



I Congreso Internacional de Ciencias Exactas y Naturales

Editado por
Yuri Morales López



Universidad Nacional
Costa Rica, 2019.



I Congreso Internacional de Ciencias Exactas
y Naturales/ Yuri Morales-López –Heredia,
Costa Rica: Universidad Nacional, 2019.

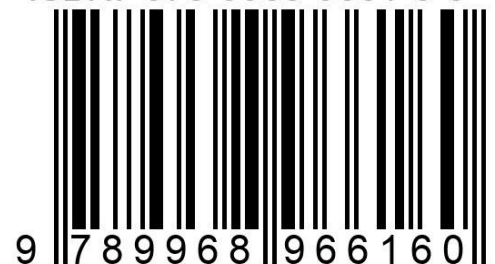
ISBN: 978-9968-9661-6-0.

- Este documento y el contenido tienen una Licencia de uso tipo CC: BY-NC-ND 4.0.
- El uso de texto, imágenes y otra información de terceros es responsabilidad plena de cada autor en su respectivo trabajo, y asumen completa responsabilidad sobre cualquier reclamo legal.
- Las opiniones expresadas en este documento son responsabilidad de los autores y no necesariamente representan la opinión de los editores ni de la Universidad Nacional.

Reconocimiento

Se les agradece profundamente a la Bachiller Evelyn Rojas Ramírez y al Máster Luis Ocampo Venegas por el apoyo para la gestión de este documento.

ISBN: 978-9968-9661-6-0



Evaluación de usabilidad para un sitio de comercio electrónico: Desarrollo de una metodología y su aplicación al sitio crgourmetcoffee.com

Francisco J. Mata

fmata@una.cr

Escuela de Informática
Universidad Nacional
Costa Rica

Irene Hernández-Ruiz

irene.hernandez.ruiz@una.cr

Escuela de Informática
Universidad Nacional
Costa Rica

Resumen

Se presenta en este trabajo una metodología para evaluar la usabilidad de un sitio web y su aplicación a un sitio de comercio electrónico en desarrollo. Esta metodología utiliza la escala de usabilidad de un sistema (denotada en inglés por las siglas SUS), la cual se justifica con base en la literatura existente. Dicha metodología se aplicó al sitio de comercio electrónico crgourmetcoffee.com. Como resultado de la aplicación de la metodología propuesta, se encontró que la usabilidad del sitio es buena. Asimismo, los comentarios obtenidos para los ítems que recibieron menor evaluación se han utilizado de base para mejorar dicha usabilidad.

Palabras clave: evaluación de la usabilidad; comercio electrónico

Abstract

This paper presents a methodology to assess the usability of a website and its application to an e-commerce site that is being developed. This methodology uses the usability scale of a system (SUS), which is justified based on the existing literature. This methodology was applied to crgourmetcoffee.com (e-commerce site). As a result of the application of the proposed methodology, it was found that the usability of the site is good. Furthermore, comments obtained—in relation to the items with lower scores—has been used as basis to improve such usability.

Keywords: usability evaluation; electronic commerce

Tema: Desarrollo y aplicaciones científicas y tecnológicas

Principal área: Informática

Mata, F.J. y Hernández-Ruiz, I. (2019). Evaluación de usabilidad para un sitio de comercio electrónico: Desarrollo de una metodología y su aplicación al sitio crgourmetcoffee.com. En Y. Morales-López (Ed.), *Memorias del I Congreso Internacional de Ciencias Exactas y Naturales de la Universidad Nacional, Costa Rica, 2019* (e62, pp. 1-10). Heredia: Universidad Nacional. doi <http://dx.doi.org/10.15359/cicen.1.9>

Introducción

El comercio electrónico (EC) se refiere al uso de Internet y otras redes (por ejemplo, intranets) para comprar, vender, transportar o comercializar datos, bienes o servicios (Turban, 2018).

Como parte de un proyecto de investigación, se está desarrollando el sitio de comercio electrónico *cr.gourmetcoffee.com*. Dicho sistema está basado en el modelo de comercio electrónico denominado “business to consumer” (B2C), el cual incluye transacciones de productos o servicios desde empresas a compradores individuales.

Debido a sus características, el desarrollo del mencionado sitio se está realizando desde cero —utilizando el “framework” Symphony, fundamentado en el lenguaje de programación web PHP, el cual permite utilizar el patrón de arquitectura de software conocido como modelo-vista-controlador.

A la fecha se cuenta con un prototipo del mencionado sistema y con el propósito de convertirlo en una plataforma funcional de comercio electrónico fue necesario realizar una prueba de usabilidad. La usabilidad es un factor crítico de éxito para cualquier sitio web (Nielsen & Landauer, 1993).

Este trabajo presenta la metodología desarrollada para llevar a cabo el estudio de usabilidad, antes mencionado, y los resultados de su aplicación al prototipo en cuestión.

Marco teórico

Como se mencionó anteriormente, uno de los aspectos importantes para el éxito de un sitio de comercio electrónico es su usabilidad, la cual se encuentra dentro del área de interacción humano computador (HCI, por sus siglas en inglés). El término usabilidad se utiliza comúnmente para referirse a la facilidad de uso de un sistema computacional (ver Nielsen y Landauer (1993)).

En este sentido, varios estudios sobre usabilidad (por ejemplo, Virzi (1992), Nielsen y Landauer (1993) y Lewis (1994)) han concluido que la mayoría de los problemas de usabilidad se descubren a través de los primeros tres a cinco usuarios y que el incremento para encontrar problemas después de cinco participantes es mínimo. Esto ha llevado a proponer la regla de $4+1$ para calcular el número de individuos en una muestra para realizar un estudio de usabilidad (Bevan, et al., 2003).

La usabilidad de un sistema es un asunto de percepción por parte de los usuarios, por lo cual es común en este tipo de estudios utilizar escalas con varias preguntas (ítems), cada una orientada a evaluar un aspecto que se considera importante para su usabilidad. Dentro de estas escalas, la escala de usabilidad de un sistema (SUS, por sus siglas en inglés) —propuesta por Brooke (1996) — es una de las más utilizadas (Bangor et al, 2008, 2009; Brooke, 2013).

Según Bangor et al. (2009) tres características hacen atractivo el uso de la SUS. Primero, al tener esta escala cuenta solo diez preguntas, facilita a los participantes completar la evaluación y a los investigadores obtener un resultado. Segundo, a diferencia de otras escalas de usabilidad, no es propietaria y, por lo tanto, no hay que pagar por su uso. Tercero, la escala es independiente de la tecnología utilizada, y puede ser utilizada para



evaluar una gran variedad de interfaces. Asimismo, Tullis y Stetson (2004) demuestran que al usar la SUS se obtiene mayor exactitud que con otras evaluaciones de usabilidad, tales como “Questionnaire for User Interface Satisfaction” (QUIS) —desarrollada por la Universidad de Maryland—, “Computer System Usability Questionnaire” (CSUQ) —desarrollada por IBM—, Words —adaptada de Product Reaction Cards de Microsoft—, y Our Questionnaire —utilizado en pruebas de usabilidad de sitios web.

Metodología

Para evaluar la usabilidad de *crgourmetcoffee.com* se utilizó una metodología de tipo cuantitativo con base en la SUS. En este sentido, Brooke (1996) propone los siguientes 10 ítems para la SUS:

1. Creo que usaría este [sistema, objeto, dispositivo, aplicación] frecuentemente;
2. Encuentro este [sistema, objeto, dispositivo, aplicación] innecesariamente complejo;
3. Creo que el [sistema, objeto, dispositivo, aplicación] fue fácil de usar;
4. Creo que necesitaría ayuda de una persona con conocimientos técnicos para usar este [sistema, objeto, dispositivo, aplicación];
5. Las funciones de este [sistema, objeto, dispositivo, aplicación] están bien integradas;
6. Creo que el [sistema, objeto, dispositivo, aplicación] es muy inconsistente;
7. Imagino que la mayoría de la gente aprendería a usar este [sistema, objeto, dispositivo, aplicación] en forma muy rápida;
8. Encuentro que el [sistema, objeto, dispositivo, aplicación] es muy difícil de usar;
9. Me siento confiado al usar este [sistema, objeto, dispositivo, aplicación];
10. Necesité aprender muchas cosas antes de ser capaz de usar este [sistema, objeto, dispositivo, aplicación].

Cada uno de estos ítems se evaluó de acuerdo con una escala Likert de opinión, la cual cuenta con 5 puntos, de los cuales 1 indica estar en total desacuerdo y 5 estar en total acuerdo. Con base en los resultados obtenidos para cada ítem, se procedió a aplicar el siguiente procedimiento:

- a) Restar 1 al valor obtenido para cada uno de ítems impares (1, 3, 5, 7 y 9) para obtener su contribución a la escala;
- b) La contribución a la escala para los ítems pares (2, 4, 6, 8, 6, 10) se calculó como 5 menos el valor obtenido para cada uno de los mismos; esto es debido a que las preguntas correspondientes están redactadas en sentido inverso a las impares;
- c) Posteriormente, se sumó las contribuciones correspondientes a los ítems de la escala para cada uno de los participantes y se multiplicó el resultado de esta sumatoria por 2.5 para obtener la puntuación final, la cual se encuentra en un rango de 0 a 100.

Para llevar a cabo la evaluación se utilizó un cuestionario electrónico, desarrollado



utilizando Google Forms (forms.google.com). En la primera sección del cuestionario se preguntó a los participantes por el grupo o estrato al cual pertenecen (profesores o estudiantes) y por su nivel de experiencia en el uso del comercio electrónico. En la segunda sección se solicitó evaluar los 10 ítems de la SUS, según lo antes explicado. Finalmente, en las secciones siguientes se solicitaron comentarios y recomendaciones sobre el diseño y funcionamiento del sistema.

Para la evaluación de *crsgourmetcoffee.com* se utilizó a profesores y estudiantes de la Universidad Nacional. Debido a su experiencia en el uso del comercio electrónico, se consideró que estos participantes, no solo podrían tener interés en un sitio de esta naturaleza, sino que también estarían en mayor capacidad de realizar su evaluación, en comparación a personas que no tuvieran ninguna experiencia con el comercio electrónico.

La evaluación se llevó a cabo del 26 de noviembre al 4 de diciembre del 2018. Se seleccionó este periodo debido a que tanto los estudiantes como los profesores estarían sin responsabilidades académicas, y así podrían contar con tiempo para realizar la evaluación.

Como población para la evaluación se consideraron potenciales usuarios del sistema. Con este fin, se seleccionó una muestra no aleatoria (a juicio) de 67 estudiantes y profesores de la Universidad Nacional, los cuales se dividieron de acuerdo con 4 estratos: i) estudiantes de pregrado/grado, ii) estudiantes de posgrado, iii) profesores, y iv) otros. Los estudiantes de pregrado se seleccionaron por haber matriculado el curso Uso Estratégico del Comercio Electrónico en el I y II ciclos del 2018 (23 estudiantes). Los otros en este grupo correspondieron a estudiantes colaboradores en actividades de la Escuela de Informática (11), relacionados con los investigadores. Los estudiantes de posgrado correspondieron a estudiantes de la VIII (2017) y IX (2018), promociones de la Maestría en Gestión de Innovación Tecnológica (MAGIT), las cuales aún cuentan con estudiantes activos, quienes tienen diferentes formaciones, algunos de ellos informáticos. En el caso de los profesores, la mayoría (10) correspondió a profesores de la Escuela de Informática. La categoría de otros se estableció para aquellos que no quisieran identificarse en alguna de los estratos anteriores.

Resultados

a. *Muestra obtenida*

Se obtuvo respuesta de un 42% de los cuestionarios enviados. Esta respuesta se considera buena, tomando en cuenta que la evaluación se envió por medio de un cuestionario electrónico, los cuales presentan tasas de respuestas de alrededor del 33% (Nulty, 2008). Es importante mencionar que la encuesta fue anónima, lo cual dificultó el envío de correos de recordatorio.

Por consiguiente, la cantidad de respuestas obtenidas fueron 28. El cuadro 1 presenta la muestra utilizada dividida por estratos.

Aunque se obtuvieron 28 respuestas, tres participantes marcaron más de un valor para los ítems de la SUS —a pesar de que las instrucciones en el cuestionario claramente solicitaban solo un valor—, lo cual obligó a eliminarlos para el cálculo de esta escala. Por esta razón, la muestra obtenida para la SUS quedó con solo 25 evaluaciones. Este número,



supera la regla de 4 ± 1 individuos para estudios de usabilidad, explicada anteriormente en el marco teórico.



Cuadro 1 Muestra utilizada para evaluación, por estratos (n=28)

Estrato	Invitaciones	Respuestas	% de respuesta
Estudiantes de pregrado/grado	34	10	29%
Estudiantes de posgrado	19	10	53%
Profesores	14	7	50%
Otro*		1	
Total	67	28	42%

* Un participante se identificó en la categoría de Otro

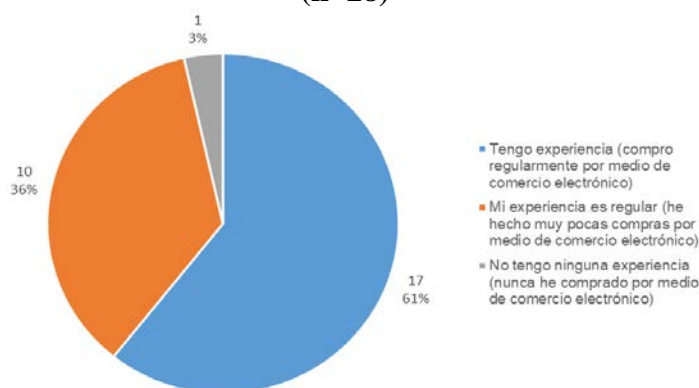
Fuente: elaboración propia

Asimismo, Tullis y Stetson (2004) —utilizando un análisis de “bootstrapping”¹— demostraron que con una muestra de 12 cuestionarios la SUS alcanza un 100% de exactitud. Por lo tanto, la muestra obtenida de 25 respuestas es adecuada para el uso de la SUS.

b. Conocimiento acerca de comercio electrónico

Como se puede observar en el Gráfico 2, el 61% de los participantes indicaron tener experiencia con el comercio electrónico, realizando compras regularmente por este medio. Por otra parte, el 36% indicó tener una experiencia regular, habiendo hecho muy pocas compras. Solo el 3% -correspondiendo a un individuo- indicó no tener experiencia en el uso del comercio electrónico. Por lo tanto, la muestra obtenida cumplió el criterio de que sus miembros tuvieran experiencia con el comercio electrónico.

Gráfico 2
Indique su experiencia en el uso del comercio electrónico
(n=28)



Fuente: elaboración propia

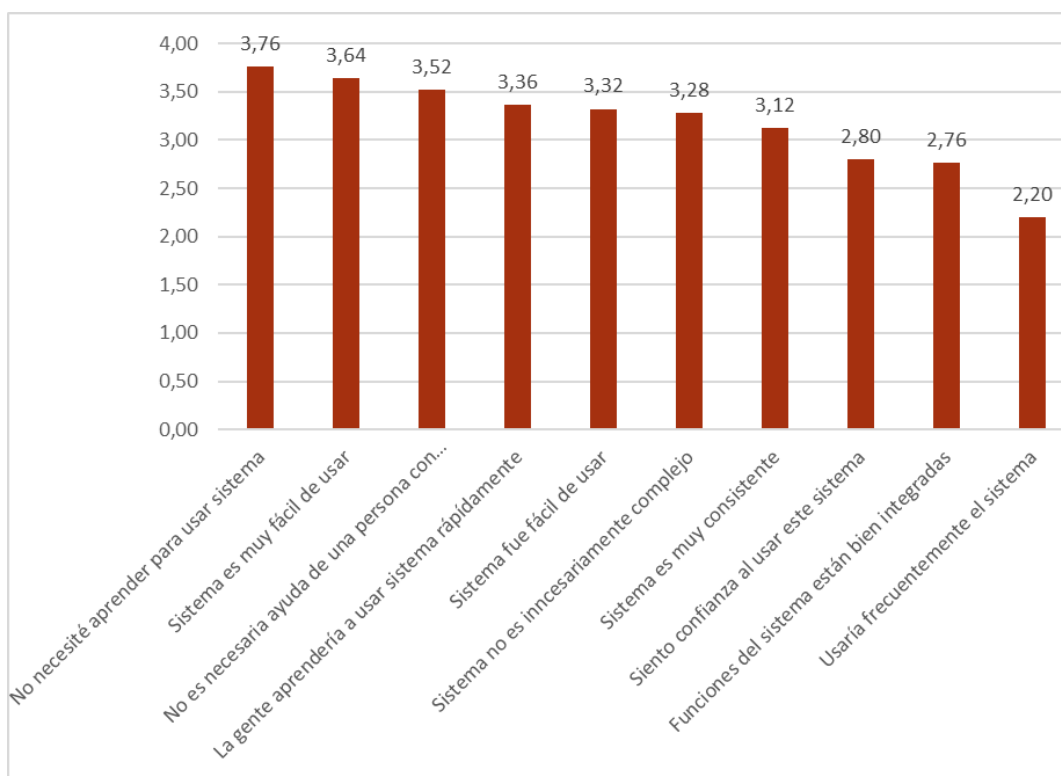
¹ “Boostrapping” es un método estadístico de remuestreo, propuesto por Efron (1979), mediante el cual se utiliza muestras tomando subconjuntos de diferente tamaño de una población para calcular el error al calcular un estadístico debido al muestreo.



c. Evaluaciones para la SUS

El Gráfico 3 presenta los promedios de las evaluaciones obtenidas —de mayor a menor— al aplicar las instrucciones anteriormente mencionadas para evaluar cada una los ítems de la escala utilizada. Como se explicó anteriormente, tres respuestas fueron eliminadas debido a los individuos seleccionaron más de un valor para los ítems de la SUS. Asimismo, es importante resaltar que, aunque la escala originalmente utilizada tenía un rango de 1 a 5, al aplicar los puntos b) y c) presentados en el marco teórico a los ítems de esta escala, se modifica a un rango de 0 a 4, siendo 2 el valor medio o de indiferencia. Como se muestra en dicho gráfico, el promedio para 7 de estos ítems exceden el valor de 3, considerando el nuevo rango. Solo tres ítems están por debajo de este número, aunque por encima del valor de indiferencia (2). Estos ítems son: i) siento confianza al utilizar el sistema, ii) las funciones del sistema están bien integradas, y iii) usaría frecuentemente el sistema. Comentarios provistos por los mismos usuarios respecto al diseño y al funcionamiento del sitio permiten identificar posibles causas para la evaluación de estos ítems, según se discute en las conclusiones y recomendaciones.

Gráfico 3
Promedio para las evaluaciones para la SUS (con los ítems pares revertidos)
(n=25)



Fuente: elaboración propia

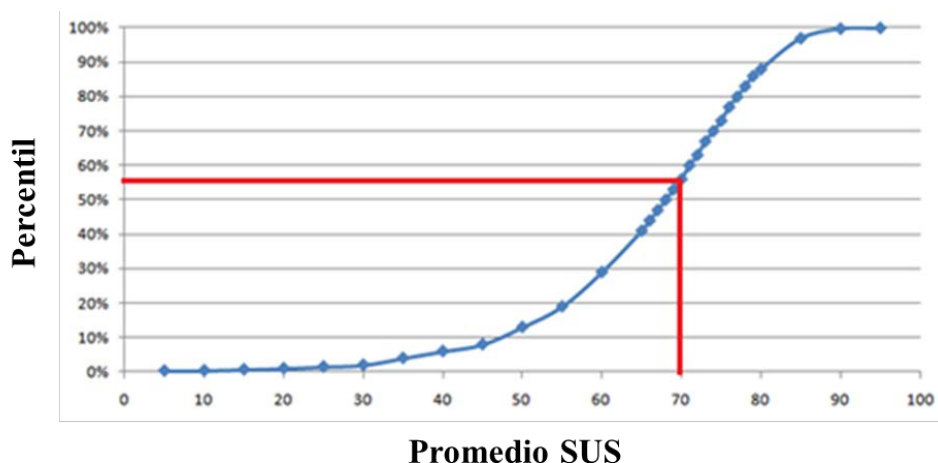
El coeficiente Alfa de Cronbach para los ítems en la escala es de 0,71, el cual se considera aceptable (Nunnaly, 1978).

El rango de valores obtenido para la SUS en este estudio varía de 52,5 a 100, con un



promedio de 70,14. Este valor promedio no se puede interpretar como un porcentaje (Brooke, 2013). Para poder interpretarlo, el mismo debe normalizarse considerando los valores promedios de otros estudios que utilicen la SUS (Bangor et al. (2008) y Sauro (2011)). El Gráfico 4 muestra la interpretación del valor obtenido en este estudio utilizando la curva de normalización propuesta por Sauro (2011), con base a 500 estudios de usabilidad.

Gráfico 4
Normalización del promedio de la SUS



Fuente: Adaptado de Sauro (2011)

De acuerdo con el anterior gráfico, el valor promedio de 70,14 obtenido para la SUS está ubicado en el 56º percentil de los estudios considerados por Sauro (2011), lo cual se puede considerar como bueno (Bangor et al., 2009).

Conclusiones y recomendaciones

Utilizando como referencia estudios anteriores que aplicaron la SUS, se puede argumentar que la evaluación de usabilidad para el prototipo de crigourmetcoffee.com obtuvo buenos resultados con los potenciales usuarios seleccionados.

Como se indicó anteriormente, tres ítems obtuvieron promedios de valores por debajo del valor de 3 (en la escala corregida con rango de 0 a 4): i) siento confianza al utilizar el sistema (2,80), ii) las funciones del sistema están bien integradas (2,76), y iii) usaría frecuentemente el sistema (2,20). En cuanto a la confianza al usar el sistema, algunos de los participantes criticaron la falta del uso en todo el sitio del protocolo de zócalo seguro (<https>). Este protocolo solo estaba implementado para aquellas partes del prototipo en las cuales se intercambiaba información sensible, tal como contraseñas o números de tarjeta de crédito. Con relación a la integración de funciones del sistema, los comentarios provistos permitieron identificar asuntos relacionados con este tema, como por ejemplo la falta de una mención explícita al ícono utilizado para introducir productos al carrito de compras y a botones correspondientes a páginas que aún no están implementadas, y que tampoco, se les había puesto un mensaje de advertencia en este sentido. Asimismo, problemas de despliegue de las páginas en algunos navegadores —los cuales dificultan la



activación de ciertas páginas— pueden atribuirse a la evaluación de este ítem. Por último, aunque no hay comentarios específicos a la baja evaluación sobre el uso del sistema, esta evaluación podría ser explicada por el hecho de que los participantes en esta evaluación no son necesariamente amantes del café (“coffee-lovers”) y, por lo tanto, no serían usuarios frecuentes de un sitio de esta naturaleza.

Es importante reconocer que esta evaluación se realizó con estudiantes y profesores de la UNA, los cuales no necesariamente representan bien a la gama de usuarios que podrían utilizar este sitio o en particular estarían interesados en utilizar este sitio, como antes se comentó. Por esta razón, sería recomendable realizar una evaluación similar con usuarios en otros países. Esta evaluación se tiene prevista con el lanzamiento del sistema. Para la misma se recurriría a los profesores becados de la UNA, con el fin de contactar compañeros, profesores y amigos en los países en que se encuentran, y que puedan estar interesados en el café de Costa Rica.

Referencias

- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human-Computer Interaction*, 24(6), 574-594.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4(3), 114-123.
- Bevan, N., Barnum, C., Cockton, G., Nielsen, J., Spool, J., & Wixon, D. (2003). The "magic number 5": Is it enough for web testing? En ACM (Ed.), *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)*, (pp. 698-699). Ft. Lauderdale, Florida, E.E.U.U.
- Brooke, J. (1996). SUS: A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194), 4-7.
- Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies*, 8(2), 29-40.
- Cronbach, L (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26.
- Lewis, J. R. (1994). Sample sizes for usability studies: Additional considerations. *Human Factors*, 36(2), 368-378.
- Nielsen, J., & Landauer, T. (1993). A mathematical model of the finding of usability problems. In ACM (Ed.), *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, (pp. 206-213). Amsterdam, Países Bajos.
- Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education*, 33(3), 301-314.
- Nunnally, J.C, (1978). *Psychometric Theory*, 2d ed. Nueva York, E.E.U.U.: McGraw-Hill.
- Sauro, J. (2011). *A practical guide to the system usability scale: Background, benchmarks & best practices*. Denver, Colorado, E.E.U.U.: Measuring Usability LLC.
- Tullis, T. S. & Stetson, J. N. (2004). A Comparison of questionnaires for assessing website usability, *Usability Professionals Association (UPA) 2004 Conference*, Minneapolis, EE.UU.



Turban, E., Outland, J., King, D., Lee, J.K., Liang. & Turban, D.C. (2018). *Electronic commerce: A managerial and social networks perspective* (9 ed.). Cham, Suiza: Springer.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

