirectamente su sexo, edad, lugar de residencia, el programa en el que estaban matriculados y el número de periodos lectivos que llevaban cursando dicho programa. Estos dos últimos aspectos también se conocían en el resto de instituciones, a través de la información que ellas mismas proporcionaban, sin necesidad de preguntarlo directamente a los estudiantes.

mencionasen otros atributos o características del servicio que siendo, a su juicio, importantes, no estuvieran adecuadamente reflejados en la encuesta. Ello permite, además, contrastar por un procedimiento alternativo la validez de contenido de la escala (Capelleras y Veciana, 2001).

Destacar, por último, que el formato de la encuesta enviada a los estudiantes de la UOC estuvo parcialmente condicionado por el *software* de elaboración de encuestas en formato *html*. Así, inicialmente se pensó que para acortar el tiempo de respuesta se podía pedir que el estudiante valorara simultáneamente, respecto a cada atributo, tanto las expectativas, como las percepciones y la importancia (Melián, 2003). Con este método se reducía el tiempo de respuesta –un factor a tener en cuenta, dada la longitud de la encuesta–, aunque, no obstante, ello podía inducir cierto nivel de “ruido” en las respuestas. Sin embargo, el *software* no permitía este diseño, por lo cual se optó por pedir que valorasen en primer lugar las expectativas en relación con todos los atributos. A continuación debían valorar el nivel de calidad del servicio percibido en relación con cada atributo y, por último, el nivel de importancia de cada uno de esos atributos.³⁵⁰

7.5.3. **Ámbito de análisis y proceso de envío de la encuesta**

Nuestra intención inicial era contar con las opiniones de estudiantes de programas de grado que desarrollaran su proceso de aprendizaje a través de un entorno virtual de aprendizaje y que pertenecieran a diferentes instituciones universitarias. Sin embargo, el todavía reducido número de instituciones que imparten programas de grado completa o mayoritariamente virtualizados, así como la dificultad para conseguir su participación, hizo que decidiéramos ampliar el ámbito de análisis para incluir también a instituciones que imparten programas de postgrado de este tipo. Finalmente fue posible contar con la colaboración de cinco de ellas: la Universitat Oberta de Catalunya, la Universidad de Alicante, la Universidad Autónoma de San Luis de Potosí (Méjico), la Institución Universitaria CEIPA y la Universidad del Magdalena (Colombia).

En el caso de la UOC el cuestionario fue enviado a los estudiantes a través del correo electrónico y directamente a su buzón de correo personal. A tal efecto se envió un mensaje (*vid.* Anexo III) desde un buzón de *Professorat Economia i Empresa*, y de *Profesorado de Economía y Empresa* en el caso de las titulaciones de grado en catalán y en castellano respectivamente. En el caso de los programas de postgrado se envió desde

³⁵⁰ Posteriormente, y para mantener la coherencia del estudio, la encuesta realizada en formato *word* mantuvo esta misma estructura.

un buzón de proyectos (*project_fp*). En todos los casos los estudiantes recibieron la encuesta en su buzón personal, accediendo directamente a ella, en formato *html*, a partir de un enlace contenido en el propio cuerpo del mensaje. Los diferentes bloques de la encuesta aparecían progresivamente a medida que los estudiantes iban cumplimentando sus distintos apartados. Si perdían la conexión o dejaban la encuesta “a medias” podían retomarla posteriormente en el punto en el que la habían dejado, ya que las respuestas iban almacenándose automáticamente en una base de datos.

Exceptuando el caso de la UOC, en el resto de las instituciones, los responsables de la gestión de las aulas virtuales fueron los que realizaron el envío del mensaje a los estudiantes animándoles a participar en la encuesta. En este caso se publicó el mensaje en un espacio común del aula virtual y no se envió al buzón personal del estudiante. La encuesta se remitía como un *word* adjunto al mensaje de presentación y los estudiantes debían enviar la encuesta cumplimentada a mi propio buzón personal o directamente a los responsables de la institución que con posterioridad las reenviaban a mi correo personal (ese fue el caso de la Universidad del Magdalena).

7.5.4. Participación y características de la muestra

En la Universidad de Alicante se ofrecen virtualmente dos programas de grado: el de Criminología y el de Detectives Privados. En el primero de ellos participan 118 estudiantes y en el segundo 67, de los cuales remiten la encuesta cumplimentada seis y dos estudiantes, respectivamente. De los 159 estudiantes del programa de grado virtual de *Administración y Dirección de Empresas* de la institución universitaria CEIPA, sólo se obtuvieron siete encuestas válidamente cumplimentadas.

En la Universidad Autónoma de San Luis de Potosí (Méjico) se imparte de manera virtual un programa de postgrado que permite la obtención del título de Diplomado virtual en “Flexibilidad curricular en educación superior”, cursado por 30 estudiantes, ocho de los cuales participan en la encuesta. Finalmente, en la Universidad del Magdalena (Colombia), de sus 250 estudiantes virtuales del Ciclo General para la obtención de grado, enviaron sus encuestas correctamente cumplimentadas un total de 16 estudiantes.

En el caso de la UOC, la encuesta se envió a todos los estudiantes que estaban en ese momento (Mayo de 2004) matriculados en dicha institución, tanto en programas de grado como de postgrado, excluyendo únicamente de partida a aquellos que expresamente habían manifestado su rechazo a recibir encuestas de cualquier tipo.

Inicialmente se pensó en remitirla a una muestra representativa de estudiantes, pero dada la baja tasa de respuesta propia de los estudios realizados a través del correo electrónico –confirmada en la fase inicial cualitativa– y teniendo en cuenta que el coste de gestión derivado de remitirla a toda la población de estudiantes era incluso inferior al que suponía enviarla únicamente a una muestra de ellos se optó por esta segunda posibilidad.

Dada la baja participación obtenida en el resto de instituciones universitarias y teniendo en cuenta los objetivos del análisis, se decidió finalmente excluirlas del proceso de análisis subsiguiente para asegurar un mayor grado de coherencia en el estudio y evitar el efecto dilución que se derivaría de procesarlas conjuntamente con las obtenidas de los estudiantes de la UOC. Además, la práctica totalidad de los estudios realizados hasta el momento con un propósito y una metodología similares al que aquí se propone –*vid.* tablas 5.2 y 5.3– se centran en el análisis de las percepciones de los estudiantes de una única institución universitaria.³⁵¹

A diferencia de lo que ocurre en los procesos de encuesta realizados de forma personal y procesados manualmente, la automatización del proceso hacía que cualquier estudiante que hubiera respondido a alguna pregunta de la encuesta –incluso únicamente a una– apareciera registrado como encuestado. Así, un porcentaje sustancial de los 4.478 estudiantes que habían contestado a una o más de las preguntas, no habían completado la encuesta en su totalidad o no habían respondido a una parte sustancial de la misma. Las características de la distribución de estas respuestas se observan en la tabla 7.5, en concreto, en la columna de *Inician encuesta*.³⁵²

Como se puede ver, siendo la moda de esta distribución 30, un porcentaje sustancial de los estudiantes abandonaba la encuesta una vez que había respondido al primer bloque de preguntas. De hecho, el 41,8% de los que la inician contestan a 30 preguntas o menos, pero fundamentalmente a 30 (el 27,9%). Ello quiere decir que resultaba necesario establecer un criterio que nos permitiera seleccionar a aquellos estudiantes que efectivamente habían respondido a la encuesta. Para hacerlo era preciso buscar un equilibrio entre la pérdida de información que supone no considerar las respuestas de determinados encuestados y los objetivos generales de la investigación, cumpliendo, además, los requisitos de fiabilidad y rigurosidad oportunos. Por otra parte, era imprescindible disponer de un nivel de información mínimo, suficiente e idóneo que

³⁵¹ No obstante, en el futuro se pretende realizar un análisis cualitativo de dicho datos teniendo en cuenta los resultados obtenidos en el presente trabajo.

³⁵² En total la encuesta está formada por 107 cuestiones, incluyendo la pregunta abierta, que ha sido excluida para la realización de estos cálculos.

permitiera estar en disposición de comparar los distintos modelos diseñados para evaluar la CSP (*SERVQUAL*, *SERVPERF* e *IPA*).

Tabla 7.5. Distribución de las respuestas obtenidas

Respuestas por estudiante		Inician encuesta	Conforman la muestra	
Respuestas por estudiante	Media	68,30	104,81	
	Mediana	72,5	106	
	Moda	30	106	
	Desviación típica	36,37	2,30	
	Percentiles:			
		25	30	105
		50	72,5	106
		75	106	106
Estudiantes	Total	4.478	2.108	

Tras realizar diferentes pruebas se optó por un criterio que garantizaba la respuesta a una parte sustancial de cada uno de los bloques de preguntas. En concreto se exigió que el encuestado hubiera respondido, al menos, al 80% de las cuestiones de cada uno de los bloques de “expectativas”, “percepciones” e “importancia”, o lo que es lo mismo, a 24 o más preguntas de cada uno de estos bloques. Además, y en relación con el resto de cuestiones, se estableció la condición mínima siguiente: el encuestado debía haber respondido a la cuestión relativa a la percepción de calidad global, al grado de satisfacción, a la disposición a rematricularse, a recomendar los estudios y a la satisfacción con la relación calidad-precio.

La aplicación de este criterio redujo el número de respuestas válidas a 2.108, pero, a cambio, permitió incrementar notablemente el número medio de respuestas por estudiante (que pasa de 68,3 a 104,81). De hecho, prácticamente todos los estudiantes seleccionados como parte de la muestra responden a todas las cuestiones (*vid.* Tabla 7.5). La restricción relativa a las cuestiones globales apenas si redujo el número de participantes.³⁵³ Al final del proceso se retuvo al 47,1% de los estudiantes que contestaron al menos una cuestión de la encuesta. En el cuadro siguiente queda reflejado cómo afecta este filtro a la selección de la muestra poblacional, globalmente considerada, y separando, por un lado, los estudiantes de grado y, por otro, los de postgrado.

³⁵³ De hecho, el exigir que hubiera una proporción de respuestas en el tercer bloque –referido a la importancia atribuida a los diferentes atributos– hacía que, en la práctica, aquellos que habían llegado hasta ese punto la contestaran, por lo general, en su totalidad.

Tabla 7.6. Selección de la muestra: distribución de los estudiantes

UOC	Población		Inician la encuesta		Conforman la muestra		
	Nº	% (sobre la población)	Nº	% (sobre la población)	Nº	% (sobre la población que inicia la encuesta)	% (sobre la población)
Grado	25.223	90,58	4.005	15,88	1.870	46,69	7,41
Postgrado	2.623	9,42	473	18,03	238	50,32	9,07
Total	27.846	100	4.478	16,08	2.108	47,07	7,57

De acuerdo con este criterio, la ficha técnica de la investigación contiene los atributos que se recogen en la siguiente tabla, destacando entre ellos que el error muestral global en el que se incurre es del 2,05%.

Tabla 7.7. Ficha técnica de la investigación

Ficha técnica de la investigación	
Procedimiento metodológico	Encuesta enviada a través del correo electrónico al buzón personal de cada estudiante.
Universo	Estudiantes matriculados en el periodo Febrero-Julio del año 2004 en cualquiera de los programas de grado y de postgrado de la UOC (exceptuando el programa de doctorado).
Población	27.846 estudiantes
Procedimiento de muestreo	Se intentó contactar con todos los estudiantes matriculados.
Tamaño de la muestra	2.108 estudiantes
Nivel de confianza	95% ($Z=1,96$; $p=q=50\%$)
Error muestral	2,05%
Periodo de realización del cualitativo	Marzo – Abril 2004
Periodo de realización del pre-test	Abril – Mayo 2004
Periodo de realización de la encuesta	Mayo – Junio 2004

Llegados a este punto se evaluó la idoneidad de considerar de manera conjunta toda la población de estudiantes, o por el contrario, la posibilidad de realizar el análisis de los estudiantes de grado y los de postgrado de manera diferenciada. En este sentido conviene tener en cuenta, desde un punto de vista teórico, que los programas de grado difieren de los de postgrado en cuanto a características tales como su duración, contenidos y enfoque. Por ejemplo, el diseño de las titulaciones de grado viene condicionado en gran medida por la troncalidad fijada por el Ministerio de Educación, aspecto éste que no tiene la misma relevancia en las titulaciones de postgrado, donde lo que se establece es el número de créditos que debe tener un programa para constituir, por ejemplo, un Postgrado o un Master.

Sin embargo, en la formación *online* el perfil de los estudiantes no difiere en exceso en función de que se trate de estudiantes de grado o de postgrado y, en cualquier caso, lo hace menos de lo que habitualmente acontece en la educación presencial de grado superior. Así, en ambos casos nos encontramos con estudiantes que mayoritariamente trabajan, tienen una edad superior a 25 años (el 89,5% de los estudiantes de grado y el 91,2% de los de postgrado) y que, además, han realizado con anterioridad algún curso universitario (un 81,1% de los encuestados de grado y un 90,3% de los de postgrado). No obstante, una de las diferencias fundamentales entre ambos perfiles radica en que entre los estudiantes de postgrado existe una mayor proporción de estudiantes que ya son titulados universitarios (un 86,5% frente a un 51,1%).

Para comprobar en qué medida las respuestas a las diferentes variables difieren de manera significativa entre un grupo de estudiantes y otro se realizó un análisis de comparación de medias a través del análisis ANOVA de un factor. De acuerdo con este procedimiento se comparan las respuestas medias aportadas por cada uno de los grupos de estudiantes a cada una de las 100 primeras preguntas que componen la encuesta.³⁵⁴ El principal resultado que se obtiene es que en 40 de las 100 respuestas analizadas se rechaza la hipótesis nula de igualdad de las puntuaciones medias otorgadas por los estudiantes de titulaciones de grado y de postgrado,³⁵⁵ siendo el nivel de confianza del 95%. Reseñar, además, que en el bloque de expectativas sólo difieren significativamente las respuestas en 8 de las 30 cuestiones, número que se incrementa hasta 10 en relación a la importancia y hasta 15 en el caso de las preguntas sobre percepciones. Además, se encuentran diferencias significativas en siete de las cuestiones genéricas, como son su grado de satisfacción global, su satisfacción con la relación calidad-precio, su disposición a rematricularse, su predisposición a recomendar y la reputación o imagen social percibida, entre otras.

Por otra parte, la reducida proporción de estudiantes de postgrado (9,42%) dentro de la población de estudiantes en su conjunto, implica que si se realizan análisis globales de la población, el comportamiento diferenciado de los estudiantes de postgrado queda diluido. A la luz de esa evidencia, teniendo en cuenta los resultados del análisis realizado y las diferencias teóricas existentes entre ambos tipos de programas, concluimos que resultaba más adecuado desarrollar un análisis diferenciado de ambos grupos,

³⁵⁴ En este punto no se consideran las últimas cuestiones de la encuesta, porque éstas son fundamentalmente clasificatorias, y no reflejan diferencias de opinión de los estudiantes.

³⁵⁵ Para garantizar la validez de los resultados obtenidos con el estadístico F se contrastó la hipótesis de homogeneidad de las varianzas, observando que ésta se incumplía en 25 de los 100 casos analizados. Por tanto, para garantizar la robustez de los resultados se calcularon los test de Welch y Brown-Forsyth.

distinguiendo, por tanto, los estudiantes de programas de grado (diplomaturas, segundos ciclos y licenciaturas) y, por otro, los de estudiantes de postgrado (programas de actualización, de especialización, postgrados y *masters*).

Esta decisión obligaba a establecer un criterio que permitiera solventar aquellas situaciones en las que un mismo estudiante cursaba simultáneamente programas de grado y de postgrado. El examen de la población reveló que 309 estudiantes estaban realizando más de un programa –ya sean dos programas de titulaciones de grado, dos postgrados o una combinación de un título homologado y un postgrado–.³⁵⁶ En principio, el hecho de cursar, por un lado, más de un programa de grado, o, por otro lado, más de un postgrado, no plantea especiales dificultades, ni distorsiona significativamente el análisis.³⁵⁷

Tabla 7.8. Error muestral para cada subpoblación: estudiantes de grado y postgrado

UOC	Población		Muestra		Error muestral
	Nº	% (sobre la población)	Nº	% (sobre la población)	P = Q = 50 Nivel de confianza del 95% (Z = 1,96)
Grado	25.223	90,58	1.870	7,41	2,18 %
Postgrado	2.623	9,42	238	9,07	6,06 %
Total	27.846	100	2.108	7,57	2,05%

El problema se planteaba en el caso de aquellos estudiantes que cursaban de forma simultánea algún postgrado y alguna titulación de grado. En esta circunstancia se encontraban 157 estudiantes de la población global de la UOC, de los que 14 forman parte de la muestra. Se decidió incorporar a este conjunto de estudiantes dentro del grupo de las titulaciones de grado, porque estos estudiantes recibieron con anterioridad la encuesta, en tanto que estudiantes de programas de grado y, de hecho, casi todos contestaron antes de haber recibido de nuevo el mensaje como estudiantes de postgrado, por lo que parece lógico que a la hora de realizar sus respuestas primara especialmente su idea de la calidad del servicio que reciben en tanto que estudiantes de titulaciones de grado, tratándose, además, de estudios que, por otra parte, suelen prolongarse más en el tiempo.

Teniendo en cuenta esta diferenciación los errores muestrales en los que se incurre son los que se indican en la tabla 7.8. Como se puede observar, el error muestral es

³⁵⁶ Había tres estudiantes de la población que estaban realizando más de dos programas de postgrado, si dos de los programas pertenecían a la misma área, eliminamos uno de ellos, y/o el que cursaban hacía menos tiempo.

³⁵⁷ En cualquier caso, todos los estudiantes que presentan estas duplicidades son correctamente identificados en la base de datos.

sustancialmente más elevado en relación con los estudiantes de postgrado (6,06%) que de grado (2,18%).

7.5.5. Representatividad de la muestra

Como no se realizó ningún procedimiento de muestreo previo, era necesario comprobar en qué medida la muestra seleccionada era representativa de la población objeto de estudio y que no se producían sesgos como consecuencia de seleccionar únicamente a aquellos estudiantes que voluntariamente han decidido cumplimentarla. Para ello se comparó, por un lado, la distribución de las características clasificatorias de la población y de la muestra seleccionada, y por otro lado, las contestaciones de los que responden a la encuesta en el día inicial de vigencia del periodo de respuesta y las de los que lo hacen al final del plazo establecido.

a) Características genéricas: población *versus* muestra

Para realizar esta comparación tomamos como referencia aquellas variables de las que disponíamos de información tanto a nivel poblacional como muestral. Son las siguientes: el sexo, la edad, el programa que cursan y los semestres que llevan cursando sus estudios en la UOC. En el caso de las titulaciones de grado, también se analiza el porcentaje de estudiantes que reciben la docencia en catalán *versus* castellano.

a1. El sexo

Tabla 7.9. Distribución de la población y las muestras por sexo

Sexo	Global				Grado				Postgrado			
	Población		Muestra		Población		Muestra		Población		Muestra	
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Fem.	12.742	45,8	1.042	49,4	11.608	46	937	50,1	1.134	43,2	105	44,1
Masc.	15.104	54,2	1.066	50,6	13.615	54	933	49,9	1.489	56,8	133	55,9
Total	27.846	100	2.108	100	25.223	100	1.870	100	2.623	100	238	100

En relación con el sexo, se puede observar (tabla 7.9.) que en la población existe un porcentaje superior de hombres que de mujeres (54,2% frente al 45,8%). Sin embargo, en la muestra seleccionada esa diferencia se ve ligeramente atenuada, ya que prácticamente responden el mismo número de hombres (1.066) que de mujeres (1.042). De ello se deduce que en la muestra, considerada globalmente, las mujeres están ligeramente sobre-representadas. Ello no ocurre sin embargo en el subconjunto de los programas de postgrado, ya que en éstos, la proporción de cada sexo en la muestra es prácticamente la misma que la poblacional. Por tanto, esta diferencia entre proporciones

se ve levemente acentuada en el caso de las titulaciones de grado, donde a pesar de estar cursando sus estudios unos 2.000 hombres más que mujeres, responden a la encuesta prácticamente el mismo número de mujeres que de hombres.

a2. La edad

Se establecieron cuatro tramos de edad que se muestran en la tabla 7.10. Como se observa, la mayor parte de los estudiantes se concentra en el tramo comprendido entre los 26 y los 45 años (el 80,3% de la población y el 77,4% de la muestra) y, en particular, en el que va de los 26 a los 35 años (el 54,2% de la población). El análisis comparativo de los datos muestrales y poblacionales evidencia que en la muestra están relativamente menos representados los estudiantes de menos de 35 años, tanto en el caso de las titulaciones de grado como de postgrado. De hecho estos estudiantes representan un 64,5% de la población y un 54,3% de la muestra.

Tabla 7.10. Distribución de la población y las muestras por edad

Edad	Global				Grado				Postgrado			
	Población		Muestra		Población		Muestra		Población		Muestra	
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Hasta 25	2.872	10,3	183	8,7	2.642	10,5	161	8,6	230	8,8	22	9,2
De 26 a 35	15.086	54,2	961	45,6	13.587	53,9	852	45,6	1.499	57,1	109	45,8
De 36 a 45	7.256	26,1	671	31,8	6.587	26,1	597	31,9	669	25,5	74	31,1
+ de 45	2.193	7,9	267	12,7	1.985	7,9	235	12,6	208	7,9	32	13,4
d.n.d*	439	1,6	26	1,2	422	1,7	25	1,3	17	0,6	1	0,4
Total	27.846	100	2.108	100	25.223	100	1.870	100	2.623	100	238	100

*Dato no disponible.

a3. El periodo lectivo

En las titulaciones de grado se oferta la docencia en periodos lectivos temporales de carácter semestral. De acuerdo con los datos registrados por la propia institución, podemos diferenciar entre los estudiantes que están cursando su primer semestre o periodo lectivo, su segundo semestre y, finalmente, los que están cursando un tercer semestre o superior. En relación con los programas de postgrado, en cambio, el periodo de docencia no suele adaptarse estrictamente al semestre docente de las titulaciones de grado, por lo que hemos adoptado la denominación periodo lectivo, por considerarla más adecuada.

En realidad cuando hablamos de periodo lectivo, nos estamos refiriendo al número de veces que los estudiantes han vivido un ciclo docente completo. Se entiende por ciclo docente el periodo que se inicia con la matrícula de unas asignaturas, continúa con el desarrollo del proceso de aprendizaje de los contenidos y de la adquisición de las

competencias propias de dichas asignaturas y su evaluación, y culmina con la obtención de una calificación definitiva. Así, al referirnos al primer periodo lectivo, se ha de considerar que los estudiantes aún no lo han cursado en su totalidad (excepción hecha de los estudiantes que no están cursando su primer programa), ya que cuando se realizó la encuesta los estudiantes aún no habían obtenido sus calificaciones definitivas. Además, como es lógico suponer, dadas las propias características de los diferentes programas, las titulaciones de grado suponen más periodos lectivos que los programas de postgrado.

Tabla 7.11. Distribución de la población y las muestras por periodo lectivo

Periodo lectivo	Global				Grado				Postgrado			
	Población		Muestra		Población		Muestra		Población		Muestra	
	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
Primero	6.786	24,4	509	24,2	5.561	22,1	419	22,4	1.225	46,7	90	37,8
Segundo	4.562	16,4	399	18,9	3.404	13,5	280	15	1.158	44,2	119	50
3º o sup.	16.472	59,1	1.200	56,9	16.254	64,4	1.171	62,6	218	8,3	29	12,2
Otros*	26	0,1	-	-	4	0,0	-	-	22	0,8	-	-
Total	27.846	100	2.108	100	25.223	100	1.870	100	2.623	100	238	100

*Aquí se incluyen estudiantes que están realizando el proyecto final de algún curso.

A escala global se observa que los porcentajes muestrales y poblacionales de estudiantes de cada uno de los periodos lectivos que se han considerado son muy similares. Sólo aparecen ligeramente sobre-representados en la muestra los estudiantes de tercer periodo lectivo y superior frente a los del segundo periodo. En el caso de las titulaciones de grado los porcentajes de estudiantes que cursan cada uno de los periodos lectivos son prácticamente iguales a nivel poblacional y muestral. Sin embargo, en el caso de los programas de postgrado en la muestra están levemente más representados los estudiantes de segundo periodo lectivo y superior frente a los del primer periodo.

a4. El programa

Con el objetivo de analizar el grado de homogeneidad entre la composición de la muestra y de la población en relación con cada uno de los programas impartidos, se analizó también, por un lado, el porcentaje de estudiantes de cada uno de los programas de titulaciones de grado y, por otro lado, el porcentaje de estudiantes de cada una de las áreas de postgrado. En relación con las titulaciones de grado se realizó simultáneamente el análisis tanto en términos de programa como de estudios. En general, los porcentajes de estudiantes por cada tipo de programa a nivel poblacional son muy similares a los muestrales. No obstante, cabe destacar que los programas de los Estudios de Informática están menos representados a nivel muestral de lo que lo están a nivel poblacional.

Asimismo y *a sensu contrario*, la Licenciatura en Psicología, el 2º ciclo de Documentación y la diplomatura en Turismo representan porcentajes más elevados a escala muestral.

Tabla 7.12. Distribución de la población y la muestra de grado por programa y estudio

Programa	Grado			
	Población		Muestra	
	Nº	%	Nº	%
Estudios de Economía y Empresa	9.569	37,7	708	37,5
Diplomatura en CC. Empresariales	5.917	23,3	430	22,8
2º ciclo de Administración y Dirección de Empresas	1.259	5	86	4,6
2º ciclo de Ciencias del Trabajo	1.899	7,5	146	7,7
2º ciclo de Investigación y Técnicas de Mercado	494	1,9	46	2,4
Estudios de Psicología y Ciencias de la Educación	3.416	13,5	285	15,1
2º ciclo de Psicopedagogía	1.176	4,6	94	5
Licenciatura en Psicología	2.240	8,8	191	10,1
Estudios de Derecho y Ciencia Política	2.557	10,1	181	9,5
Licenciatura en Derecho	2.345	9,2	165	8,7
2º ciclo de Ciencias Políticas y de la Administración	212	0,8	16	0,8
Estudios de Humanidades y Filología	1.935	7,6	155	8,2
2º ciclo en Humanidades	1.581	6,2	128	6,8
Licenciatura en Filología catalana	354	1,4	27	1,4
Estudios de Informática y Multimedia	5.749	22,7	342	18,1
Ingeniería Técnica en Informática de Gestión	1.837	7,2	126	6,7
Ingeniería Técnica en Informática de Sistemas	2.730	10,8	140	7,4
2º ciclo de Ingeniería en Informática	568	2,2	34	1,8
Graduado Multimedia (<i>título propio</i>)	614	2,4	42	2,2
Estudios de Ciencias de la Información y la Comunicación	1.364	5,4	126	6,7
2º ciclo de Documentación	1.059	4,2	103	5,5
2º ciclo de Comunicación audiovisual	305	1,2	23	1,2
Otros ³⁵⁸	771	3	92	4,9
Diplomatura en Turismo	523	2,1	67	3,6
Licenciatura en Estudios sobre el Asia Oriental	248	1	25	1,3
Total	25.361 ³⁵⁹	100	1.889	100

Nota: Algunas sumas parciales pueden no coincidir como consecuencia del redondeo de los decimales.

En relación con los estudiantes de Postgrado, se mantiene la coherencia con el criterio adoptado para la división de la muestra en estudiantes de programas de grado y de

³⁵⁸ Se ha agrupado la Diplomatura en Turismo y la Licenciatura en Estudios de Asia Oriental porque en términos organizativos presentan un alto grado de similitud y porque individualmente consideradas no representan un número muy elevado de estudiantes.

³⁵⁹ El total poblacional es superior a los 25.233 estudiantes considerados con anterioridad como estudiantes que cursan titulaciones de grado, lo que pone de manifiesto precisamente la existencia de duplicidades asociadas al hecho de que algunos estudiantes cursan más de un programa. El total muestral es también superior a los 1.870 estudiantes, porque en la muestra aparecen recogidas algunas de estas duplicidades.

postgrado.³⁶⁰ Además, tal y como ya hemos comentado con anterioridad, en este caso resulta más apropiado hablar de áreas temáticas que de programas.

Los porcentajes que se exhiben en la tabla 7.13. evidencian ciertas diferencias entre la estructura de los estudiantes de la muestra y de la población en función del área temática en la que se encuadra el programa que están cursando. Así, el área de cultura y algunas áreas de empresa –especialmente finanzas, calidad y logística– están infra-representadas a nivel muestral, mientras que el área de *e-learning* está sobre-representada.

Tabla 7.13. Distribución de la población y la muestra de postgrado por áreas temáticas

Áreas temáticas	Postgrados			
	Población		Muestra	
	Nº	%	Nº	%
Área de Humanidades	366	17,6	30	17
Cultura	198	7,5	13	5,4
Resolución de conflictos	138	5,2	14	5,8
Salud	30	1,1	3	1,2
Periodismo	98	3,7	11	4,6
Área de Empresa	1.239	47	99	41,1
Calidad	148	5,6	9	3,7
Finanzas	327	12,4	23	9,5
Logística	101	3,8	4	1,7
Marketing	205	7,8	16	6,6
Recursos Humanos	253	9,6	26	10,8
Prevención de riesgos laborales	170	6,4	16	6,6
Medio ambiente	35	1,3	5	2,1
Área de Informática y TIC	758	28,7	73	30,3
Informática	513	19,5	49	20,3
Tecnologías de la Información y la Comunicación	155	5,9	17	7,1
Microsoft	90	3,4	7	2,9
Área de <i>e-learning</i>	162	6,1	28	11,6
Proyecto	14	0,5	-	-
Total	2.637	100	241	100

a5. El idioma

Por último, y en relación con las titulaciones de grado, se analizó la proporción de estudiantes de la muestra que desarrolla su aprendizaje en catalán y en castellano. El sentido de este análisis está relacionado, en primer lugar, con la evidencia de que la UOC tiene una mayor implantación en Cataluña que en el resto de España y, en

³⁶⁰ A la hora de determinar los estudiantes que conforman la población de Postgrado se seleccionó a aquellos que cursaban únicamente programas de Postgrado, ya sea uno o más de uno. Esto determina una población de 2.623 estudiantes, alguno de los cuales simultaneaba la realización de más de un programa de Postgrado.

segundo lugar, con el hecho de que el precio que los estudiantes pagan por sus estudios es diferente, al estar subvencionada la formación en catalán, lo que, *prima facie*, podría influir sobre las percepciones de los estudiantes. En este sentido, la Tabla 7.14. evidencia que en la muestra están ligeramente sobre-representados los estudiantes que reciben la docencia en castellano, ya que si bien representan únicamente un 19,1% de la población de estudiantes de titulaciones de grado, suponen un 24,5% de la muestra.

Tabla 7.14. Distribución de la población y la muestra de grado por idioma

Docencia	Grado			
	Población		Muestra	
	Nº	%	Nº	%
Catalán	20.412	80,9	1.411	75,5
Castellano	4.811	19,1	459	24,5
Total	25.223	100	1.870	100

En conclusión, por tanto, podemos afirmar que, a pesar de las diferencias reseñadas anteriormente, en general, la estructura y el perfil del conjunto de individuos que componen la muestra son similares a los del conjunto de la población. No obstante, las diferencias identificadas previamente han de ser adecuadamente ponderadas a la hora de extraer conclusiones a partir de la muestra objeto de estudio que resulten generalizables al conjunto de la población global o de cada uno de los subconjuntos objeto de análisis –titulaciones de grado y postgrado–.

b) Características respuestas: primeras *versus* últimas

Las conclusiones de los estudios realizados utilizando encuestas pueden presentar sesgos derivados de la no participación del grupo de individuos que no responde a la encuesta (Bailey, 1978; Straub, 1989). Para evitarlo, siguiendo las recomendaciones de Armstrong y Overton (1977), comparamos las respuestas más tempranas con las más tardías, a fin de detectar si existen o no diferencias significativas entre ellas. Como las encuestas no fueron enviadas en su totalidad en la misma fecha, resulta adecuado analizar, por un lado, las titulaciones de grado (diferenciando en ellas, los estudiantes que cursan sus estudios en castellano y los que lo hacen en catalán) y, por otro, los programas de postgrado. Para efectuar el análisis se tomó como referencia las respuestas aportadas a las 100 primeras cuestiones de la encuesta.

En el caso de las titulaciones de grado, se compararon las cien primeras respuestas obtenidas con las cien últimas obtenidas en catalán, por un lado, y en castellano, por otro. En relación con los estudiantes que reciben la docencia en catalán, se compararon las primeras cien respuestas registradas el mismo día de envío (el día 21 de Mayo) con

las cien últimas (recibidas entre el 26 y el 31 de Mayo, inclusive). Los resultados del análisis de diferencias entre medias (ANOVA de un factor) indican que no se puede rechazar la hipótesis nula de igualdad de medias en 92 de las 100 variables analizadas, siendo el nivel de confianza del 95%. Este análisis se replicó con los estudiantes de titulaciones de grado en castellano (se comparan las 100 primeras respuestas recibidas el día 28 de Mayo con las 100 últimas, recibidas entre el 3 y el 7 de Junio). Los resultados obtenidos son similares: en este caso no se rechaza la hipótesis nula en el 94% de las variables analizadas.

En relación con los estudios de postgrado, comparamos las cincuenta primeras respuestas³⁶¹ obtenidas el mismo día de envío de la encuesta (28 de Mayo), con las cincuenta últimas, obtenidas entre los días 2 y 7 de Junio, inclusive. Tras la aplicación del análisis de medias (ANOVA), se observa que no hay diferencias significativas – siendo el nivel de significación del 95%– entre las medias de las respuestas del primer grupo de estudiantes y las del segundo grupo, en relación con todas y cada una de las variables encuestadas, a excepción únicamente de la pregunta relativa a la importancia atribuida a la biblioteca.

Como conclusión podemos afirmar, por tanto, que no existen diferencias significativas entre las respuestas más tempranas y más tardías, excepto en un porcentaje muy reducido de las cuestiones planteadas. Además, en ninguno de los casos se observa que estas diferencias afecten a las cuestiones de carácter más global como la percepción de la calidad global, el grado de satisfacción con los servicios, la disposición a rematricularse o a recomendar los estudios, la satisfacción con la relación calidad-precio y la reputación o imagen social percibida de la institución. Por tanto, de acuerdo con el criterio establecido por Armstrong y Overton (1977), podemos afirmar que no existe un sesgo en las conclusiones asociado al hecho de que sólo se tengan en cuenta las respuestas de aquellos estudiantes que voluntariamente deciden participar en la encuesta.

³⁶¹ En este caso, se compara un menor número de respuestas, dado que el número de estudiantes de postgrado que participa en la encuesta es también sustancialmente más reducido.