

Vol. 12 (1), Abril 2014, 177-195

ISSN: 1887-4592

Fecha de recepción: 26-02-2014

Fecha de aceptación: 22-04-2014

Usabilidad y Satisfacción de la e-Rúbrica.

José Serrano Angulo

Daniel Cebrián Robles

Universidad de Málaga, España

Usability and Satisfaction of e-Rubric.

José Serrano Angulo

Daniel Cebrián Robles

Universidad de Málaga, Spain

Resumen

En este trabajo se evalúa la usabilidad de herramientas en línea utilizadas en la educación universitaria. De todas las posibles herramientas se centra en especial en las e-Rúbricas utilizadas para la autoevaluación y la evaluación entre pares. Los resultados muestran un alto grado de usabilidad de esta herramienta y de satisfacción por parte de estudiantes que la emplean como recurso en su formación académica. Desde la perspectiva del diseño y de la accesibilidad se señalan los aspectos a mejorar en las e-Rúbricas evaluadas, sin entrar en aspectos de contenido de las mismas, o de los indicadores empleados, o de otras características propias de los instrumentos de evaluación, sino desde la perspectiva técnica de

Abstract

This work evaluates the usability of tools on line in university education. It focuses especially on the e-Rubric tool used for self-evaluation and peer evaluation among all the possible tools. The results show a high degree of usability and satisfaction on the part of students who use this tool as task in their academic training. The aspects to improve in the evaluated e-Rubric will be pointed out from the perspective of design and accessibility without going into content aspects, used indicators, or other feature of the assessment instruments, but a technical perspective of handling of the tool. The results have been obtained for the implementation of an online survey. It is a Cronbach's alpha of 0.889. The

manejo en la web. Los resultados se han obtenido de la aplicación de un instrumento en línea del que se tiene un coeficiente de fiabilidad alfa de Cronbach de 0,889. El instrumento utilizado consta de unas pocas preguntas descriptivas sobre quien contesta y qué herramienta evalúa, seguidas de 22 enunciados en una escala Likert del 1 al 5. Los enunciados están dispuestos alternativamente en forma directa (1 peor valoración, a 5 mejor valoración), e indirecta (1 mejor valoración, a 5 peor valoración). Los resultados muestran que más del 50% de los que completan el cuestionario valoran la usabilidad de las herramientas evaluadas positivamente, en especial la e-Rúbrica, y menos del 20% valoran la usabilidad como baja. Este cuestionario puede utilizarse en la evaluación de la usabilidad de cualquier herramienta en la Web.

Palabras clave: Usabilidad, e-Rúbrica, evaluación de herramientas on-line, cuestionarios on-line, satisfacción del usuario.

used instrument is the third version of a questionnaire that it has been improving since three academic years ago. The used version consists of a few descriptive questions about the user who answers and what tool is evaluated, followed by 22 questions using a Likert scale from 1 to 5. The statements were prepared alternately in a direct way (1 completely disagree, to 5 fully agree), and in an indirect way (1 fully agree, to 5 completely disagree) in order to prevent, or detect quick responses and/or thoughtless of the respondents. The results show that more than 50% of the users who has responded to the whole questionnaire value the usability of the evaluated tools positively, in special e-Rubric (answers on the values of agreement, and fully agree), and less than 20% value the usability low (answers with values completely disagree and disagreement). All this with federated tools. This questionnaire can be used in the evaluation of the usability of any tool on the Web.

Key words: Usability, e-Rubric, evaluation tools on-line, on-line questionnaire, user satisfaction.

Introducción

La Convergencia y el Espacio Europeo de Educación Superior (EEES) persigue nuevos principios pedagógicos entre los que se destaca un modelo de docencia donde la “enseñanza está centrada en el estudiante” (Blanco, 2009). Esto significa otorgar a los estudiantes mayor responsabilidad en todo el proceso de enseñanza y aprendizaje. Dentro de esta responsabilidad se puede considerar la autoevaluación y la evaluación entre pares. En el primer caso las estrategias para que los estudiantes reflexionen sobre lo que han aprendido y realicen una autoevaluación es un proceso innovador que requiere de los estudiantes una buena actitud y disposición, así como una guía con criterios de evaluación (Moril, Ballester, y Martínez, 2012, Caballero, y Garza, 2012). Un instrumento de evaluación (Rodríguez e Ibarra, 2012) útil para la auto-evaluación y para la evaluación entre pares es la e-Rúbrica (Bartolomé, Martínez, y Tellado, 2012; Raposo, y Gallego, 2012; Martínez, Tellado, y Raposo, 2013; Raposo, Martínez, Tellado, y Doval, 2012). La e-Rúbrica se caracteriza por estar disponible en la red, por lo que está accesible, los da-

tos quedan registrados automáticamente para su análisis y/o información a los demás, (Cebrián, 2011).

Los docentes están implementando estas metodologías en las nuevas titulaciones de grado, ayudándose con tecnologías de la información y comunicación, y de otras herramientas que facilitan el aprendizaje (Cebrián, 2008; Etxabe, Aranguren, y Losada, 2011). Para el desarrollo y viabilidad de estas metodologías se hace casi imprescindible las TIC. Teniendo mayor impacto cuanto más accesibilidad y facilidad de uso muestren. En este trabajo se evalúa la usabilidad y satisfacción de herramientas que son empleadas en la docencia universitaria, prestando mayor atención a la e-Rúbricas, a partir de la opinión de los usuarios sobre las mismas.

La utilización de herramientas en líneas para el proceso de enseñanza aprendizaje será más extendida y se conseguirán mayores éxitos en la enseñanza, cuanto más accesibles y simples de usar sean. Como ejemplo de generalización de uso puede servir la aplicación para dispositivos móviles Whatsapp, que la utilizan incluso personas que nunca han abierto un ordenador, y no sólo esta facilidad de uso, sino que las instrucciones para su manejo se transmiten de unos usuarios a otros, aprendiendo a enviar mensajes, fotos, vídeos,..., entre ellos sin dificultad. El éxito de la herramienta y de servicio, sin duda, se debe a la simplicidad de uso de esta herramienta, o lo que se viene denominado la usabilidad y por supuesto, a las tecnologías de comunicación que la hacen posible. Queda claro que los sistemas de comunicación y las formas de acceso a la información, así como las distintas herramientas de las que se dispone para ello, tienen una mayor aceptación entre los potenciales usuarios cuanto menor exigencia requieren de los mismos para su manejo. Este concepto de usabilidad con el que se viene trabajando, quizás desde que se inventa la primera máquina, se aplica también a la web y a las herramientas disponibles en línea, (Bailey y Pearson, 1983; Bevan, 1991; Brooke, 1996; Kirakowski, 1994; Kirakowski y Cierlik, 1998; Lund, 2001; Sauro, 2011). Las páginas web con una presentación intuitiva ayudan al usuario a encontrar la información que busca en la misma. Y por ende pueden resultar más visitadas que otras. Por lo que consideramos que la evaluación de la usabilidad de cualquier herramienta empleada en la educación, como en este caso nos ocupa aquí de la e-Rúbrica, una estrategia fundamental para mejorar el impacto en la enseñanza de las mismas, en el sentido de accesibilidad, comodidad, facilidad de manejo,..., y si es posible, sin exigencias de conocimientos específicos. En definitiva, que cualquier usuario pueda desenvolverse con ella sin necesidad de formación, o información extensa sobre la herramienta.

Instrumentos para medir usabilidad y satisfacción

Se entiende por Usabilidad según la ISO 9241-10: “los métodos que permiten evaluar si un sitio es usable y el logro de metas específicas como efectividad, eficiencia y satisfacción del usuario en un contexto determinado de uso” (Bevan, 1998, 2009).

Los términos de usabilidad y satisfacción están íntimamente relacionados por lo que se pueden prestar a confusión. Siendo común considerar la satisfacción como una variable de la usabilidad, pues podemos encontrar en ciertas herramientas, instrumentos y escalas de evaluación de la usabilidad la satisfacción como variable. Cuando, todo lo contrario, es más una consecuencia de la usabilidad y no un factor de la misma (Hassan, 2006).

Podemos encontrar un número considerado de cuestionarios para medir la usa-

bilidad y satisfacción de webs, o de sistemas informáticos, entre los que destacaríamos los siguientes:

- El cuestionario SUS -System Usability Scale-, (Bevan, Kirakowski, y Maissel 1991; Brooke, 1996; estándar CISUR del NIST-National Institute of Standards and Technology, 2007). Quizás es uno de los cuestionarios más conocidos por conjugar el número reducido de preguntas y precisión. Mide la usabilidad de una herramienta, programa informático, instrumento, etc. Está compuesto por diez ítems a valorar en una escala Likert del 1 al 5, donde el 1 es totalmente en desacuerdo y 5 totalmente de acuerdo. Al total de puntuaciones se le aplican las transformaciones necesarias para presentarlo en una escala del 1 al 100. Necesita poco tiempo para contestarlo, es simple de calificar y fácilmente es comparable con otros instrumentos. Se recomienda pasarlo una vez que los usuarios han trabajado con la aplicación o herramienta que se va a evaluar, en nuestro caso, la e-Rúbrica, entre otras (Bangor, Kortum y Miller, 2008, 2009; Molich, Ede, Kaasgaard y Karyukin, 2004; Sauro, 2011).
- El cuestionario QUIS -Questionnaire for User Interface Satisfaction- (Chin, Diehl, y Norman, 1988; Harper y Norman, 1993). Este cuestionario permite la evaluación de la satisfacción de los usuarios. Se diferencia del cuestionarios SUS en que se aplica mientras se está utilizando el software, instrumento, servicio, etc. Se desarrolló a finales de los años 80, y actualmente se están desarrollando versiones mejoradas del mismo. Los ítems se presentan en una escala de 0 al 9, siendo 0 confuso y el 9 claro. Consta de 5 categorías de preguntas: 1) Reacciones generales sobre el software, 2) las ventanas, 3) la terminología e información del sistema, 4) el aprendizaje y 5) las capacidades del sistema, (Bailey y Pearson, 1983; Harper, Slaughter y Norman, 1997; Ives, Olson, y. Baroudi, 1983; Kirakowski y Corbett 1988; Slaughter, Harper y Norman, 1994; Wallace, Norman y Plaisant, 1988).
- El cuestionario USE -Usefulness, Satisfaction and Ease of Use- (Lund, 2001). Este cuestionario no sólo mide la usabilidad, sino también la utilidad y la satisfacción de los usuarios. Es uno de los más completos al evaluar la satisfacción, usabilidad y utilidad, además de ser muy simple de implementar al igual que SUS. Consta de 30 ítems en una escala Likert de siete puntos, desde muy fuertemente de acuerdo con la máxima puntuación, a muy fuertemente desacuerdo. Posee también la posibilidad de adaptar las preguntas del cuestionario a necesidades particulares (Alva, Martínez, Cueva, Sagástegui y López, 2003).
- El cuestionario SUMI -Software Usability Measurement Inventory- (Kirakowski, 1994). Es un cuestionario de pago que mide la usabilidad de un software, o servicio, puede ser utilizado para evaluar nuevos productos, realizar comparaciones desde versiones previas para llegar a nuevos y mejores desarrollos, siempre y como todos los demás instrumentos, desde la opinión del usuario final. Consta de 50 ítems a valorar en una escala del 1 al 3 (desacuerdo, no lo sé y de acuerdo). Se ha validado con otros instrumentos como el MUSiC, (Kelly, 1994). En la página web se encuentra información completa del instrumento (Cavallin, Martin, y Heylighen, 2007).

El cuestionario WAMMI -Website Analysis and Measurement Inventory- (Kirakowski y Cierlik, 1998; Kirakowski y Claridge 2013). Éste, al igual que el cuestionario SUMI, es de pago, mide sobre todo la usabilidad de herramientas webs, pero también tiene preguntas de satisfacción y de utilidad. Consta de 20 ítems en una escala

del 1 al 5, desde totalmente en desacuerdo a totalmente de acuerdo (Alva, et al., 2003; Bevan, 1991).

Existen más instrumentos en el estudio del arte como son: MUSIC (Kelly, 1994) que considera la usabilidad como medida a través de la calidad en uso de la herramienta; QEM -Quality Evaluation Methodology- (Olsina, Lafuente, Rossi, 2001) para la evaluación de la calidad en aplicaciones web, considerando cuatro premisas: fiabilidad, eficiencia, usabilidad y funcionalidad; ISOMETRIC, se utiliza para la comparación de productos competitivos basándose en la ISO 9241-10, adaptándose a los principios de la ISO 9241-11; MAGICA -Multimedia AGent-based Interactive CAtalogue IE 2069-, fue el desarrollo de una metodología de la medida de la usabilidad con los siguientes apartados: 1) medida de la satisfacción de los usuarios, para ello hace uso de WAMMI; 2) Medida del tiempo de finalización de la tarea usando la medida RUE (Relative User Effective); 3) Medida del esfuerzo cognitivo, gracias a cuestionarios como SMEQ (Subjective Mental Effort Questionnaire) y CSUQ (Lewis, 1992) entre muchos (Tullis y Stetson 2004).

De todos ellos, el más interesante desde nuestro punto de vista es el cuestionario SUS, por varias razones: es gratuito, sencillo, completo y tiene disponible un baremo en una escala de 0 al 100 (Sauro, 2011) en el que se puede referenciar los resultados de su aplicación para medir la usabilidad de una herramienta en la Web.

Metodología

Los grupos de interés para nuestro estudio están formados por profesores y estudiantes que han utilizado algunas de las herramientas a evaluar. Para ello se contaba con un cuestionario en línea al que se le invitaban a contestar. El acceso al cuestionario es libre y anónimo. Los datos se recogieron en el año 2013, accedieron al cuestionario 198 personas, de la cuales solo 84 completaron totalmente el cuestionario, 77 estudiantes y 7 profesores. Los estudiantes son de primer curso por lo que la mayoría no han utilizado previamente la e-Rúbrica, salvo los repetidores, cuestión que es importante a la hora de valorar la usabilidad, no sólo de la e-Rúbrica, sino también del resto de herramientas que aparecen en el gráfico 1. Los 7 profesores que han completado el cuestionario son profesores que han utilizado alguna de las herramientas en su labor docente, pero no se corresponden con los profesores de los estudiantes. Los datos se analizaran conjuntamente.

Objetivos

El objetivo que se plantea es evaluar la usabilidad y la satisfacción de las herramientas disponibles en línea empleadas en la educación superior, en especial la e-Rúbrica. Para ellos se emplea la tercera versión de un cuestionario del que los autores vienen estudiando sus características métricas y mejorando en años anteriores y que se adjunta anexo a este artículo.

Obtener información de las características métricas del cuestionario para Evaluar la usabilidad y la satisfacción de los usuarios de herramientas en la Web. En particular se aplica a las e-Rúbricas diseñadas por un grupo de profesores en el curso 2012-13, en contextos específicos.

Instrumento

Para medir la usabilidad y la satisfacción de las herramientas en línea que se emplean en la enseñanza, entre ellas las e-Rúbricas, se viene aplicando desde el curso 2009-10 un cuestionario a un promedio de 270 estudiantes cada curso académico. Esto se hacía después de que hubieran realizado auto evaluaciones y/o evaluaciones entre pares con e-Rúbricas (Serrano, y Cebrián, 2012). Con los datos de cada año se han ido estudiando las características internas de cada pregunta del cuestionario para ver su correcta adecuación y para mejorar la fiabilidad del instrumento, cambiando y/o modificando preguntas del mismo según los resultados de las medidas de tendencia central y de dispersión de cada pregunta, así como su índice de homogeneidad interno, y del coeficiente de fiabilidad que se obtenía. A la vez se estudiaba la validez de criterio, con el cuestionario SUS. El cuestionario que se emplea en este estudio es la tercera versión, en la que se ha mejorado el formato de preguntas el número de las mismas, y se han eliminado cuestiones descriptivas que aportaban poco, o ninguna información de relevancia por los estudios anteriores.

En un primer momento se construyeron dos instrumentos uno para medir la usabilidad de las e-Rúbricas y otro para medir la satisfacción por parte de los estudiantes. Se consideraron dos prioridades a la hora de construir los instrumentos: por un lado, evitar respuestas rápidas dadas sin que se llegara a leer los enunciados, y por otro, que fuese breve. Con el fin de evitar respuestas rápidas, o desinteresadas, los instrumentos estaban conformados por enunciados directos e indirectos en una escala Likert de 1 al 5. En los enunciados directos la mejor valoración es el 5, mientras que en los enunciados indirectos la mejor valoración es el 1 y la peor el 5. Asimismo, estos instrumentos deberían de ser relativamente breves para obtener el máximo de respuestas y de mejor calidad, un cuestionario excesivo produce un efecto negativo a quien se le aplica a la hora de responder y de completarlo en su totalidad con la fidelidad necesaria.

Las fases de construcción del instrumento se desarrollaron en tres momentos principalmente: Una vez obtenido los datos y realizado los cálculos se seleccionaron los enunciados a modificar y se construyó la segunda versión de los instrumentos de usabilidad y de satisfacción, aplicándolos a una muestra que habían utilizado e-Rúbrica. Con los resultados de los datos recogidos con la segunda versión (Serrano y Cebrián, 2012) se vuelve a modificar el cuestionario resultando la tercera versión, que es la que se emplea en este trabajo.

En los enunciados de los ítems se ha evitado hacer referencia explícita a las e-Rúbricas con el fin de que se pueda aplicar a otras herramientas, y conseguir una mayor generalidad de los instrumentos de medida de usabilidad y de satisfacción.

En su primera versión, se pasaron todas las cuestiones mezcladas junto con las del cuestionario SUS, dando la apariencia de un solo cuestionario, con esto se pretendía obtener respuestas de cada uno de los instrumentos simultáneamente, para evitar de la aplicación de varios instrumentos alternativos y las alteraciones debidas al momento en el que se aplican. Para su análisis, se separaban los datos de cada instrumentos para estudiarlos por separado. Los valores obtenidos con el cuestionario SUS permitían calcular la validez de criterio del instrumento de usabilidad. La fiabilidad se calculaba con el alfa de Cronbach, por ser el método más adecuado a la situación que se plantea: aplicación una sola vez del instrumento y respuestas en una escala. A la vez se hacía un estudio de las características de cada ítem, calculando sus estadísticos y el índice de homogeneidad interna. A la vista de los resultados se pasó a modificar los enunciados con un índice de

homogeneidad bajo, o negativo. Todo ello se realizó teniendo en cuenta los enunciados inversos, dado que sus valores -como se ha mencionado anteriormente- hay que cambiarlos adecuadamente para que no se contrarresten con los demás.

Hay que hacer notar, que en el transcurrir de los últimos 5 años han ido apareciendo e incorporándose otros medios de acceso a la red, así que cada vez hay más usuarios que acceden a través de una Tablet, de un Smartphone,... Por lo que se exige de las herramientas que sean compatibles en cualquier navegador, sistema operativo, tipo de pantalla,... Al mismo tiempo que se recomienda contar con las versiones adecuadas para su correcta visualización, y por ende, su utilización en cualquier dispositivo al efecto, lo que redundaría en una mayor accesibilidad. Por lo que una de las preguntas de esta versión del cuestionario hace referencia al soporte empleado.

La versión del cuestionario que se utiliza en este estudio contiene preguntas descriptivas como: tipo de usuario, universidad, sexo, edad, herramienta a evaluar, conocimientos que posee sobre la herramienta, experiencia con herramientas similares en Internet, soporte empleado con la herramienta y navegador web usado. Seguimiento de 22 enunciados sobre usabilidad y satisfacción, a valorar en una escala Likert de 1 a 5 totalmente en desacuerdo, a totalmente de acuerdo. Terminando con una pregunta abierta para que el encuestado pueda escribir sus observaciones, sugerencias u opinión sobre la misma.

Entre las características del cuestionario, se ha perseguido en todo momento que fuese breve para que lleve poco tiempo su contestación (Serrano, 2012). Con ello se pretende que no signifique una carga de tarea para los encuestados cuando se les pida que lo cumplimenten. Hay que tener en cuenta que la respuesta a este tipo de cuestionario no le aporta beneficio alguno al entrevistado, por lo que se debe de agradecer su colaboración y por tanto, requerirle poco tiempo para su cumplimentación. Se puede producir respuestas irreflexivas, o sin haber leído los enunciados que se les plantean, para evitar esto, o en su caso, detectarlo, el cuestionario está conformado con enunciados directos e indirectos. De este modo se puede detectar aquellas posibles respuestas de sujetos que responden sin leer los enunciados, sin más que rellenar por rellenar el cuestionario, como por ejemplo puede ser el caso quienes marquen siempre el 4. Las respuestas en las que se detecten este efecto se pueden eliminar, o consideradas como respuestas no válidas en los análisis de los datos.

Resultados y discusión

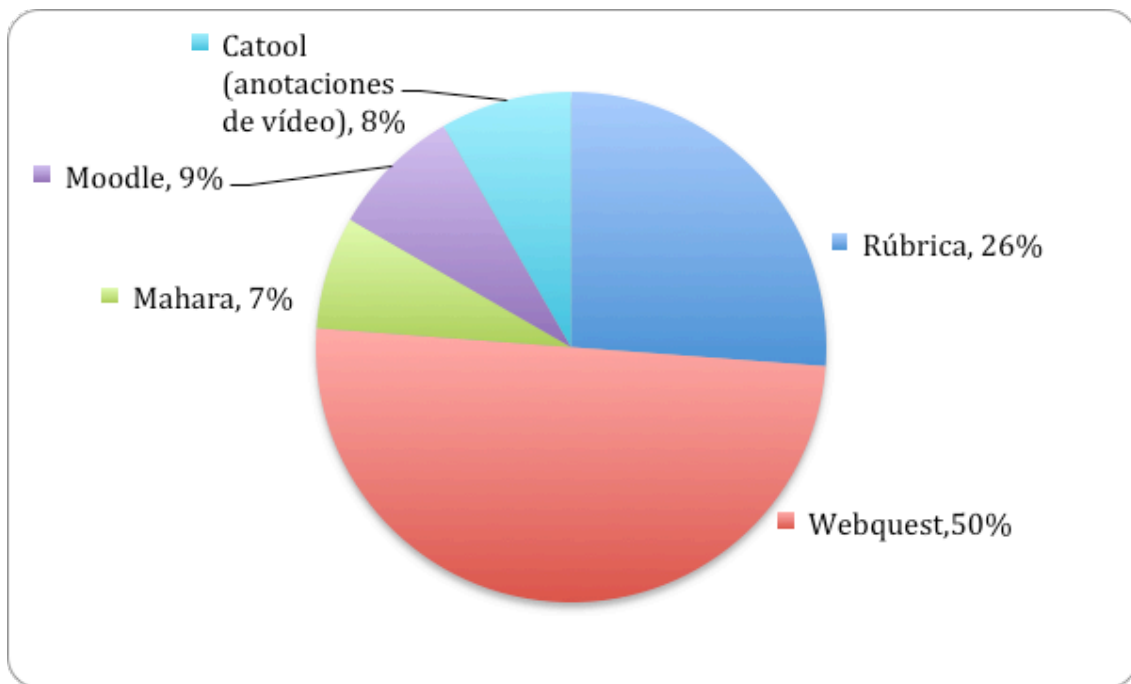
Para el análisis de los datos se revierten los valores correspondientes a los enunciados indirectos, para que no se contrarresten en los distintos análisis estadísticos, y en la puntuación de la suma total. Así se hacen los cambios oportunos para transformar las escalas indirectas en directas.

El cuestionario se administraba en línea con Limesurvey. De las 198 personas que han entrado en el sistema, a final del curso 2012-13, para contestar el cuestionario lo han completado solo 84, por lo que los datos están repartidos como sigue:

El 78,6% de mujeres y el 21,4 de hombres, respecto a la edad el 70% están entre los 18 y los 24 años, y resto de 25 a 54 años. El 91,3% estudiantes y el 8,3% profesores.

Como se muestra en el Gráfico 1 el 26 % han evaluado e-Rúbrica, el 50% Wequest, y el resto se reparte entre Mahara, Moodle y Catool. El reparto de hombres y

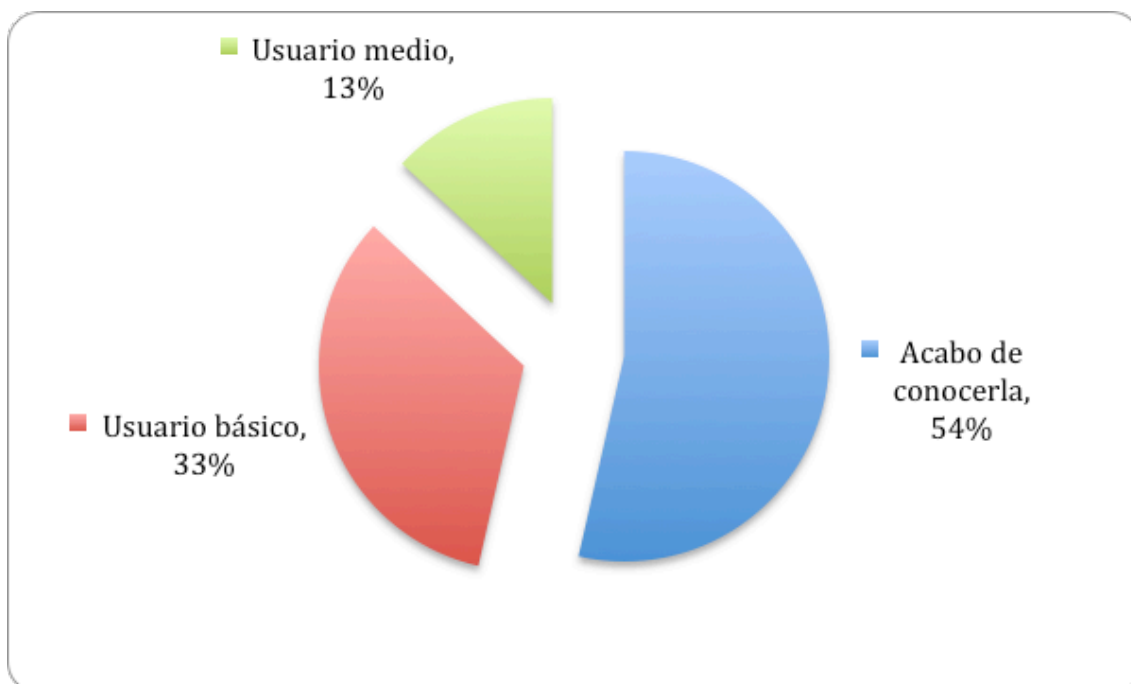
mujeres se mantienen por igual en todas las variables, asimismo ocurre con la edad.



Fuente: elaboración propia

Gráfico 1: Herramienta a evaluar.

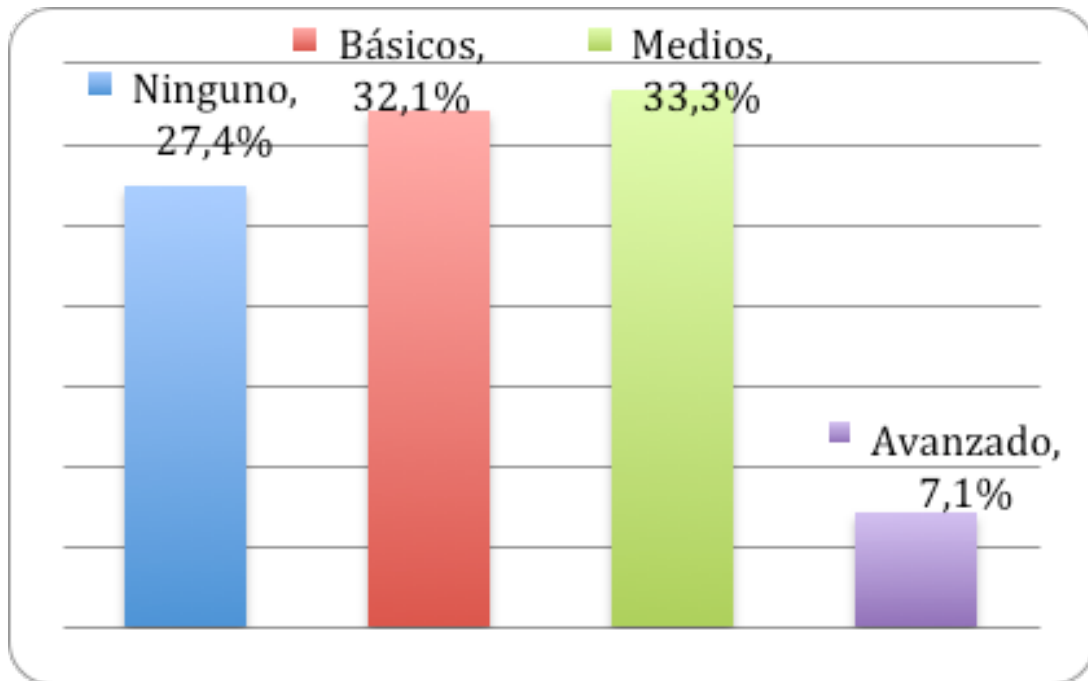
Los conocimientos que tenían de la herramienta a evaluar se muestran en el Gráfico 2, destacando que el 54% dicen no conocerla con anterioridad, lo que es favorable para nuestro objetivo de evaluar la usabilidad y la satisfacción sobre la misma.



Fuente: elaboración propia

Gráfico 1: Conocimientos sobre la herramienta.

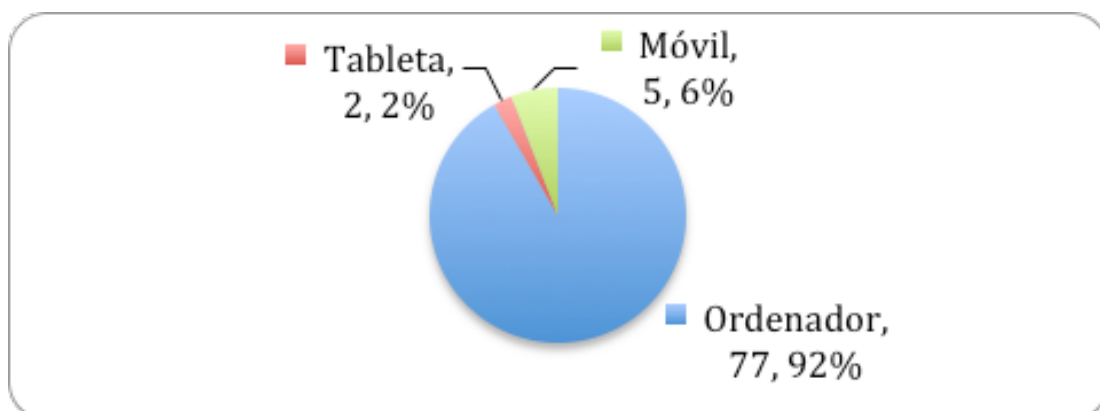
El conocimiento previo de herramientas similares puede inducir a que les resulten más fácil el manejo de estas herramientas, por ello se preguntaba sobre esta cuestión. En el Gráfico 3 se muestra la distribución de la experiencia con herramienta similares, como se aprecia quienes manifiestan no tener conocimiento, o tener conocimientos básicos con herramientas similares suman un 60%.



Fuente: elaboración propia

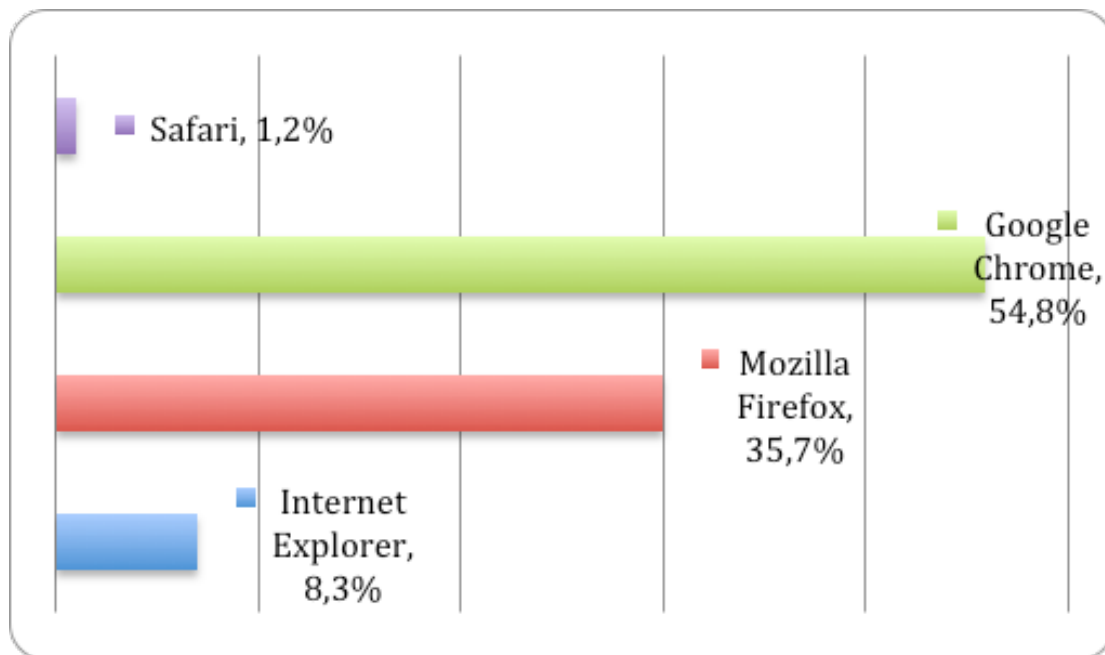
Gráfico 3: Conocimiento de herramientas similares en internet.

El soporte empleado mayoritariamente es el ordenador (Gráfico 4). Aunque hay que señalar que todos los que dicen haber utilizado el Teléfono móvil han sido en el caso de la e-Rúbrica.



Fuente: elaboración propia

Gráfico 4: Soporte empleado con la herramienta.

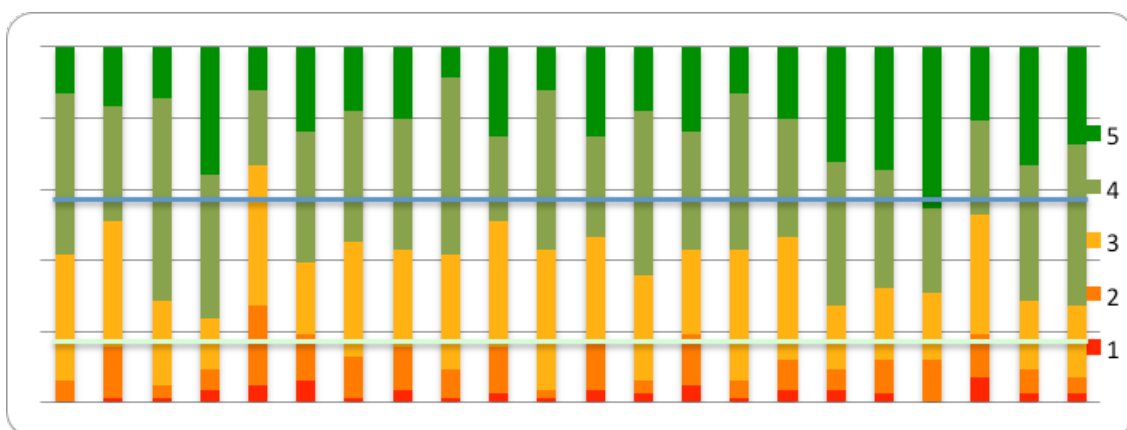


Fuente: elaboración propia

Gráfico 5: Navegador web usado.

El navegador más utilizado es el Google Chrome, seguido de Mozilla Firefox, esta variable nos permite comprobar si las herramientas son igualmente operativas en un navegador u otro, no hay que olvidar que estamos evaluando herramientas que están en línea.

En relación a los enunciados sobre usabilidad y satisfacción se presenta el Gráfico 6, en el que se muestran las respuestas a los distintos enunciados planteados (ver anexo), los números en el eje de abscisas corresponde al orden en el que aparecen los enunciados en el cuestionario. Para una mejor interpretación, aquí hemos invertido los valores de las respuestas en los enunciados indirectos. Se observa que en todas salvo en la cuestión 5ª, que las respuesta peor valoradas (1, ó 2) se encuentran por debajo del 20%. Asimismo, al menos el 50% han valorado totalmente de acuerdo (4, ó 5) a todos los enunciados excepto en el 5º: "Se puede usar sin necesidad de previas explicaciones".



Fuente: elaboración propia

Gráfico 6: Respuestas a cada uno de los enunciados (con los cambios de 5 a 1 en los enunciados indirectos).

El Gráfico 6 nos permite ver el perfil de respuesta en cada uno de los enunciados, si se unen los extremos inferiores del verde tenemos el porcentaje de respuestas de 4, ó 5. Del mismo modo, uniendo los extremos superiores del naranja tenemos los porcentajes de repuestas 2 ó 1. Se puede observar como la respuesta 1 es menor del 10% en todos y cada uno de los enunciados.

De un análisis de conglomerado de los resultados de los 22 enunciados se desprende que hay dos grupos de sujetos. Por un lado, los que dan valoraciones altas y por otro, los que dan valoraciones bajas, siendo los tamaños de los grupos del 69,9% los que valoran más alto y del 31,1% los que valoran más bajo. Por lo que se puede afirmar que quienes manifiestan una mayor usabilidad y satisfacción lo hacen en la mayoría de los enunciados, asimismo para quienes marcan una opción de menor satisfacción y usabilidad de la herramienta.

Los coeficientes de fiabilidad del instrumento calculado en las muestras correspondientes a cada una de las herramientas evaluadas se recogen en la tabla 1. Destacando que la fiabilidad del instrumento es de 0,889, esta es una fiabilidad aceptablemente alta considerando que sólo se tienen 22 enunciados. La media total de 80,66 y la desviación típica de 12,03. Estos resultados pueden servir para una valoración más directa, aunque menos detallada, de la usabilidad de una herramienta. En las medias de cada herramienta se observa que las mejores valoradas es la Moodle y Mahara, aunque también son las más familiares entre los encuestados, la Webquest es la que obtiene una media menor.

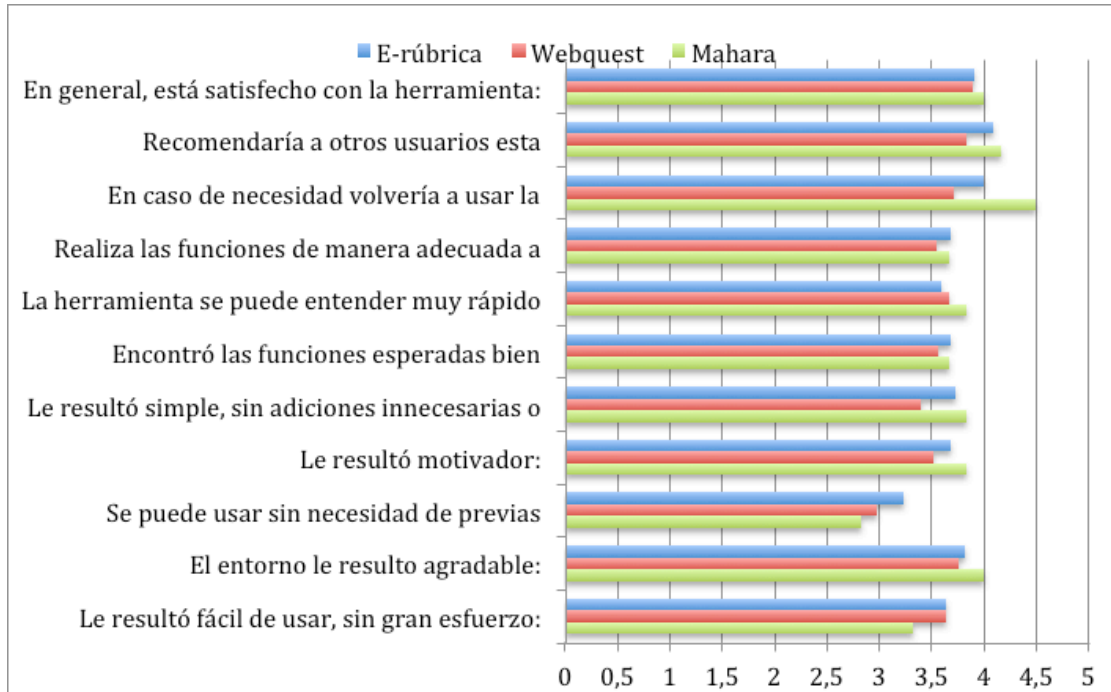
	Media	Desv. típica	Alfa de Cronbach
e-Rúbrica	81,23	12,22	0,891
Webquest	79,44	12,769	0,901
Mahara	82,33	10,633	0,865
Moodle	84,71	12,203	0,915
Catool	80,57	9,589	0,813
Total	80,66	12,031	0,889

Tabla 1: Coeficientes de fiabilidad

Si se analizan las puntuaciones totales de cada sujeto se tiene que el 27,3% de los casos que evalúan la e-Rúbrica le asigna un valor por debajo de la media total, mientras que entre los que evalúan las demás herramientas se reparte al 50% los valores por debajo de la media y por encima de la misma total. Lo que puede considerarse como la herramienta con mayor usabilidad respecto a las otras evaluadas.

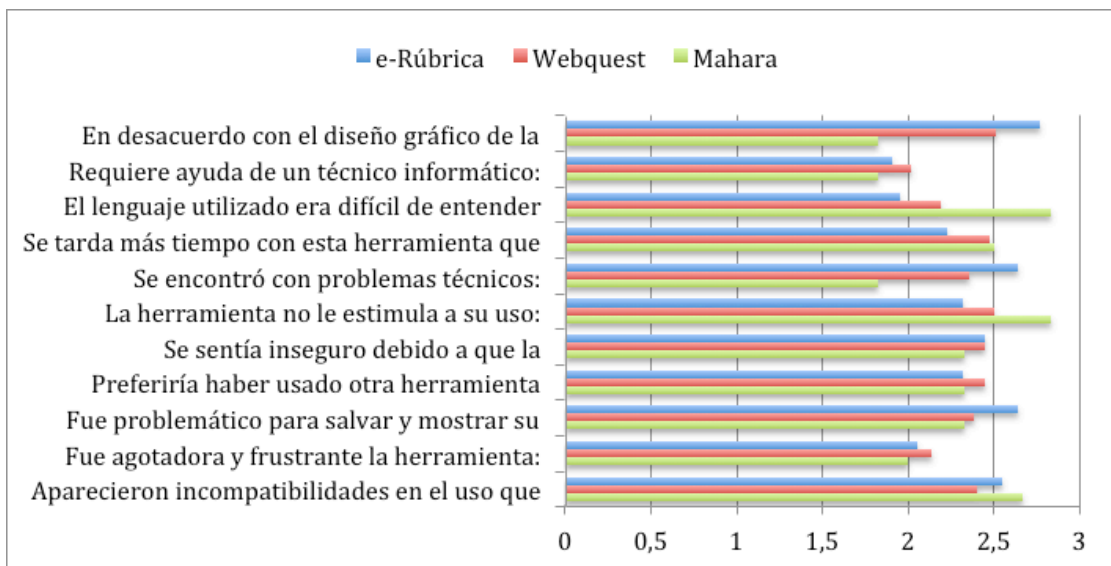
Un análisis más detallado se representa en los siguientes Gráficos en los que se muestran las medias en los enunciados directos (Gráfico 7) e indirectos (Gráfico 8) para las herramientas e-Rúbrica, Webquest y Mahara. En los enunciados directos y en el caso

de la e-Rúbrica todas las medias se encuentran por encima de 3,5 salvo para “Se puede usar sin necesidad de previas explicaciones” que no llega a 3.



Fuente: elaboración propia

Gráfico 7: Media de respuestas en los enunciados directos.



Fuente: elaboración propia

Gráfico 8: Media de respuestas en los enunciados indirectos.

En el Gráfico 8 se muestran las medias de las respuestas originales en los enunciados indirectos. En esta ocasión se observa que la media de la mayoría de los enunciados se encuentran por debajo de 2,5. Hay que hacer notar que en estos casos es más favorable una puntuación baja. Para el caso de la e-Rúbrica se tiene una media por encima de 2,5 en los enunciados “En desacuerdo con el diseño gráfico de la herramienta”, “Se encontró con problemas técnicos”, “Fue problemático para salvar y mostrar su trabajo”, y “Aparecieron incompatibilidades en el uso que le dificultaron”. Por lo que, son aspectos a mejorar en las e-Rúbricas los siguientes: el diseño gráfico, los problemas técnicos relacionados con el acceso, el poder guardar y ver los resultados, e incompatibilidades con el sistema. Éstos pueden ser debidos a que las e-Rúbricas empleadas son herramientas federadas y para acceder a ellas se sigue un proceso de autenticación del usuario, que es el propio de cada universidad y por tanto controlado por cada universidad en su intercambio de metadatos con RedIris. Por lo que estas dificultades pueden deberse más al proceso de federación y de acceso a la misma, que a la propia herramienta.

Aunque se han dado resultados en porcentajes de las respuestas en cada una de los enunciados del cuestionario, describiendo el perfil de respuesta. También se puede establecer los resultados en función de la suma total de las respuestas, aunque no se tenga un información detallada de cada uno de los enunciados. En este caso, el recorrido de los valores que se pueden obtener en el cuestionario va desde 22 a 110, una vez que se transforma los valores de los enunciados indirecto, esto se consigue con al transformación $y=6-x$, donde x es el valor de la respuesta e y el correspondiente si este fuese directo. Para una mejor interpretación de los resultados se podría expresar en una escala del 0 al 100 (Sauro, 2011). En este caso, la media obtenida de 80,66 equivaldría a una puntuación de 66,66 en la escala de 0 al 100 y la transformación de la media a esta escala sería $y=(x-22)*1,35$.

Si se analizan las puntuaciones totales de cada sujeto se tiene que el 27,3% de los casos que evalúan la e-Rúbrica le asigna un valor por debajo de la media total, mientras que entre los que evalúan las demás herramientas se reparte al 50% los valores por debajo de la media y por encima de la misma. Lo que indica que la e-Rúbrica muestra mayor usabilidad y satisfacción entre los usuarios, a pesar de ser una nueva herramienta para la mayoría de los encuestados, y significar un proceso complejo de la evaluación por evidencias, además de representar un momento sensible para los usuarios de la educación.

Conclusiones

Es obvio que una herramienta complicada de usar será menos utilizada libremente que otra que sea más fácil, asimismo la satisfacción por el resultado y uso de la herramienta. En internet podemos encontrar distintas herramientas para un mismo fin, de todas ellas se utilizarán las que sean más accesibles, tengan menos restricciones y requerimientos para su uso, y las que encierren menos complejidad y satisfagan más su resultados. En este sentido, se hace necesario evaluar la usabilidad de las herramientas y la satisfacción de los usuarios en la misma. Ésta es la razón por la que proponemos construir y utilizar un instrumento sencillo y fácil de emplear que proporcione información sobre estos aspectos a los usuarios, y especialmente a los responsables de los servicios de herramientas en la web.

La fiabilidad del cuestionario es de 0,889, la media total de la escala es de 80,66 con una desviación típica de 12,031.

En el caso de la e-Rúbrica las medias en cada una de los enunciados directos se encuentran por encima de 3,5 salvo para “Se puede usar sin necesidad de previas explicaciones” que no llega a 3. La media de las respuestas en la mayoría de los enunciados indirectos se encuentran por debajo de 2,5. Para el caso de la e-Rúbrica se tiene una media por encima de 2,5 en los enunciados “En desacuerdo con el diseño gráfico de la herramienta”, “Se encontró con problemas técnicos”, “Fue problemático para salvar y mostrar su trabajo”, y “Aparecieron incompatibilidades en el uso que le dificultaron”

En vista de los resultados de este estudio se debería de mejorar el diseño gráfico, las condiciones técnicas así como las opciones de guardar y ver, las incompatibilidades surgidas y añadir una breve explicación en línea del funcionamiento de la herramienta. En este estudio no se entra en consideraciones sobre el contenido, indicadores y valoraciones de las e-Rúbricas empleadas.

Referencias bibliográficas

- Alva, M., Martínez, A., Cueva, P. J., Sagástegui, T. Ch. y López, B. (2003) Comparison of Methods and Existing Tools for the Measurement of Usability in the Web. *Web Engineering*, pp. 386–389.
- Bailey, J.E. y Pearson, S.W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management science*, 29(5), pp. 530–545.
- Bangor, A. Kortum, P.T. y Miller, J.T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*. 24(6): pp. 574–594.
- Bangor, A. Kortum, P.T. y Miller, J.T. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*. Vol. 4 (3), pp. 114-123.
- Bartolomé, A., Martínez, M.E. y Tellado, F. (2012). Análisis comparativo de metodologías de evaluación formativa: diarios personales mediante blogs y autoevaluación mediante rúbricas. En C. Leite y M. Zabalza (coords.), *Ensino Superior. Inovação e qualidade na docencia* (417-429). CIIIE: Porto.
- Bevan, N. (1998). European Usability Support Centre. <http://www.usabilitynet.org/papers/tapus98.pdf>
- Bevan, N. (2009). International Standards for Usability Should Be More Widely Used. *Journal of Usability Studies*. Vol. 4, Issue 3, pp. 106-113.
- Bevan, N, Kirakowski, J y Maissel, J, (1991). What is Usability?, in H.-J. Bullinger, (Ed.). *Human Aspects in Computing: Design and use of interactive systems and work with terminals*, Amsterdam: Elsevier.
- Blanco, A. (2009). *Desarrollo y evaluación de competencias en Educación Superior*. Madrid: Narcea.
- Brooke, J. (1996). SUS-A quick and dirty usability scale. En Jordan, P. W., Thomas, B., Weerdmeester, B. A., McClelland (eds.) *Usability evaluation in industry*. London:

- Taylor & Francis, pp. 189–194.
- Caballero Míguez, G. y Garza Gil, M.D. (2012). Innovando la docencia superior en Economía: trabajo cooperativo y elaboración participativa de contenidos. *Revista de Docencia Universitaria*. Vol.10 (2), pp. 319-327.
- Cavallin, H., Martin, W.M. y Heylighen, A. (2007). How relative absolute can be: SUMI and the impact of the nature of the task in measuring perceived software usability. *AI & Society*. 22(2): pp. 227–235.
- Cebrián, M. (2008). La evaluación formativa mediante e-rúbricas, *INDIVISA – Boletín de Estudios e Investigación*-. Monografía X, 197-208.
- Cebrián, M. (2011). *Servicio federado de e-rúbrica para la evaluación de aprendizajes universitarios*. Memoria Técnica para proyectos tipo A y B. Referencia EDU2010-15432.
- Chin, J.P., Diehl, V.A. y Norman, K.L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. En Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 213–218. ACM.
- CISU-R NISTIR. (2007). Common Industry Specification for Usability – Requirements. National Institute of Standards and Technology. <http://zing.ncsl.nist.gov/iusr/documents/CISU-R-IR7432.pdf>.
- Etxabe, J.M., Aranguren, K. y Losada, D. (2011). Diseño de rúbricas en la formación inicial de maestros/as, *Revista de Formación e Innovación Educativa Universitaria*, 4 (3), pp. 156-169.
- Hassan, Y. (2006). Factores de diseño web orientado a la satisfacción y no-frustración de uso. *Revista Española de Documentación Científica*, 29 (2), pp. 239-257.
- Harper, B. y Norman, K. (1993). Improving user satisfaction: The questionnaire for user satisfaction interaction version 5.5. En Proceedings of the First Annual Mid-Atlantic Human Factors Conference. pp. 224–228.
- Harper, B., Slaughter, L. y Norman K. (1997). Questionnaire administration via the WWW: A validation & reliability study for a user satisfaction questionnaire. <http://www.lap.umd.edu/webnet/paper.html>.
- International Standards for Usability: http://www.upassoc.org/upa_publications/jus/2009may/index.html
- Ives, B., Olson, M.H. y Baroudi, J.J. (1983). The measurement of user information satisfaction. *Communications of the ACM*. 26(10), pp. 785–793.
- Kelly, M. (1994). MUSiC Final Report Parts 1 and 2: the MUSiC Project. Hampshire, UK: Brameur Ltd.
- Kirakowski, J. (1994). Background notes on the SUMI questionnaire. Human Factors Research Group University College Cork, Ireland. Originally.
- Kirakowski, J. y Claridge, N. (2013). WAMMI-Web Usability Questionnaire. <http://www.wammi.com/questionnaire.html>
- Kirakowski, J. y Cierlik, B. (1998). Measuring the usability of web sites. En Human Factors and Ergonomics Society Annual Meeting Proceedings, *Human Factors and Ergonomics Society*. Vol. 42, pp. 424–428.

- Kirakowski, J. y Corbett, M. (1988). Measuring user satisfaction. En 4ª Conference of the British Computer Society Human-Computer Interaction Specialist Group, pp. 329–338.
- Lewis, J. R. (1992). Psychometric evaluation of the computer system usability questionnaire: The CSUQ. Technical report, Tech. Rep.
- Lund, A.M. (2001). Measuring usability with the USE questionnaire. *Usability and User Experience*, 8(2), pp. 3-6.
- Martínez M.E., Tellado, F. y Raposo, M. (2013). La rúbrica como instrumento para la autoevaluación: un estudio piloto, *Revista de Docencia Universitaria*, 11 (2), 373-390. <http://red-u.net/redu/files/journals/1/articles/490/public/490-2380-1-PB.pdf>.
- Molich, R., Ede, M.R., Kaasgaard, K. y Karyukin, B. (2004). Comparative usability evaluation. *Behaviour & Information Technology*, 23(1), pp. 65–74.
- Moril Valle, R., Ballester Pont, L. y Martínez Fernández, J. (2012). Introducción de las matrices de valoración analítica en el proceso de evaluación del Practicum de los Grados de Infantil y de Primaria. *Revista de Docencia Universitaria Vol.10 (2)*, Mayo-Agosto 2012, pp. 251-271.
- Olsina, L., Lafuente, G. y Rossi, G. (2001). Using Web-site QEM to Measure the Quality of Sites. http://sedici.unlp.edu.ar/bitstream/handle/10915/22155/Documento_completo.pdf?sequence=1
- QUIS, <http://lap.umd.edu/quis/>.
- Raposo, M. y Gallego, M.J. (2012). Evaluación entre pares y autoevaluación basadas en rúbricas. En C. Leite y M. Zabalza (coords.), *Ensino Superior. Inovação e qualidade na docencia* (444-453). CIIE: Porto.
- Raposo, M., Martínez, M.E., Tellado, F. y Doval, M.I. (2012). La evaluación de la mejora del aprendizaje y del trabajo en grupo mediante rúbricas. En C. Leite y M. Zabalza (coords.), *Ensino Superior. Inovação e qualidade na docencia* (4051-4065). CIIE: Porto.
- Rodríguez Gómez, G. y Ibarra Sáiz, M. S. (2012). Reflexiones en torno a la competencia evaluadora del profesorado en la Educación Superior. *Revista de Docencia Universitaria Vol.10 (2)*, pp. 149-161.
- Sauro, J. (2011). Measuring Usability with the System Usability Scale (SUS) <http://www.measuringusability.com/sus.php>
- Serrano, J. (2012). Aplicación on-line y tratamiento informático de cuestionarios. *Revista Española de Pedagogía. Nº 251*, pp. 61-76.
- Serrano, J. y Cebrián, D. (2012). Usabilidad de la e-rúbrica mediante cuestionarios online con limesurvey. En C. Leite y M. Zabalza (coords.), *Ensino Superior. Inovação e qualidade na docencia* (467-486). CIIE: Porto.
- Slaughter, L., Harper, B., y Norman, K. (1994). Assessing the equivalence of the paper and on-line formats of the QUIS 5.5. En *Proceedings of Mid Atlantic Human Factors Conference*, pp. 87–91. Citeseer.
- SUMI, <http://sumi.ucc.ie/sumipapp.html>.

Tullis, T.S. y Stetson, J.N. (2004). A comparison of questionnaires for assessing website usability. *En Usability Professional Association Conference*.

USE, http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html

Usefulness, Satisfaction, and Ease of Use <http://www.usesurvey.com/>

Wallace, D.F., Norman, K.L. y Plaisant, C. (1988). The American Voice and Robotics “Guardian” System: A Case Study in User Interface Usability Evaluation. Citeseer.

WAMMI, <http://www.wammi.com/>

Artículo concluido el 25 de Abril de 2014.

Serrano Angulo, J. y Cebrián Robles, D. (2014). Usabilidad y Satisfacción de la e-Rúbrica. *REDU: Revista de Docencia Universitaria*, Número monográfico dedicado a *Evaluación formativa mediante Erúbricas*, 12 (1), pp. 177-195.

Publicado en <http://red-u.net>

Anexo: Cuestionario de usabilidad y satisfacción

Hombre Mujer

Edad:.....

Herramienta a evaluar:.....

Conocimientos que posees sobre la herramienta:.....

Valora tu experiencia de herramientas similares en Internet:.....

Soporte empleado con la herramienta:.....

Navegador web usado:.....

Valora de 1 totalmente en desacuerdo a 5 totalmente de acuerdo

	1	2	3	4	5
Le resultó fácil de usar, sin gran esfuerzo:					
Aparecieron incompatibilidades en el uso que le dificultaron el manejo:					
El entorno le resulto agradable:					
Fue agotadora y frustrante la herramienta:					
Se puede usar sin necesidad de previas explicaciones:					
Fue problemático para salvar y mostrar su trabajo:					
Le resultó motivador:					
Preferiría haber usado otra herramienta conocida en lugar de esta:					
Le resultó simple, sin adiciones innecesarias o complicaciones:					
Se sentía inseguro debido a que la herramienta se podía estropear por lo que hacía:					
Encontró las funciones esperadas bien integradas:					
La herramienta no le estimula a su uso:					
La herramienta se puede entender muy rápido por la mayoría de usuarios:					
Se encontró con problemas técnicos:					
Realiza las funciones de manera adecuada a sus expectativas:					
Se tarda más tiempo con esta herramienta que con otra:					
En caso de necesidad volvería a usar la herramienta en el futuro:					
El lenguaje utilizado era difícil de entender para interactuar con la herramienta:					
Requiere ayuda de un técnico informático:					
En desacuerdo con el diseño gráfico de la herramienta:					
Recomendaría a otros usuarios esta herramienta:					
En general, está satisfecho con la herramienta:					

Comentarios adicionales (Optativo):.....

José Serrano Angulo

***Universidad de Málaga
Facultad de Ciencias de la Educación***

Mail: joseserrano@uma.es



Doctor en Matemáticas, profesor de Métodos de Investigación en Educación en la Facultad de Ciencias de la Educación de la Universidad de Málaga. Sus principales líneas de investigación son: evaluación educativa, indicadores de calidad, teoría de la medida, aplicación de métodos estadísticos en la investigación educativa. Coordinador del programa de doctorado de Investigación e Innovación Educativa.

Daniel Cebrián Robles

***Universidad de Málaga
Facultad de Ciencias de la Educación***

Mail: dcebrian@uma.es



Ingeniero Industrial. Universidad de Málaga, Máster oficial “Energías Renovables, Pilas de Combustible e Hidrógeno” Universidad Internacional Menéndez Pelayo y Consejo Superior de Investigaciones Científicas (CSIC). Programa “Doctorado de investigación e Innovación educativa”. y Máster en “Nuevas Tecnologías Aplicadas a la Educación. Universidad de Málaga. Miembro y webmaster servidores del grupo Gtea. Estancias en instituciones extranjeras para el desarrollo de tecnologías de federación como Sined (México) 2011 y Universidade Federal do Paraná (Brasil) 2012. Research Fellow in Harvard University 2013. Desarrollador de la herramienta eRúbrica de gtea [1] y desarrollador de Multimedia Annotations [2] para Harvard University y su integración en el MOOC de edX.