



UNIVERSIDAD DE MÁLAGA
Facultad de Ciencias de la Educación
Departamento de Métodos de Investigación e
Innovación Educativa

Tesis Doctoral

Diseño de instrumentos para medir la usabilidad, la
satisfacción y aspectos técnicos de herramientas en la
Web para la evaluación educativa

Tesis realizada por Daniel Cebrián Robles para el título
de doctor por la Universidad de Málaga

2015

Autor: Daniel Cebrián Robles

Director: Dr. José Serrano Angulo

Programa de doctorado: Investigación e Innovación Educativa



Publicaciones y
Divulgación Científica

AUTOR: Daniel Cebrián Robles

 <http://orcid.org/0000-0002-3768-1511>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está sujeta a una licencia Creative Commons:

Reconocimiento - No comercial - SinObraDerivada (cc-by-nc-nd):

[Http://creativecommons.org/licenses/by-nc-nd/3.0/es](http://creativecommons.org/licenses/by-nc-nd/3.0/es)

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer
obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de
Málaga (RIUMA): riuma.uma.es



MALAGA UNIVERSITY
Faculty of Educational Sciences
Department of Methods of Educational Research
and Innovation

Doctoral Thesis

**Design of instruments to measure usability,
satisfaction and technical aspects on the Web tools for
educational evaluation**

Thesis submitted by Daniel Cebrián Robles for the
degree of Doctor at University of Malaga
2015

Author: Daniel Cebrián Robles

Director: Dr. José Serrano Angulo

Doctoral program: Educational Research and Innovation

Dr. José Serrano Angulo

Departamento de Métodos de Investigación e Innovación Educativa

INFORMA

Que D. Daniel Cebrián Robles, ha realizado bajo su dirección la tesis doctoral por compendio de publicaciones titulada DISEÑO DE INSTRUMENTOS PARA LA MEDIR LA USABILIDAD, LA SATISFACCIÓN Y ASPECTOS TÉCNICOS DE HERRAMIENTAS EN LA WEB PARA LA EVALUACIÓN EDUCATIVA que se recoge en la presente memoria, cumpliendo todos los requisitos legales para optar al grado INTERNACIONAL de DOCTOR, por lo que autoriza su lectura y defensa pública en la Facultad de Ciencias de la Educación de la Universidad de Málaga.

Y para que así conste y tenga los efectos oportunos, firmo el presente informe en

Málaga, lunes, 11 de mayo de 2015

Dr. José Serrano Angulo

Dedicado a mi familia,
mi director y todos aquellos
que me brindaron su consejo

Dedicated to my family,
my director and everyone
who gave me an advise.

Agradecimientos

Agradezco al director de la tesis por su paciencia y apoyo en todo momento, por sus acertadas recomendaciones, por ser quien abrió camino en este trabajo y confió en mi persona para desarrollar el proyecto. También quiero agradecer a todos los docentes e investigadores del grupo GTEA quien me ofrecieron su experiencia y aplicaron los instrumentos que elaboramos en la tesis. Igualmente a todos los investigadores del proyecto I+D+I EDU2010-15432 que utilizaron el servicio de eRúbrica, y que sin el cual no habiéramos podido validar el instrumento creado. Quiero hacer un especial agradecimiento al Dr. Juan José Monedero y Manuel Cebrián de la Serna por el ánimo, recomendaciones y revisiones del instrumento durante la búsqueda de mejora de la fiabilidad. También quisiera citar a la Dra. Manuela Raposo Rivas, Dra. M. Jesús Gallego Arrufat, M^a Esther Martínez Figueiras y Dr. Rafael Pérez Galán por las aplicaciones del instrumento en sus aulas y sus importantes consejos de mejora en el mismo. Por último, deseamos agradecer a la Junta de Andalucía por la beca Talentia concedida y la ayuda de la Universidad de Málaga para estancias en la Universidad de Harvard, el centro de edX (Boston, EEUU) y al Center for Hellenic Studies (Washington, EEUU), cuyo resultado permitió el desarrollo de nuevas herramientas web educativas para cursos masivos abiertos en línea (MOOC) y la aplicación del instrumento de medida.

Indice

Indice.....	V
Introducción.....	1
1. Motivación al estudio.....	3
2. ¿Qué es la usabilidad y satisfacción?.....	5
3. Resumen introductorio al trabajo.....	7
4. Objetivos.....	10
Artículos.....	11
1. Presentación de los artículos.....	13
Artículo 1: Federated eRubric service to facilitate self-regulated learning in the European university model.....	15
Artículo 2: Usabilidad y Satisfacción de la e-Rúbrica.....	27
Artículo 3: Evaluación de pares y autoevaluación con eRúbricas: caso de estudio en el grado de Educación Primaria.....	49
Artículo 4: Usability and Satisfaction in Multimedia Annotation Tools for MOOCs.....	71
Resumen.....	81
1. Resultados.....	83
2. Conclusiones.....	87
3. Nuevas perspectivas o Investigaciones futuras.....	90
Referencias bibliográficas.....	93
Anexo A. El cuestionario.....	99
Anexo B. Justificación de la contribución del doctorando en los artículos.....	107
Anexo C. Índice de impacto de las revistas.....	111
1. Revista European Educational Research Journal.....	113
2. Revista de Docencia Universitaria. REDU.....	114
3. Revista Comunicar.....	115
Anexo D. Summary and Conclusions in English.....	117
1. Summary.....	119
2. Conclusions.....	122

Introducción

1. Motivación al estudio

Como muchas de las profesiones y disciplinas científicas, la educación está en continuo cambio y adaptación a nuevas realidades sociales y tecnológicas, en un mundo cada vez más globalizado, digitalizado y cambiante. En el nivel universitario estas transformaciones han sido importantes en la última década, mostrándose una multitud de cambios en las prácticas con el uso de las tecnologías han sido utilizadas (Kirkup & Kirkwood, 2005), y donde encontramos multitud de estudios sobre su impacto en los aprendizajes de los estudiantes (European Commission, 2008). Estas innovaciones educativas exigen una constante revisión legislativa, normativa, metodológica y curricular; y sobre todo, la experimentación de tecnologías de la información y comunicación aplicadas a la educación y a los cambios y mejoras que esta plantea como necesarias.

La tecnología en general y el número de herramientas web en particular ha crecido exponencialmente en los últimos años. De igual forma, los contenidos en los portales de internet, como el caso de Youtube donde se suben 300 horas de vídeo por minuto (Estadísticas de YouTube, 2015), están configurando una nueva biblioteca digital, permitiendo no sólo una enseñanza universitaria más eficiente y accesible, sino una mayor motivación para los estudiantes y docentes. La cantidad es importante con un aumento imparables; no obstante, incrementa también la necesidad de utilizar y disponer de criterios de calidad para la selección de contenidos y herramientas aplicables a cada contexto educativo específico.

Toda innovación tecnológica no puede ser interesante para todos los contextos educativos, por lo que debemos disponer de criterios claros a la hora de seleccionar contenidos, servicios y herramientas en internet. Necesitamos instrumentos, por tanto, que nos permitan conocer cómo es el acceso a las herramientas, la facilidad de uso de los servicios tecnológicos, así como la satisfacción que producen en los usuarios finales. En nuestro caso, docentes y estudiantes universitarios. Todo ello, nos lleva a la necesidad de conocer con rapidez qué nos ofrecen estos nuevos servicios tecnológicos, pues el dinamismo y actualización es muy elevada. Al tiempo que, saber cómo permiten captar mejor la atención del estudiante, cómo desarrollar las competencias y permitir formas más intuitivas para la construcción del conocimiento y la comunicación didáctica.

Con carácter general, esta tesis pretende mejorar la calidad de los procesos de enseñanza y de aprendizaje de los estudiantes en el uso de servicios de herramientas web. Para ello, analiza y estudia la usabilidad y la satisfacción desde la opinión de los estudiantes, como usuarios finales, cuando utilizan nuevas herramientas educativas online en contextos educativos. A lo largo de este trabajo se presenta el diseño, elaboración y validación de un instrumento de medida fácil de emplear por los docentes e investigadores. Este instrumento será útil a los gestores de los servicios universitarios, los docentes, los investigadores y los estudiantes a la hora de decidir y seleccionar herramientas de entre todas las que se encuentran en los servicios web en internet. Para tal fin, es necesario la creación y validación de un instrumento de medida. El instrumento de medida fue probado en todas las fases de la producción de una herramienta como es la eRúbrica para evaluar

por competencias, como paso previo para posteriormente aplicarlo a otras herramientas educativas. Al acompañar nuestra tesis en todo momento del proceso de creación de una herramienta, como fue desde el diseño hasta la validación de la herramienta, nos permitió conocer con mayor profundidad las cuestiones intrínsecas de cualquier proceso de desarrollo tecnológico, como de creación de medida adaptado al usuario final, especialmente las aportaciones de los docentes a los resultados que encontrábamos de sus estudiantes. En nuestro caso la población y contexto está enmarcado dentro de la población de estudiantes y docentes participantes en el proyecto de I+D+i (Proyecto I+D+i EDU2010-15432, 2010-13, en adelante siempre que se nombre se utilizará “el proyecto I+D+i”). Se creó el servicio de evaluación por competencias mediante eRúbrica para todas las universidades asociadas a RedIris, (2015) (todas las españolas), EduGain, (2015) (la mayoría de las europeas y resto del mundo) y (Gteavirtual, 2015) (grupo de investigación Gtea).

Nuestro proyecto de tesis empezó con la construcción de un primer instrumento de medida centrado en el usuario, como se propone del diseño centrado en el usuario, (Bevan, 2009). Siendo estos usuarios en un principio estudiantes universitarios y docentes. Esperando su extrapolación a una población más amplia de usuarios como son los responsables de servicios tecnológicos, docentes, estudiantes y usuarios finales universitarios. Esto dota de mayor definición y precisión contextualizada en el marco de investigación, y otorga mayor fortaleza a la metodología e instrumento a desarrollar, como en el marco teórico en el que se sustenta desde el punto de vista pedagógico, en línea con el Espacio Europeo de Educación Superior.

Como parámetros de estudio del instrumento se aplicó la usabilidad y satisfacción. La usabilidad y satisfacción es interesante ya que presentan los siguientes beneficios destacados en el panorama educativo:

- Disminuye el coste del aprendizaje.
- Reduce los errores cometidos por el usuario y volver a cometerlo.
- Reduce el coste necesario de un servicio técnico de asistencia y mejorarla sirve de ayuda para el usuario.
- Presenta una optimización de costes de diseño, rediseño y mantenimiento, para no ir a ensayo y error.
- Incrementa los usuarios de la aplicación no sólo en el aula.
- Con ello incrementar la satisfacción y comodidad del usuario.
- Disminuir el estrés del usuario y aumentar así su producción.
- Facilitan la comunicación entre estudiantes y profesores.
- Hace más divertida la clase y la asignatura.

2. ¿Qué es la usabilidad y satisfacción?

Partiendo de la definición ISO 9241-10 se entiende por Usabilidad “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”. Según Bevan, (2009) existen muchos beneficios (Incremento de eficiencia, productividad, reducción de error, formación...) para las organizaciones cuando sus servicios y herramientas están centradas el diseño en la usabilidad de los usuarios -User Centred Design (UCD) (Ferre, Bevan, & Escobar, 2010)- , y se preguntaba por qué a pesar de estos beneficios potenciales aún siguen existiendo servicios que no se preocupan por la usabilidad y facilidad de uso para los usuarios.

Sin duda, como muchas otras áreas y temas de estudios en los seres humanos, el concepto de usabilidad o facilidad de uso requiere considerar otros conceptos, dimensiones y términos asociados, como es el caso de satisfacción, eficacia, eficiencia y utilidad. Por ejemplo, en el trabajo de Hornbæk, (2006) encontramos una buena explicación de usabilidad o facilidad de uso y su relación con los conceptos de eficiencia, eficacia y satisfacción, cuando nos recuerda el estándar (DIS, 1998, p. 2) donde se define la facilidad de uso cuando “un producto puede ser utilizado por usuarios específicos que alcanzan objetivos específicos con efectividad, eficiencia y satisfacción”. Siendo definido el concepto de “eficiencia como los recursos gastados en relación con la exactitud e integridad con que los usuarios alcanzan objetivos”, y por otro lado, el mismo autor define la satisfacción como “la disconformidad o las actitudes positivas del usuario hacia el producto” y también podríamos decir servicios, como es el caso de los servicios informáticos y herramientas web. Dicho de otra forma, la Satisfacción consiste en “the opinion of the user about a specific computer application, which they use” (Doll & Torkzadeh, 1988).

Es fácil caer en la confusión cuando consideramos diferentes conceptos en un mismo estudio. Por ejemplo, entre Satisfacción del Usuario y Usabilidad. Las confusiones más comunes son las que consideran la satisfacción como parte de la usabilidad. Esto es así, porque se puede encontrar en ciertas herramientas, instrumentos y escalas de evaluación de la usabilidad donde la satisfacción es parte del mismo instrumento sin considerar este factor por separado en el estudio.

Otro término que también está muy relacionado es el de utilidad. La diferenciación más clara de utilidad frente a usabilidad es que la usabilidad mide la facilidad de uso, mientras que la utilidad mide la adecuación de una herramienta para un determinado objetivo y perfil de usuario.

En muchas ocasiones no se plantea este factor como una clave posible que explique la baja utilización de un servicio. Normalmente se orienta la evaluación de los servicios a la facilidad de uso, facilidad en el acceso y la satisfacción de los usuarios. Pero en ocasiones no son percibidas en un primer momento la utilidad de los servicios para los usuarios, no importa si otros factores -usabilidad, satisfacción, accesibilidad...- puntúan alto. Esta es una de las críticas actuales a los estudios de usabilidad (Johannessen & Hornbæk, 2014) al no considerar este factor. Y esto es así, porque el uso de servicios y herramientas implica en los usuarios cambios y adopciones de nuevas

formas de pensar y de trabajar con la integración de nuevos servicios. Esto provoca una relación muy estrecha con los cambios e innovaciones organizacionales, con las percepciones de utilidad sobre las innovaciones y nuevos servicios por los usuarios; sin duda, factores relacionados con la satisfacción y facilidad de uso. Como todo proceso humano, los servicios requieren de tiempo hasta ser visto como útiles. Como argumenta McLaughlin & Skinner, (2000) cuando señalan que “la usabilidad y utilidad de los sistemas tecnológicos no son inherentes... surgen a medida que se convierten en parte de la vida cotidiana y los contextos organizacionales”.

3. Resumen introductorio al trabajo

En los orígenes del trabajo, se creó un instrumento para medir la usabilidad y satisfacción de elaboración propia adaptado a la eRúbrica que se validó utilizando como criterio el cuestionario SUS, (Brooke, 1996). Más adelante se trató de modificar el instrumento de manera que fuese capaz de obtener un instrumento de medida más específico para evaluar los dos factores, la usabilidad y satisfacción, en servicios de herramientas web educativas.

De entre los cuestionarios para medir la usabilidad, la utilidad y satisfacción se destacan a continuación los más interesantes de la bibliografía para este trabajo.

El cuestionario SUS (System Usability Scale), (Bevan, N., Kirakowski, J., & Maissel J., 1991; Brooke, 1996). Consta de 10 preguntas y es muy sencillo de completar, calificar y fácilmente es comparable con otros instrumentos. Además es gratuita su utilización. Se aplica una vez que los usuarios han trabajado con la aplicación o herramienta que se va a evaluar, en nuestro caso la eRúbrica.

El cuestionario QUIS, (Chin, Diehl, & Norman, 1988; Harper, Slaughter, & Norman, 1997), (Questionnaire for User Interface Satisfaction). Se desarrolló a finales de los años 80, y actualmente se están desarrollando versiones mejoradas del mismo. En la siguiente página web se puede encontrar información más detallada del mismo y sus actualizaciones: <http://lap.umd.edu/quis/>. Es un instrumento de 27 preguntas que permite la evaluación sobre todo de la satisfacción de los usuarios; si bien, también aborda el tema de la usabilidad. Es de valoración subjetiva, pero se diferencia del cuestionario SUS en que esta se aplica mientras se está trabajando con el software, instrumento, etc.

El cuestionario USE (Lund, 2001), (Usefulness, Satisfaction, and Ease of Use) está aún por revisar. Es muy parecido al SUS, pero este cuestionario además de medir la usabilidad, mide también la utilidad, aprendizaje y la satisfacción de los usuario. Consta de 30 preguntas y es uno de los más completos al medir satisfacción, usabilidad y utilidad, además es muy simple de implementar al igual que SUS, es una escala muy sencilla de aplicar.

El cuestionario SUMI (Kirakowski, 1994), (Software Usability Measurement Inventory, SUMI). Es un cuestionario utilizado para la evaluación de la calidad de un software teniendo en cuenta principalmente la usabilidad (es bastante específico en este sentido) desde el punto de vista del usuario final.

El cuestionario WAMMI, (Kirakowski & Cierlik, 1998), (Website Analysis and Measurement Inventory). Es una extensión del cuestionario SUMI, mide sobre todo la usabilidad de herramientas webs, pero también tiene preguntas de satisfacción de usuario y utilidad. Consta de 20 preguntas.

Existen más instrumentos en el estudio del arte, pero sería excesivo comentarlos todos. No obstante, no podemos por menos que citarlos, como son: MUSiC, (Kelly, M., 1994), que considera

la usabilidad como medida a través de la calidad en uso de la herramienta, QEM, es una metodología cuantitativa para la evaluación de la calidad en aplicaciones web, considerando cuatro premisas, fiabilidad, eficiencia, usabilidad y funcionalidad. ISOMETRIC, es utilizado para la comparación de productos competitivos basándose en la ISO 9241-10, adaptándose a los principios de la ISO 9241-11. MAGICA, 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, con cuestionarios como SMEQ y CSUQ, (Lewis, 1992), entre muchos.

De todos ellos, el instrumento más interesante y que al final se empleó para la validación fue el cuestionario SUS por diferentes ventajas:

- Gratuito
- Sencillo
- Completo
- Muchos autores han validado el cuestionario SUS (Bangor et al., 2008; Molich et al., 2004)
- Presenta un coeficiente de fiabilidad de 0,91 para las ocho preguntas de usabilidad y 0,7 para las dos que presenta de aprendizaje, (Lewis & Sauro, 2009).

En las siguientes fases de la elaboración del instrumento se fueron mejorando las preguntas a partir de los resultados obtenidos en las diferentes versiones del instrumento. Para mejorar la fiabilidad, se fue redactando la comprensión los ítem del cuestionario de una forma más comprensiva, y se fue revisando por diferentes expertos participantes en el proyecto I+D+i que ayudaron con la sintaxis.

Para el estudio de la validez, la fiabilidad y la escala o baremo del primer instrumento utilizado se siguió la validez de criterio, el alpha de Cronbach para la fiabilidad y un baremo de puntuación típica derivada, respectivamente. Este primer instrumento se ha ido mejorando durante el proceso del proyecto y aplicándose a diferentes contextos como puede apreciarse en los artículos que fueron publicados, llegando hasta la 4ª versión.

El identificador de la pregunta no sigue un orden lógico a priori, porque se consideró que sería mejor mezclar preguntas de satisfacción y usabilidad para que si hubiese preguntas parecidas por el cuestionario SUS y el de usabilidad, que no llevase a contestar lo mismo, así se garantiza el estudio. La idea es aplicar ambos test a la vez como si fuese dos test paralelos.

El contexto del estudio y población fueron todos los estudiantes de la Facultad de Ciencias de la Educación de la Universidad de Málaga coordinador del proyecto I+D+i, y el centro asociado a esta universidad en Antequera, abarcando en tiempo desde el curso 2010-11 hasta el 2013-14. La muestra de estudiantes varió según los distintos momentos del estudio, no obstante, se aplicó mayoritariamente en grupos de asignaturas troncales del Grado de Educación Primaria y de Pedagogía con aproximadamente 75 estudiantes por grupo, además de otras asignaturas optativas de

Pedagogía en el tercer y cuarto curso.

Para la primera validación del instrumento de medida sobre la usabilidad y satisfacción se aplicó a los estudiantes de 5 grupos del Grado de Educación Primaria y en la asignatura troncal de Nuevas tecnologías aplicadas a la educación con una totalidad de 375 estudiantes.

En el segundo artículo del compendio de publicaciones se expone la información sobre cómo se aplicó las versiones segunda y tercera del instrumento y su mejora. Recogiendo y analizando datos del año 2013, donde accedieron al cuestionario 198 personas, de las cuales solo 84 completaron totalmente el cuestionario, 77 estudiantes y 7 profesores. El objetivo de este artículo fue evaluar la usabilidad y la satisfacción de las diferentes herramientas disponibles en línea para la enseñanza universitaria y en especial la eRúbrica, entre otras herramientas.

Los estudiantes participantes en este estudio pertenecían al primer curso del Grado de Educación Primaria de la Facultad de Educación de la Universidad de Málaga, en la asignatura de Nuevas tecnologías aplicadas a la educación, por lo que la mayoría no habían utilizado previamente la eRúbrica, salvo los repetidores.

Siguiendo el proceso de mejora del instrumento se aplicó para evaluar la usabilidad y satisfacción de la eRúbrica en los contextos y prácticas planteadas por el proyecto de I+D+i se usó la autoevaluación y la evaluación de pares. Este estudio se publicó en el tercer artículo que se presenta como parte de la tesis. La muestra se tomó de la población de la Escuela de Educación de Antequera, centro adscrito a la Universidad de Málaga, y durante todo el curso académico 2011-2012, con asignaturas troncales diferentes (Teoría y Política de la Educación Infantil y hacia una escuela inclusiva: modelos y prácticas) y con idénticos grupos (Grupo de 1º de Grado de Educación Infantil A-grupo experimental- y Grupo de 1º de Grado de Educación Infantil B-grupo control).

Para el último estudio de mejora del instrumento de usabilidad y satisfacción, se aplicó a otras herramientas diferentes a los estudios previos, como fueron Open video annotation -OVA- y Colaborative annotation tool -Catool- (ver cuarto artículo). En el que se utilizaron las versiones tercera y cuarta del instrumento. En este caso, la muestra fue de 180 estudiantes de las asignaturas troncales de Recursos didácticos y tecnológicos en educación de segundo año del Grado de Pedagogía, y dos optativas de cuarto curso del mismo grado.

4. Objetivos

La tesis se ha desarrollado con una estrecha vinculación con el proyecto I+D+i titulado: Servicio federado de eRúbrica para la evaluación de aprendizajes universitarios. Dentro del Plan Nacional I+D+i 2010-2013. EDU2010-15432. Resolución de 30 de diciembre de 2009 (BOE de 31 de diciembre).

Este proyecto planteaba entre otros objetivos la elaboración de un servicio de eRúbrica con acceso de autenticación federado por RedIris. El desarrollo tecnológico que perseguía el proyecto I+D+i nos permitió un espacio ideal para definir y alcanzar nuestros propios objetivos de investigación para la tesis, que si bien no eran objetivos del proyecto sí los complementaban y ampliaban. De este modo, nuestros objetivos facilitaban el éxito del desarrollo y logro de los objetivos del proyecto de investigación I+D+i, comprendiendo la facilidad de uso y la satisfacción que representaba dicho servicio para sus usuarios. Con lo que, nuestro objetivo nació y se desarrolló a la par que el proyecto, sirviendo de proceso de evaluación y mejora de los servicios que pretendía el proyecto I+D+i.

El objetivo general pretende “La elaboración y validación de un nuevo instrumento de medida para evaluar la usabilidad y satisfacción de herramientas y servicios web educativos”.

Para el logro de este objetivo general perseguimos otros objetivos más específicos, como fueron los siguientes:

1. Diseño y creación de un servicio web federado en GTEA que facilite la evaluación formativa y la autorregulación de los aprendizajes universitarios; a la vez que, nos permita el estudio y evaluación de la usabilidad y la satisfacción de los servicios y herramientas federadas.
2. Elaboración y validación de un instrumento de medida para evaluar la usabilidad y satisfacción del servicio eRúbrica.
3. Adaptación y mejora del instrumento de medida del objetivo 2 para evaluar la usabilidad y satisfacción de otros servicios y herramientas web federadas en GTEA.
4. Experimentación y evaluación del servicio de eRúbrica federado en un contexto de enseñanza universitaria, como prueba de concepto para la autorregulación de los aprendizajes mediante la autoevaluación y evaluación por pares.
5. Estudio de las posibilidades del instrumento objetivo 3 como modelo de contraste y mejora de servicios y herramientas federadas con funciones similares.

Artículos

1. Presentación de los artículos

La presente tesis es un compendio de cuatro artículos publicados que guardan una vinculación estrecha con el proyecto de investigación referido anteriormente. A continuación vamos a explicar los contenidos de dichos artículos y el orden lógico seguido en la producción de sus contenidos. Si bien, el orden de publicación de los mismos pudiera ser diferente debido al proceso de revisiones seguido, y los intereses en las fechas de publicación de las revistas.

El primer artículo parte desde dos marco teóricos bien amplios: uno de naturaleza pedagógica y otro de índole tecnológico. Del tal modo, partimos por un lado y en primer lugar, desde el marco pedagógico del espacio europeo de educación superior, destacando el principio de enfocar la enseñanza en el aprendizaje de los estudiantes, en el que sean ellos los protagonistas del proceso de enseñanza y estén cada vez más comprometidos con su propio aprendizaje. Enseñanzas centrada en los aprendizaje como plantear el diseño de herramientas centrado en el usuario (Bevan, 2009). Es aquí donde la herramienta de eRúbrica pretende servir de instrumento para la autorregulación de sus aprendizajes por los estudiantes; a la vez que, una metodología para apoyar y facilitar el trabajo del docente cuando pretende implementar una evaluación formativa basada en competencias.

En definitiva, en este primer artículo llegamos a la conclusión de que las eRúbricas son unas herramientas que pueden facilitar estos procesos de autorregulación en el aprendizaje de los estudiantes. Por tanto, sería una estrategia muy oportuna conocer la opinión de los usuarios -que son el centro de este modelo-, sobre los recursos tecnológicos o herramientas web. Es decir, conocer lo fácil que resultan estas tecnologías y servicios para los usuarios, saber si se encuentran satisfechos en algo tan particular y privado como es la autorregulación de su aprendizaje y la autoevaluación con herramientas como la eRúbrica.

Por otro lado y en segundo lugar, se presenta y explica el modelo técnico de la federación, que es la base tecnológica que subyace a la herramienta creada de eRúbrica en GTEA (<http://gteavirtual.org/rubric>). Esta representa toda una innovación tecnológica de reciente implantación, y es el principio que rigen los mecanismos de acceso a los servicios y herramientas web hoy presentes en la mayoría de las instituciones de educación superior en el mundo, como también en los servicios web, especialmente Google quien brinda un sistema similar a la federación a muchas herramientas en internet. A la par, se muestran en este artículo las principales funciones técnicas de la herramienta eRúbrica para la evaluación por competencias y su relación con la autorregulación de los aprendizajes.

En este segundo apartado como en el primero, abordamos también una relación de las tecnologías de federación con los espacios personales de los estudiantes Personal Learning Enviroments (PLE) (Coll & Engel, 2014), especialmente por la comodidad de acceso a los servicios y herramientas que brindan las instituciones a sus estudiantes, creando un acceso único de autenticación, donde con un “solo clic” pueden disponer automáticamente el acceso a todas las demás herramientas de la institución. Mediante herramientas federadas es mayor la facilidad de

acceso a los servicios (más fáciles de usar o usabilidad) y mayor satisfacción a la hora de construir sus propios entornos personales y con ello generar un mayor aprendizaje. Esto nos reafirma en la necesidad de que las universidades ofrezcan sus servicios federados, y que todas sus herramientas deberían ser evaluada su usabilidad y la satisfacción por los usuarios finales.

El segundo artículo se adentra en el análisis y comprensión de la usabilidad y la satisfacción; así como, expone y analiza los instrumentos más conocidos para su medición. Partiendo del instrumento creado en la tesis en su primera y segunda versión, mejora la fiabilidad del instrumento en una tercera revisión y aplica dicho instrumento en diferentes servicios y herramientas web educativa, especialmente en el servicio de eRúbrica frente a otras herramientas.

En el tercer artículo se analiza el impacto de la eRúbrica en los aprendizaje universitarios como prueba de concepto para la autorregulación de los aprendizaje mediante la autoevaluación y evaluación por pares, donde un grupo de estudiantes recibía una metodología de evaluación de pares y auto-evaluación con eRúbricas.

En el cuarto artículo, y después de una cuarta y última revisión del instrumento de medición creado, se aplicó y se midió la usabilidad y satisfacción en dos herramientas y servicios web con similares funcionalidades como fueron Collaborative Annotation Tool (CaTool) y Open Video Annotation -OVA-. La primera herramienta elaborado por el equipo de anotaciones de Harvard University, y la segunda desarrollado por nosotros dentro de un proyecto de colaboración con este mismo grupo en Harvard University, y que contó con diferente apoyo como fueron Becas Talenta de la Junta de Andalucía y ayudas a estancias para investigación de doctorando por la Universidad de Málaga.

A continuación se recogen las diferentes publicaciones:

- Cebrián Robles, D.; Serrano Angulo, J. and Cebrián de la Serna, M. (2014). Federated eRubric service to facilitate self-regulated learning in the European university model. *European Educational Research Journal*. 13 (5). 575-584 ISSN:1474-9041. <http://dx.doi.org/10.2304/eeerj.2014.13.5.575>
- Serrano Angulo, J. and Cebrián Robles, D. (2014). Usabilidad y satisfacción de la eRúbrica. *Revista de Docencia Universitaria. REDU*. 12(1), pp.177-195. ISSN:1887-4592 <http://red-u.net/redu/index.php/REDU/article/view/775/pdf>
- Pérez Galan, R., Cebrian Robles, D. & Rueda Galiano, A. (2015). Evaluación de pares y autoevaluación con eRúbricas: caso de estudio en el grado de Educación Primaria. *Revista de Docencia Universitaria. REDU*. 12(4). 437-456. <http://red-u.net/redu/index.php/REDU/article/view/719>
- Monedero Moya, J., Cebrián Robles, D. & Desenne, Ph. (2015). Usabilidad y satisfacción en herramientas de anotaciones multimedia para MOOC. *Comunicar*. 44. ISSN: 1134-3478, <http://dx.doi.org/10.3916/C44-2015-06>

Artículo 1: Federated eRubric service to facilitate self-regulated learning in the European university model

1 Federated eRubric Service to Facilitate Self-regulated 2 Learning in the European University Model

3 **DANIEL CEBRIÁN ROBLES**

4 *Faculty of Education Sciences, University of Málaga, Spain*

5 **JOSÉ SERRANO ANGULO**

6 *Department of Research Methods and Innovation in Education,*

7 *University of Málaga, Spain*

8 **MANUEL CEBRIÁN DE LA SERNA**

9 *Department of Didactics and School Organisation,*

10 *University of Málaga, Spain*

11 **ABSTRACT** The European Higher Education Area (EHEA) is a political programme for higher
12 education in Europe that was developed in the context of the Bologna process. It highlights the
13 importance of focusing education on students' learning. It also claims that students should achieve
14 certain skills in a self-study process supported by their teachers. This approach demands that students
15 reflect upon and self-assess their learning. It therefore requires that research is conducted on new
16 technologies that require a greater involvement of students in their own education. The technologies
17 require a higher commitment from students regarding their own assessment. eRubric evaluation is part
18 of this conception of technology-enhanced learning and formative assessment. It is an evaluation
19 method, a technique and an assessment management tool to support self-regulation of the learning
20 process. Information is given about a research project to construct the use of eRubrics.

21 **Introduction**

22 The European Higher Education Area (EHEA) is a political programme for higher education in
23 Europe that was developed in the context of the Bologna process. It was officially launched in 2010
24 and was meant 'to ensure more comparable, compatible and coherent systems of higher education
25 in Europe' (<http://www.ehea.info/>). In the Bucharest Communiqué, ministers of 47 European
26 countries agreed on the following priorities for future work in the EHEA: (1) to provide quality
27 higher education for all; (2) to enhance graduates' employability; (3) to strengthen mobility as a
28 means for better learning (EHEA Ministerial Conference, 2012). This document states: 'We
29 reiterate our commitment to promote student-centred learning in higher education, characterised
30 by innovative methods of teaching that involve students as active participants in their own
31 learning' (EHEA Ministerial Conference, 2012, p. 2).

32 The active participation of students in their own learning can be greatly increased by fostering
33 their competences for self-regulated learning (SRL). The development and improvement of SRL
34 competences depends greatly on the feedback students receive on their learning activities. It is our
35 belief that electronic rubrics (eRubrics) have a great potential for fostering SRL and subject-specific
36 competences by providing feedback through self-assessment and assessment by teachers and peers.

37 The objective of this article is to present and discuss the concept of eRubrics in a federated
38 system of higher education institutions based on the 'Federated eRubric Service to Assess
39 University Learning', a research project led by the University of Malaga along with five other

18 institutions on the use of eRubrics to assess learning.[1] The project evaluates the extent to which
 40 eRubrics can be used in different university contexts (compulsory subjects, optional subjects, lab,
 41 blended learning, e-learning, internships, etc.) and in different fields of study. The current project
 42 uses as a technological tool the eRubric developed by Gtea [2] (University of Malaga), which can be
 43 accessed by all universities in Spain through the Spanish academic and research network RedIris, as
 44 well as by universities of Mexico and Latin America through the National System of Distance
 45 Education, Sined [3], and European universities and the rest of the world through EduGAIN.[4]
 46 The EduGAIN service is intended to support the trustworthy exchange of information related to
 47 identity, authentication and authorisation between the GÉANT partners. GÉANT is a network that
 48 links European national research and educational networks. The initial goal was to enable a Pan-
 49 European Web Single Sign On.[5] Meanwhile, there are EduGAIN members in almost all
 50 European countries.

51
 52 Assessment has been a recurring theme in university research (Brown & Glaser, 2003).
 53 However, recent changes have focused on competence assessment, which demands greater
 54 participation of students throughout the entire learning process. The work programme proposed
 55 for the European Higher Education Area [6] has turned assessment into one of the main areas
 56 related to innovation and current studies, research, projects and developments. These changes
 57 require a greater responsibility from students so that they can assume a more defined role in both
 58 learning (self-learning) and assessment (self-assessment). They also require teachers to change from
 59 using traditional evaluation methods to using new methods for competence evaluation, and these
 60 changes do pose difficulties. At the same time, both students and teachers benefit from this new
 61 approach (Mooij, 2007).

62 In addition, students are increasingly expected to self-regulate their learning (Blanco, 2009;
 63 López, 2009; Carneiro et al, 2011; Rodríguez & Ibarra, 2011); in fact, the capacity to self-regulate
 64 one's learning is seen as one of the twenty-first-century key competences (European Parliament,
 65 2006). Basically, to self-regulate one's learning means to be aware of one's own learning processes.
 66 While there are a number of models for self-regulated learning (SRL), the best known is probably
 67 the one developed by Zimmerman (2000). According to Zimmerman, learning may be self-
 68 regulated in cycles of (1) forethought, (2) performance or volitional control, and (3) self-reflection.
 69 Evidently, feedback students receive on their learning processes will help them to self-regulate their
 70 learning. One way to receive feedback is through assessment. The eRubric tool can provide self-
 71 assessment and peer assessment as well as assessment by the teacher (Panadero & Jonsson, 2013).

72 It is therefore a technology that improves the learning environment (Steffens & Underwood,
 73 2008). This is realized by helping teachers design criteria for competence assessment and by
 74 improving communication between teachers and students about the assessment. It is hoped that
 75 interesting results from studies that focus their attention on the impact of learning with the eRubric
 76 tool will soon be obtained, as this model allows for a better understanding and greater cooperation
 77 between students and teachers in the process of developing assessment criteria. Evaluation by
 78 eRubrics represents an innovative model that allows both teachers and students to achieve
 79 professional skills and improve them; in the case of students, because they are able to appreciate
 80 their learning progress more easily, they can identify the indicators and evidence of skills developed
 81 in a course. Furthermore, they can discuss with their teachers the quality of the learning process
 82 based on the indicators and criteria specified in the eRubric. In the case of teachers, they provide
 83 definitions of the relevant skills and their associated criteria and indicators, which they then apply
 84 to the evidence of learning submitted by students (Hanrahan & Hafner, 2003; Prins et al, 2005;
 85 Vickerman, 2009).

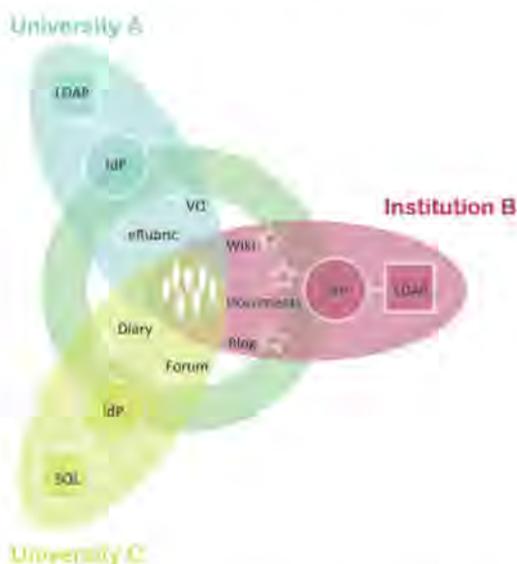
86 Developing a research project on formative assessment by using eRubrics involves a strong
 87 technological component. On the one hand, the assessment tool itself, the eRubric, has a privileged
 88 position in distance or e-learning contexts and also in blended courses because it facilitates the
 89 personalised monitoring of the learning process. On the other hand, the participation of teachers
 90 and researchers from different universities and the active role that students from each institution
 91 must take creates a need for organising access to the tools which are involved in the projects.

92 The research project presented in this article uses the eRubric federated tool service,
 93 developed by Gtea (University of Malaga). This service offers the most important advantages in the
 94 practice of the federated technological model, as will be shown in the next section. In addition, it
 95 encourages new research lines about the implementation of the eRubric in different contexts

96 (compulsory courses, elective subjects, laboratory practices, blended learning, e-learning,
97 practicum, etc.) and in different fields of study, such as SRL.

98 **Theoretical Model**

99 Technologies that identify users and provide access to a variety of tools can be used to develop new
100 models of learning environments that are personalised. This is the case with the eRubric, which
101 teachers can select from a list of possible tools in the shared network system. The tool is distributed
102 via the Internet and may belong either to the institution where it was created or to another
103 institution within the federated system. This system offers teachers the opportunity to create a new
104 learning environment which is much more personalised and flexible than traditional platforms.
105 Moreover, teachers may adapt the tool to particular teaching contexts.
106



107 Figure 1. Federated environments for collaboration.

108 Figure 1 presents a theoretical model tested in practice, which demonstrates how a group of
109 students and teachers can choose different services and tools from different institutions, using their
110 own institution's login. Some acronyms in Figure 1 are technical and refer to computer terms
111 regarding the security of the information flow between different systems from different universities
112 or other organisations. For example, Lightweight Directory Access Protocol (LDAP) accesses and
113 maintains distributed directory information services over an Internet Protocol (IP) network;
114 Structured Query Language (SQL) is a programming language for managing data held in a
115 relational database management system; Identity Provider (IdP) asserts identification information
116 for all providers via an authentication module; and Virtual Organizations (VO) are tools enabling
117 the organisation of users into groups in order to create learning communities. In addition, as we
118 can see in Figure 1, each participating university or other organisation offers its own tools to the
119 federated community. University A offers an eRubric and VO; Institution B offers a wiki,
120 documents and a blog, and University C provides a diary and a forum. All of the tools within the
121 green circle are linked to a common platform (LMS), **SPELL OUT?** This means that once users
122 enter their own institutional password in any tool within the green circle, the user has access to all
123 the tools included without having to enter the password again. Thus, the federated environment
124 facilitates joint collaboration projects with eRubrics between teachers and students from different
125 universities or other organisations across different countries.
126
127

128 The federated system constitutes a technological innovation because it represents a new way
129 to access online resources. It is as simple as logging into Google. Once an organisation has been
130 registered with the system and users have been given an account, they can share a tool and have

120 access to all other tools within the system with a simple login. This technology represents a new
 132 model through which institutions can collaborate; it is a new personal learning environment for
 133 students and teachers where the Internet is the platform (Accino Dominguez & Cebrián de la
 134 Serna, 2008).

135 **The Tool's Features**

136 The technological support for the project is based on SimpleSAMLphp, which is software including
 137 Security Assertion Markup Language (SAML) in PHP developed by Uninett, the Norwegian
 138 academic network. The SimpleSAMLphp infrastructure is the most simple and efficient solution as
 139 it allows a group of institutions to federate by integrating their shared services into an identity
 140 infrastructure that uses SAML 2. The Gtea group has made available a website with federated tools
 141 based on this technology [7], thus allowing for the creation of new, open and flexible
 142 environments.

143 A core tool of the federated environment is the Gtea eRubric [8], a competence-based
 144 assessment tool allowing for different modalities of assessment – self-assessment, peer assessment,
 145 teacher assessment and group assessment – all of which can be anonymous. The basic idea of the
 146 eRubric model is that competences need to be operationalised. That is, teachers need to define
 147 observable events (learning outcomes) which can be considered indicators of the competence in
 148 question. In Figure 2, for example, 'Define a research project with objectives' would be an indicator
 149 for the partial competence 'Pre-project summary research'. Several partial competences may be
 150 defined and differentially weighted to constitute the competence in question. Indicators of a partial
 151 competence may be assessed using quantitative or qualitative criteria. In the example of Figure 2,
 152 quantitative criteria are used ranging from 'not shown' to 'excellent'.
 153

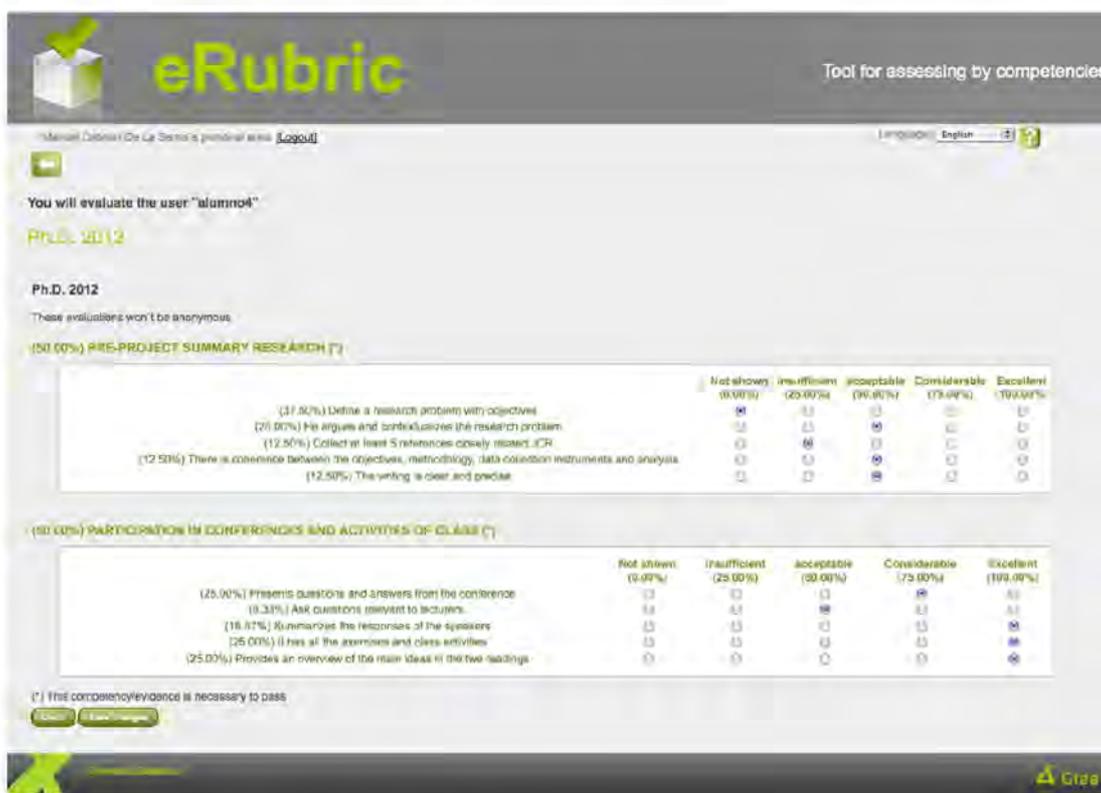


Figure 2. Assessment of competences using eRubrics

154 Teachers can follow their students' achievements and learning difficulties using two different
 155 approaches:
 156
 157
 158

159
160
161
162
163
164

- (a) By clicking on a student's name, teachers can see all the competences the student has acquired so far (Figure 2).
- (b) By clicking on a specific indicator, teachers can see which students in the class group have acquired partial competences and to what degree (Figure 3).



165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190

Figure 3. Assessment of several students by several evaluators

Additionally, the tool offers a graphical view of competences and evidence for the whole group. Teachers are able to access students' peer assessments in a graph and can easily check if these assessments (which are repeated over time) show a specific tendency. According to our experience, students' self-assessments tend to approximate the assessments provided by their teachers in the course of time (Serrano Angulo & Cebrián de la Serna, 2011). This graphical view also visualises the impact of the eRubric on students' learning over time.

The Gtea eRubric is continuously evolving and improving based on the contributions offered by users through the community of practice [9], showing to date the following major functions:

- It is interoperable with any other tool (wiki, blog, platform [LMS]) or system within the institution.
- eRubrics can have geolocations, allowing teachers to design programs within models for mobile learning (e.g. an eRubric for an educational walk through a park where each tree has a geolocated location on a rubric).
- Different criteria can be used and expressed either in numbers or in qualitative expressions.
- There is an improvement in the communication of meaning between evaluator and evaluated through multimedia annotations (explanatory text, video explanations, drawings and graphics).
- It facilitates student grouping and group assessment.
- The eRubric can be exported in PDF, while the results of the assessment can be exported in Excel for subsequent statistical analysis.
- Different teachers can assess the same group of students and each teacher can assess different competences.
- It has been translated into the different languages within the Spanish territory (e.g. Galician, Basque) as well as into other languages (e.g. English, French).

192

Methodology: evaluating the impact on learning with eRubrics

192

The first project to include eRubrics was entitled 'Federated eRubric Service to Assess University Learning' (2010-13).[10] The universities that collaborated in this project were the University of Malaga (coordinator), the University of Vigo, the University of Granada, the Polytechnic University (Madrid), the University of the Basque Country and Barcelona University. This project was followed by other projects that emerged, especially in Latin America, such as the 'Design Center eRubric' for the National Distance Education in Mexico, and the 'Supervision Practicum' project by a Swedish University and several Brazilian universities. These projects, institutions and countries have resulted in the formation of a community of practice on eRubrics [11], allowing for the exchange of experiences, good practices and collaborative projects.

193

194

195

196

197

198

199

200

The objectives of the project 'Federated eRubric Service to Assess University Learning' were:

202

1. Exploring an assessment methodology based on peer assessment and self-assessment using eRubric tools.

203

2. Understanding and assessing the use of the eRubric and its methodology in different universities with similar educational contexts.

204

205

3. Exploring identification technologies (authentication and federation technologies) that are available to students and professors from different educational institutions.

206

207

208

In order to achieve these objectives, different methodological approaches were used. First, by using a multimethod design, we applied quantitative methods to analyse, for example, self-assessments, peer assessments and teacher assessments obtained with the help of eRubrics and data from an online questionnaire. Qualitative methods were used to carry out interviews and analyse the resulting data. Quasi-experimental designs were used to analyse groups that did and did not use eRubrics. Second, given that the project offers a new technological environment as well as a new methodology for users, an online questionnaire was designed to assess the *usability* and *satisfaction* with eRubrics, by using the LimeSurvey software, adapted by a research team at the University of Malaga. The tool was validated with the SUS (System Usability Scale) survey in tests conducted prior to the project (Brooke, 1996).

209

210

211

212

213

214

215

216

217

218

Expected Results

219

At the moment, there are several interesting results and products. On one hand, Gtea eRubric is a validated tool and part of a service which is accessible by all the associated universities of a Spanish academic and research network called RedIris (453 institutions), by the National System of Distance Education using the Sined network (Mexico) [12], and by other institutions through EduGAIN.[13] Furthermore, data are currently being analysed from different groups (see also Raposo Rivas et al, 2014) and we expect the results to be released before the end of the project. Once the results are obtained, it is hoped that the benefits of the system and the effect it has on student learning (especially by self-assessment and peer assessment) can be shown.

220

221

222

223

224

225

226

The eRubric methodology and technique applied to the assessment process poses a huge professional challenge to both students and teachers. Students are gradually learning to use eRubrics through regular practice which helps them to internalise criteria as well as both formative and summative assessment standards.

227

228

229

230

231

Conclusion

232

The objective of this article was to present and discuss the concept of eRubrics **AS USED** in a federated system of higher education institutions, based on the research project 'Federated eRubric Service to Assess University Learning'. Apart from their use in the research project, eRubrics have already been tried in different sizes of student groups from small to large (Cebrián de la Serna, 2011). Depending on the size of each group and the design of the learning environment, the eRubric can be used more or less frequently, and thus **CAN** promote opportunities for self-regulation of learning (Mooij, 2009; Mooij et al, 2014). However, like any other tool, it has advantages and limitations:

233

234

235

236

237

238

239

- 240 *Advantages*
- 241 1. Creating and using eRubrics supports thinking about, reflecting on and internalizing
 - 242 assessment criteria.
 - 243 2. In an eRubric, formative and summative assessment criteria are defined clearly and objectively.
 - 244 3. Peer review helps students understand the criteria.
 - 245 4. This methodology requires students to be more engaged and to feel more responsible with
 - 246 regard to their learning.
 - 247 5. Overall, the eRubric encourages self-assessment and learning autonomy, thus facilitating
 - 248 dialogue, reflection and student–teacher feedback.
 - 249 6. Using federated technology, institutions **can collaborate with low cost, shared sources and**
 - 250 **tools CAN COLLABORATE, BENEFITING FROM THE LOWER COSTS INVOLVED IN**
 - 251 **SHARING RESOURCES AND TOOLS.** This technology extends the capabilities of the rubric
 - 252 to foster collaboration between users from different institutions and **OF DIFFERENT**
 - 253 technology systems in countries all over the world.

- 254 *Limitations*
- 255 1. It is difficult for teachers and students to move from traditional models of learning and
 - 256 assessment to new ones without going through a period of transition (see also Andrade, 2014;
 - 257 Mooij, 2014).
 - 258 2. Collecting, describing and interpreting **THE VARIOUS FORMS OF** evidence that will be used
 - 259 as indicators of competences are not easy tasks.
 - 260 3. The success of formative assessment with eRubrics will depend on the size of the student
 - 261 group, as formative evaluation using eRubrics in groups comprised of many students is
 - 262 difficult.

263 Appreciating these limitations and advantages, we can generally conclude that federated

264 technology should also be interoperable with other systems developed by multiple institutions for

265 potential collaboration and/or mobility of students and teachers, and this is a growing reality.

266 Finally, one should be able to access eRubrics from any Internet-connected device. Technologies

267 that allow access to a federated information system can be used to develop new models of learning

268 environments that are more personalised, as teachers are able to choose an eRubric from a list of

269 possible tools which belong either to their institution or to another institution within the federated

270 system and which are distributed via the Internet. By focusing technologies on characteristics of

271 users or groups, we are also *focusing the teaching process on learning rather than on technology.*

272 As a whole, eRubrics are a source of information for **THE PROCESS OF** evaluation.

273 However, evaluation is a complex task and needs other sources of information and records as well

274 (teachers' diaries, interviews, notes, etc.). The eRubric should not be seen as a technique or method

275 suitable for the whole evaluation process. **Its highest value is on ITS GREATEST VALUE IS IN**

276 **ALLOWING** remote monitoring: it provides clear criteria for communication and evaluation, **and**

277 **also in pre-service teacher training to encourage teaching AND IS ALSO VALUABLE IN PRE-**

278 **SERVICE TEACHER TRAINING IN ENCOURAGING TEACHING** from the point of view of

279 SRL (Panadero et al, 2013), which is in line with the European higher education model.

280 Notes

- 281 [1] The project 'Federated eRubric Service to Assess University Learning' (2010-13) was funded by the
- 282 Project 'Plan Nacional I + D + i 2010-2013. EDU2010-15432'. Resolution of 30 December 2009 (**BOE**
- 283 **SPELL OUT?** of 31 December 2009).
- 284 [2] Gtea: research group of **CONCERNED WITH?** technologies applied to education in Andalusia,
- 285 Spain (<http://gtea.uma.es>).
- 286 [3] Sined: National System of Distance Education in Mexico (<http://www.sined.mx>).
- 287 [4] EduGAIN: <http://www.geant.net/service/eduGAIN/Pages/home.aspx>
- 288 [5] (Web SSO; see <http://www.geant.net/service/eduGAIN/Pages/home.aspx>)

- 284 [6] (<http://www.ehea.info/article-details.aspx?ArticleId=73>)
- 290 [7] Federated virtual environment: <http://gteavirtual.org>
- 291 [8] The tool's eRubric: <http://gteavirtual.org/rubric>
- 292 [9] For information on all projects and the community of practice, see <http://erubrica.org>
- 293 [10] See note 1 above.
- 294 [11] See note 9 above.
- 295 [12] See note 3 above.
- 296 [13] See note 4 above.

297 References

- 298 Accino Dominguez, A. & Cebrián de la Serna, M. (2008) The Platform is the Net: user-centered learning and
299 identity-centered architectures, *Bulletin of RedIRIS*, 84. [http://gtea.uma.es/mcebrian/wordpress/wp-](http://gtea.uma.es/mcebrian/wordpress/wp-content/uploads/2012/07/rediris08eng.pdf)
300 [content/uploads/2012/07/rediris08eng.pdf](http://gtea.uma.es/mcebrian/wordpress/wp-content/uploads/2012/07/rediris08eng.pdf)
- 301 Andrade, M.S. (2014) Dialogue and Structure: enabling learner self-regulation in technology-enhanced
302 learning environments, *European Educational Research Journal*. **UPDATE PAGE NUMBERS**
- 303 Blanco, A. (Coord.) **PLEASE USE BRITISH EQUIVALENT TERM** (2009) *Desarrollo y evaluación de*
304 *competencias en Educación Superior* [Competence Development and Evaluation in Higher Education].
305 Madrid: Narcea.
- 306 Brooke, J. (1996) SUS: a 'quick and dirty' usability scale, in P.W. Jordan, B. Thomas, B.A. Weerdmeester &
307 A.L. McClelland (Eds) *Usability Evaluation in Industry*, pp. 189-194. London: Taylor & Francis.
308 <http://www.usabilitynet.org/trump/documents/Suschapt.doc>
- 309 Brown, S. & Glaser, A. (2003) *Evaluar en la universidad. Problemas y nuevos enfoques* [Assessment in the
310 Universities. Problems and New Approaches]. Madrid: Narcea.
- 311 Carneiro, R., Lefrere, P., Steffens, K. & Underwood, J. (2011) *Self-regulated Learning in Technology Enhanced*
312 *Learning Environments: a European perspective*. **Technology Enhanced Learning, V.5. IS THIS A SERIES**
313 **TITLE AND VOLUME NUMBER?** Rotterdam: Sense Publishers.
- 314 Castells, M. (2010) *The Information Age: economy, society and culture*. Vol. 1. *The Rise of the Network Society*, 2nd
315 edn. Oxford: Wiley Blackwell. **NOT CITED IN TEXT – PLEASE CITE OR REMOVE FROM REFS**
- 316 Cebrián de la Serna, M., Martínez, M.E., Gallego, M.J. & Raposo, M. (2011) E-rúbrica para la evaluación: una
317 experiencia de colaboración interuniversitaria en materia TIC [E-rubric for Evaluation: an experience of
318 inter-university collaboration in ICT]. Paper presented at 2nd **Congreso Internacional de Uso y Buenas**
319 **Prácticas con TIC, TRANSLATE THIS?** Malaga (Spain), 14-16 December. [http://erubrica.uma.es/wp-](http://erubrica.uma.es/wp-content/uploads/2011/06/Comunicaci%C3%B3n.pdf)
320 [content/uploads/2011/06/Comunicaci%C3%B3n.pdf](http://erubrica.uma.es/wp-content/uploads/2011/06/Comunicaci%C3%B3n.pdf)
- 321 European Higher Education Area (EHEA) Ministerial Conference (2012) Making the Most of Our Potential:
322 consolidating the European Higher Education Area. Bucharest Communiqué.
323 <http://www.ehea.info/article-details.aspx?ArticleId=11>
- 324 European Parliament (2006) Recommendation of the European Parliament and of the Council of 18
325 December 2006 on Key Competences for Lifelong Learning.
326 <http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:394:0010:0018:en:PDF>.
- 327 Hanrahan, S. & Hafner, P.H. (2003) Quantitative Analysis of the Rubric as an Assessment Tool: an empirical
328 study of student peer-group rating, *International Journal of Science Education*, 25(12), 1509-1528.
329 <http://dx.doi.org/10.1080/0950069022000038268>
- 330 López, V. (2009) *Evaluación formativa y compartida en Educación Superior* [Formative and Shared Assessment in
331 Higher Education]. Madrid: Narcea.
- 332 Mooij, T. (2007) *Learning for Self-regulation: improving instructional benefits for pupils, teachers, parents, schools,*
333 *and society at large*. Heerlen: Open University of the Netherlands (OUNL).
- 334 Mooij, T. (2009) Education and ICT-based Self-regulation in Learning: theory, design and implementation,
335 *Education and Information Technologies*, 14(1), 3-27. <http://dx.doi.org/10.1007/s10639-008-9066-8>
- 336 Mooij, T. (2014) Theoretical and Practical Aspects of Self-regulated Learning in Technology-enhanced
337 Preschools and Primary Schools, *European Educational Research Journal*. **UPDATE PAGE NUMBERS**
- 338 Mooij, T., Steffens, K. & Andrade, M.S. (2014) Self-regulated and Technology-enhanced Learning: a
339 European perspective, *European Educational Research Journal*. **UPDATE PAGE NUMBERS**

- 340 Panadero, E., Alonso-Tapia, J. & Reche, E. (2013) Rubrics vs. Self-assessment Scripts: Effect on Self-
 341 regulation, Performance and Self-efficacy in Pre-service Teachers, *Studies in Educational Evaluation*,
 342 V. **DOES THIS STAND FOR VOLUME? IF SO, OMIT** 39, 125-132.
- 343 Panadero, E. & Jonsson, A. (2013) The Use of Scoring Rubrics for Formative Assessment Purposes Revisited:
 344 a review, *Educational Research Review*, 9, 129-144. <http://dx.doi.org/10.1016/j.edurev.2013.01.002>
- 345 Prins, F.J., Sluijsmans, D.M.A, Kirschner, P.A. & Strijbos, J.W. (2005) Formative Peer Assessment in a CSCL
 346 Environment: a case study, *Assessment and Evaluation in Higher Education*, 30, 417-444.
 347 <http://dx.doi.org/10.1080/02602930500099219>
- 348 Raposo Rivas, M., Cebrián de la Serna, M. & Martínez-Figueira, E. (2014) Electronic Rubrics to Assess
 349 Competences in ICT Subjects, *European Educational Research Journal*. **UPDATE PAGE NUMBERS**
- 350 Rodríguez, G. & Ibarra, M.S. (2011) *E-Evaluación orientada al e-aprendizaje estratégico* [E-Assessment Oriented
 351 to Strategic e-Learning]. Madrid: Narcea.
- 352 Serrano Angulo, J. & Cebrián de la Serna, M. (2011) Study of the Impact on Student Learning using the
 353 eRubric Tool and Peer Assessment, in A. Mendez-Vilas (Ed.) *Education in a Technological World:
 354 communicating current and emerging research and technological efforts*. Badajoz, Spain: Formatex Research
 355 Center. <http://www.formatex.info/ict/book/421-427.pdf> **PLEASE CHECK THIS URL**
- 356 Steffens, K. & Underwood, J. (2008) Self-regulated Learning in a Digital World, *Technology, Pedagogy and
 357 Education*, 17(3), 167-170. <http://dx.doi.org/10.1080/14759390802383736>
- 358 Vickerman, P.H. (2009) Student Perspectives on Formative Peer Assessment: an attempt to deepen learning,
 359 *Assessment & Evaluation in Higher Education*, 34, 221-230. <http://dx.doi.org/10.1080/02602930801955986>
- 360 Zimmerman, B.J. (2000) Attaining Self-regulation: a social cognitive perspective. In M. Boekaerts, P. Pintrich
 361 & M. Zeidner (Eds) *Handbook of Self-regulation*. New York: Academic Press.
 362 <http://dx.doi.org/10.1016/B978-012109890-2/50031-7>

363
 364
 365
 366 **DANIEL CEBRIÁN ROBLES** studied as an industrial engineer and has been temporarily hired for
 367 a research project at Málaga University. He was a research fellow at Harvard University in 2013 and
 368 a research fellow at Nottingham University in 2012. He worked as a developer in research projects
 369 such as Gtea (<http://gtea.uma.es>; gteavirtual.org), and at the Center for Hellenic Studies at
 370 Harvard University to develop multimedia annotation tools (video, text, image and audio)
 371 (<http://openvideoannotation.org>). He is a PhD student in two areas: educational technology and
 372 fluid mechanics. *Correspondence:* dcebrian@uma.es
 373

374 **JOSÉ SERRANO ANGULO** is a professor of research methods and assessment in education at the
 375 Department of Research Methods and Educational Innovation at the University of Málaga, Spain.
 376 He was the coordinator of the commission that drew up the system of internal quality assurance of
 377 **FOR** the Faculty of Education Sciences. He has been director of the Methods and Resources for
 378 Educational Research and Innovation research group. He has directed and/or participated in
 379 research projects involving international cooperation, research projects of excellence, **I+D+I**
 380 **SPELL OUT/EXPLAIN?** projects and educational innovation projects. His current research
 381 interests are educational measurement, evaluation, indicators of educational quality, and statistical
 382 analysis applied to educational research. He is the director of the Educational Research and
 383 Innovation PhD programme. **EMAIL ADDRESS?**
 384

385 **MANUEL CEBRIÁN DE LA SERNA** is a professor in education at Málaga University, Spain. He is
 386 the director of the research group Gtea (Globalization, technology, education and learning)
 387 (<http://gtea.uma.es>). He was the director of the Institute of Science Education from 1994 to 1998
 388 and of **THE INSTITUTE OF** Educational Innovation from 1998 to 2000. From 2000 to 2003 he
 389 was the director of virtual education at the University of Málaga. Furthermore, he was National
 390 Assessment Evaluator (ANEP) **IS ANEP THE NAME OF THE ORGANISATION?** for **Research**
 391 **and Development FOR WHICH ORGANISATION?** and other agencies in Spain. He has the
 392 award 'Order of Puerto Ayacucho', granted by the City Council of Puerto Ayacucho, the capital of
 393 Amazonas State. He is the promoter of educational innovation with the Iberoamerican Network of
 394 Higher Education. Moreover, he was the manager of a **major production of MAJOR**

326
396 **PRODUCTION COMPANY INVOLVED WITH** audiovisual and multimedia and technological developments **KNOWN AS** 'Open Federated Environment'. **EMAIL ADDRESS?**

Artículo 2: Usabilidad y Satisfacción de la e-Rúbrica

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.

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.

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

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 maquina, 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 analizarán 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. Seguido 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

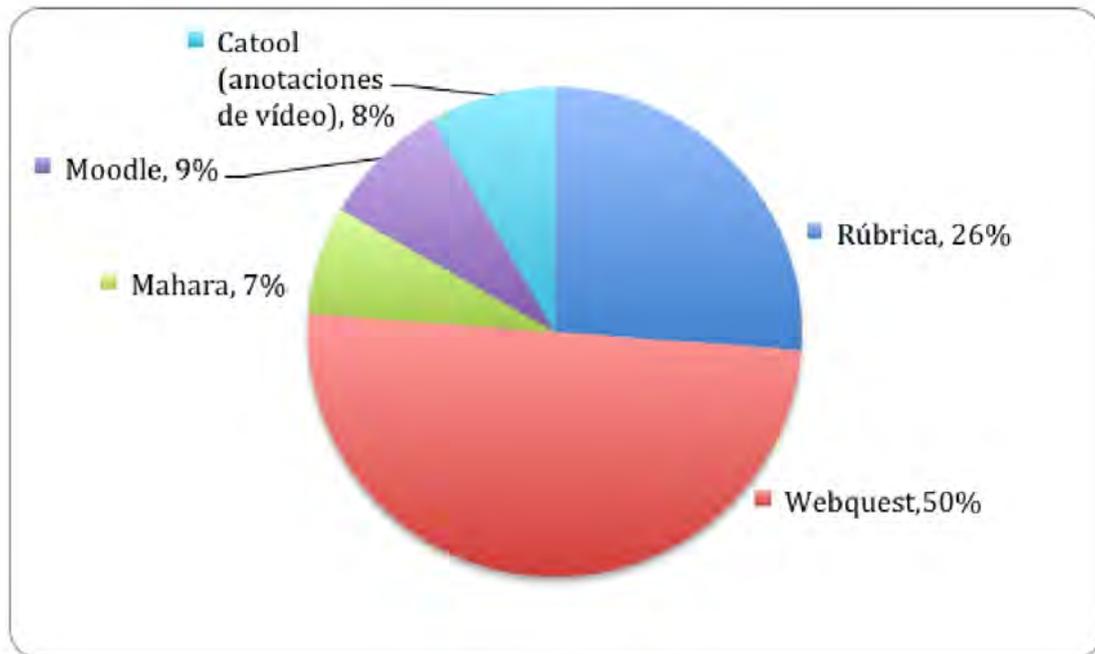
Para el análisis de los datos se revierten los valores correspondientes a los enunciados indirectos, para que no se contrarresten en los distintos análisis estadísticos, y en la puntuación de la suma total. Así se hacen los cambios oportunos para transformar las escalas indirectas en directas.

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% We-bquest, y el resto se reparte entre Mahara, Moodle y Catool. El reparto de hombres y

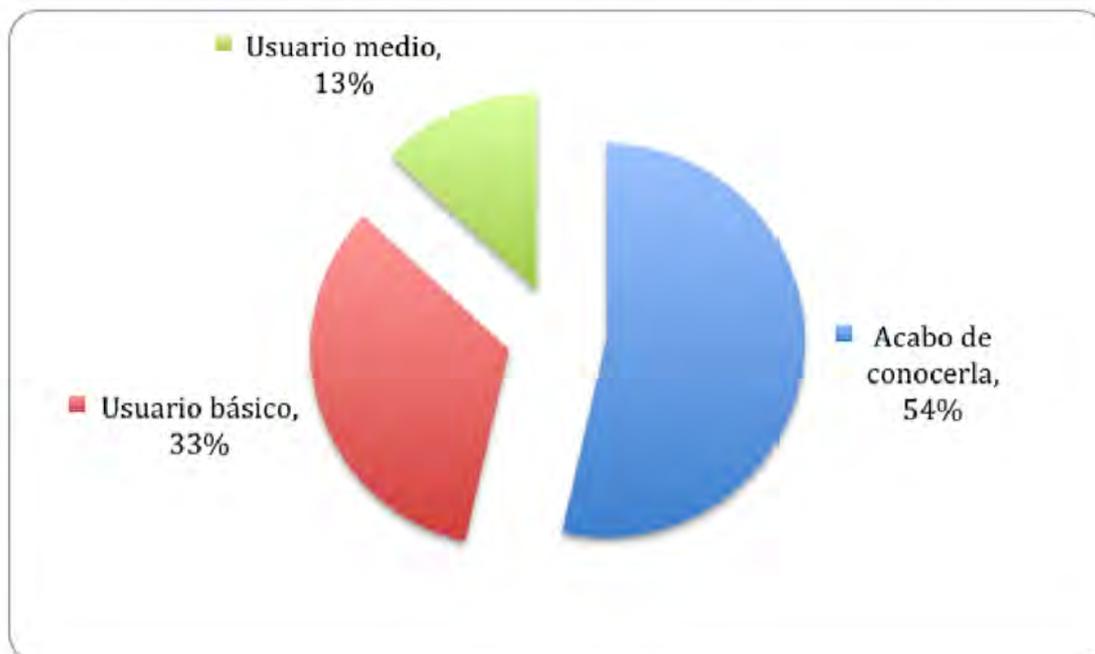
mujeres se mantienen por igual en todas las variables, asimismo ocurre con la edad.



Fuente: elaboración propia

Gráfico 1: Herramienta a evaluar.

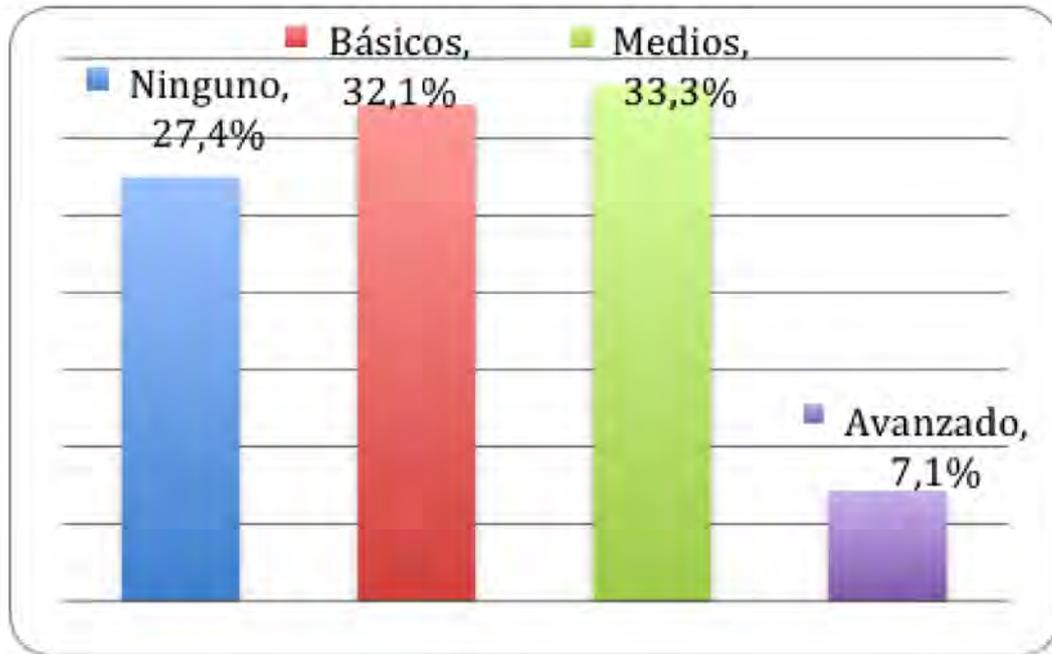
Los conocimientos que tenían de la herramienta a evaluar se muestran en el Gráfico 2, destacando que el 54% dicen no conocerla con anterioridad, lo que es favorable para nuestro objetivo de evaluar la usabilidad y la satisfacción sobre la misma.



Fuente: elaboración propia

Gráfico 1: Conocimientos sobre la herramienta.

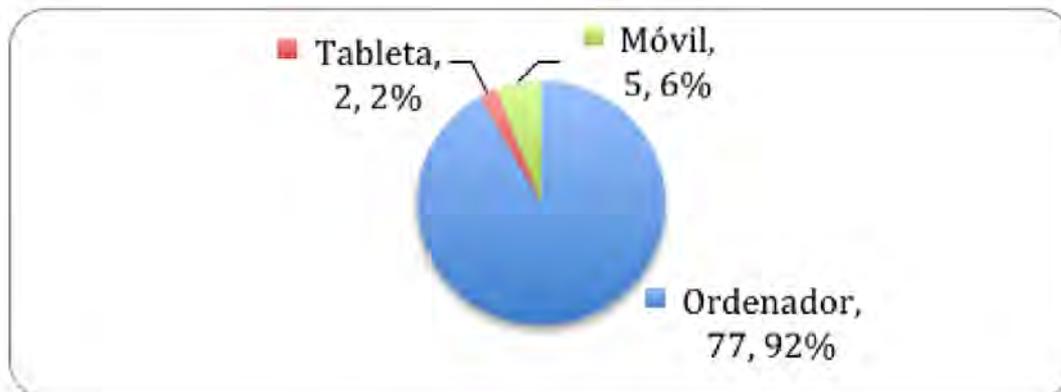
El conocimiento previo de herramientas similares puede inducir a que les resulten más fácil el manejo de estas herramientas, por ello se preguntaba sobre esta cuestión. En el Gráfico 3 se muestra la distribución de la experiencia con herramienta similares, como se aprecia quienes manifiestan no tener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



Fuente: elaboración propia

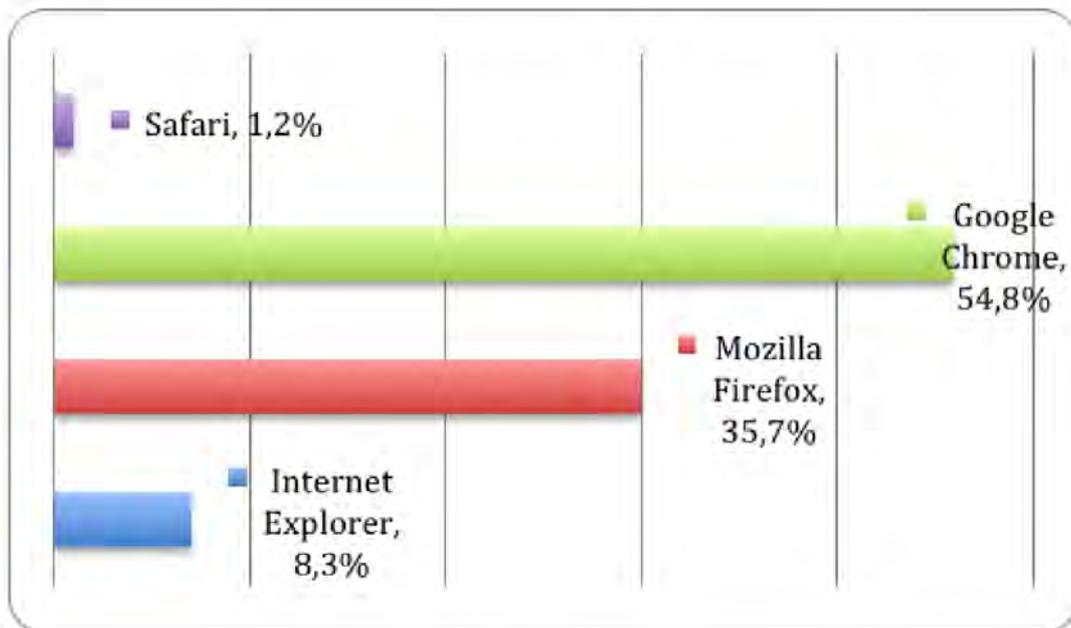
Gráfico 3: Conocimiento de herramientas similares en internet.

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.

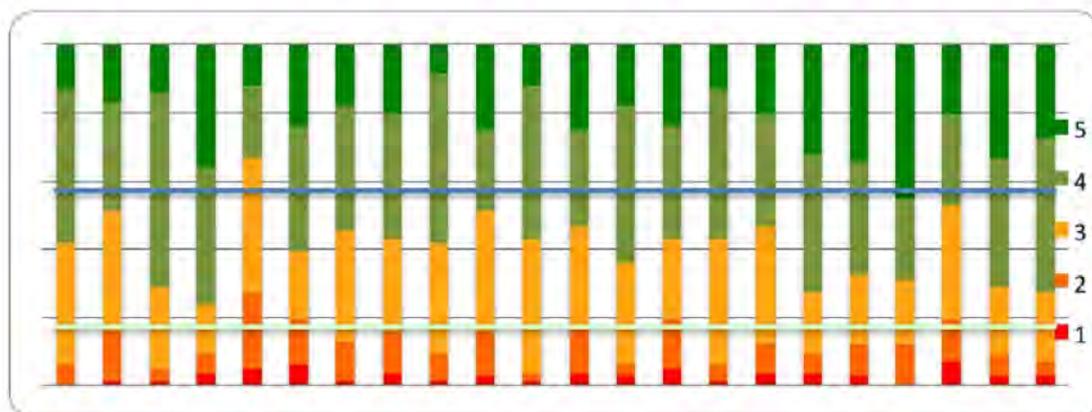


Fuente: elaboración propia

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.

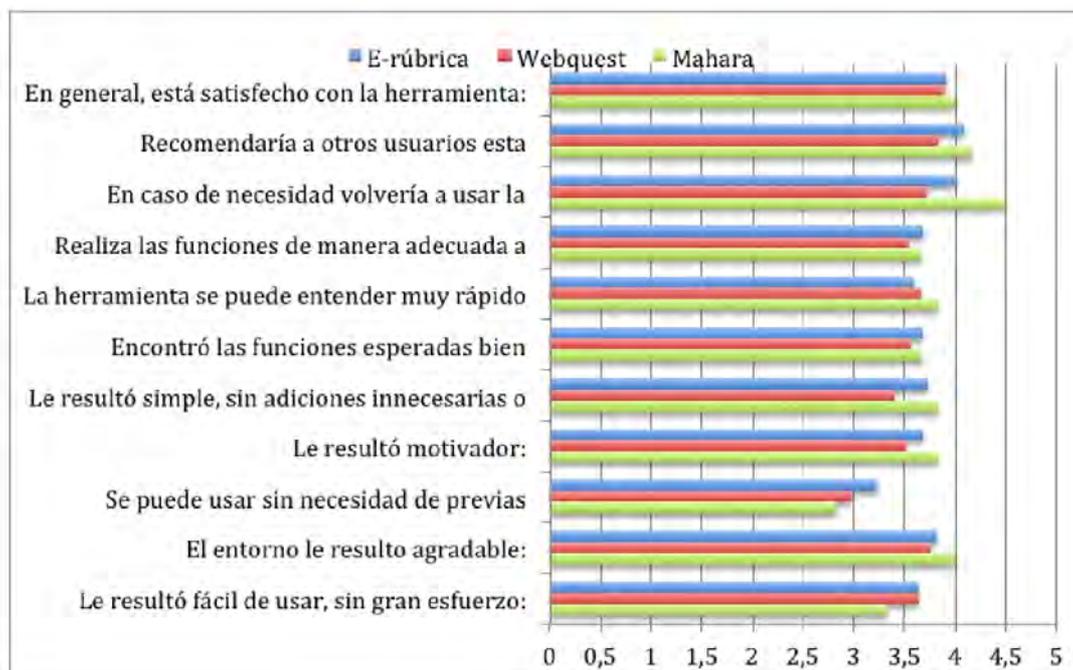
	Media	Desv. típica	Alfa de Cronbach
e-Rúbrica	81,23	12,22	0,891
Webquest	79,44	12,769	0,901
Mahara	82,33	10,633	0,865
Moodle	84,71	12,203	0,915
Catool	80,57	9,589	0,813
Total	80,66	12,031	0,889

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.

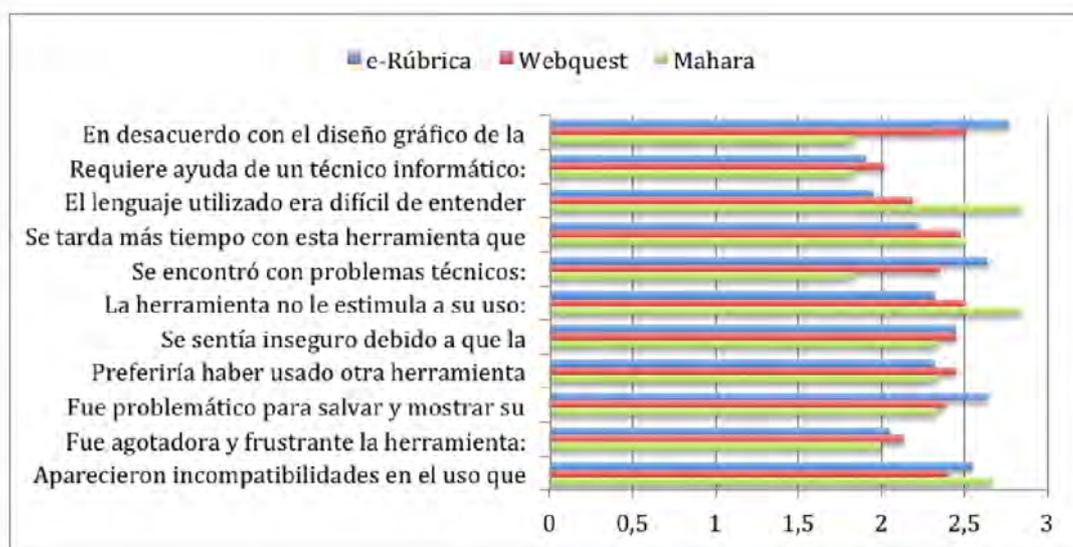
Un análisis más detallado se representa en los siguientes Gráficos en los que se muestran las medias en los enunciados directos (Gráfico 7) e indirectos (Gráfico 8) para las herramientas e-Rúbrica, Webquest y Mahara. En los enunciados directos y en el caso

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 docência* (417-429). CIIE: 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). CIEE: 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). CIEE: 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). CIEE: 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.

Artículo 3: Evaluación de pares y autoevaluación con eRúbricas: caso de estudio en el grado de Educación Primaria.

Evaluación de pares y autoevaluación con erúbricas: caso de estudio en el grado de Educación Primaria.

Peer assessment and self-evaluation with erubrics: a case study in the Degree in Early Childhood Education.

Rafael Pérez Galán

Daniel Cebrián Robles

Ana B. Rueda Galiano

Universidad de Málaga, España

Rafael Pérez Galán

Daniel Cebrián Robles

Ana B. Rueda Galiano

Universidad de Málaga, España

Resumen

Este trabajo trata de analizar hasta qué punto los alumnos son capaces de evaluar sus propios aprendizajes y los de sus compañeros (evaluación entre pares y autoevaluación) a través de la erúbrica, conociendo estándares y criterios de evaluación propuestos por el profesorado en la configuración, diseño y desarrollo de la evaluación de proyectos. Esta forma de evaluar y autoevaluarse presupone en el Nuevo Espacio Europeo de Educación Superior que el alumno asuma un papel más activo en el proceso de enseñanza

Abstract

This paper analyzes the extent to which students are able to assess their own learning and that of their peers (peer assessment and self-evaluation) through erubrica, standards and evaluation criteria proposed by the faculty in the configuration, design and development of project evaluation. This type of evaluation and self-assessment presumes in the New European space of Higher Education that students take a more active role in the learning process and the evaluation itself will serve in helping

y que la propia evaluación le sirva de aprendizaje en la consecución de las competencias profesionales. Así el uso de las nuevas tecnologías en procesos de evaluación autorregulada permite que el alumno alcance un rendimiento óptimo, eficaz y de calidad tanto en la propuesta de contenidos, como en el desarrollo de los mismos ante el grupo, considerando el acto de evaluar como un proceso de mejora y de aprendizaje, a través de herramientas como la erúbrica que favorecen la “evaluación total”; es decir, la asunción de todas las competencias básicas.

Este estudio se ha llevado a cabo en el curso académico 2011-2012, durante dos semestres con asignaturas troncales distintas (Teoría y Política de la Educación Infantil y Hacia una escuela inclusiva: modelos y prácticas) y con idénticos grupos (Grupo de 1º de grado de infantil A-grupo experimental- y Grupo de 1º de grado de infantil B-grupo control).

El presente trabajo expone los resultados de la evaluación entre pares y la autoevaluación a través de la puesta en práctica de proyectos, donde los alumnos evalúan a sus compañeros y se autoevalúan, comparando los resultados con la evaluación que realiza el profesor de los grupos y de cada alumno de forma individual.

Palabras clave: Evaluación entre pares, rúbricas, evaluación formativa, actuaciones de éxito, autoevaluación y Aprendizaje basado en proyectos.

to achieve the knowledge of professional skills. Thus the use of new technologies in self-regulatory assessment processes allows the student to achieve an optimal, effective and quality performance in the proposed content, as much as in the development of them before the group, considering the act of evaluation as a process of improvement and learning through tools such as the erubrica that favor the “overall assessment”; that is, the assumption of all core competencies. This study has been carried out in the academic year 2011-2012, for two semesters with different core subjects (theory and child education policy and towards an inclusive school: patterns and practices) and with identical groups (Group 1, 1st degree of infant A as the experimental group; and Group 2, 1st degree of infant B, as the control group). This paper exposes the results of peer evaluation and self-evaluation through the implementation of projects, where students evaluate their peers and evaluate themselves by comparing the results with the assessment made by the teacher with their groups and each student individually.

Key words: Peer assessment/evaluation, rubrics, formative assessment/evaluation, successful actions, self-assessment/evaluation and project-based learning.

Introducción

La evaluación es, y sigue siendo, uno de los elementos de mayor controversia en el diseño y desarrollo del currículo. Cada docente la utiliza de forma diferente según la convicción teórico-práctica que tiene y al servicio de qué o de quién la pone. Para evaluar hay que comprender, dice taxativamente Stenhouse (1984). No hay que olvidar que lo más importante, afirma Santos Guerra (1999), no es el hecho de realizar la evaluación, ni siquiera el modo de hacerlo, sino al servicio de quién se sirve. Por ello, la evaluación que deseamos, al margen de otras muchas posibles, está centrada y parte del propio participante, dándole voz a los interesados para que provoque un aprendizaje constructivista y más autónomo sobre el mismo proceso y acto de la evaluación y, también, sobre las evidencias y criterios que se ponen en funcionamiento. Al mismo tiempo, nuestros alumnos se comprometen e interesan no sólo por el mecanismo externo de la evaluación sino por su comprensión interna que les permita un aprendizaje continuo y verdadero a lo largo de la vida.

En este mismo sentido, también se ha manifestado el Espacio Europeo de Educación Superior, en cuanto que subraya y hace hincapié en la evaluación centrada en las competencias de logro por parte de los alumnos, verdaderos protagonistas del proceso de enseñanza-aprendizaje. La nueva sociedad del conocimiento y, por ende, las propias universidades, como entornos reales y virtuales de aprendizaje, han trasladado todo el protagonismo al alumno, lo que exige que los docentes varíen y cambien sus metodologías y estrategias con el fin de facilitar el conocimiento en los estudiantes y su transferencia a contextos diversos. El alumno debe ser el protagonista de su propio aprendizaje, y lo debe ser con la asunción de un liderazgo tanto en la decisión de las herramientas a utilizar en su aprendizaje como de qué contenidos seleccionar e integrar. El docente por su parte debe orientar, acompañar y guiar la construcción autónoma del aprendizaje del alumno (Andreu Barrachina, Sanz Torrent y Serrat Sellabona, 2009).

Por tanto, la asunción de estas competencias, a través de la formalización de estrategias evaluativas deben orientarse a dar capacidad y poder a los actores para vivir lo mejor posible, defender sus derechos y autonomía, respetando en todo momento a los demás (Simona Rychen y Hers Salganik, 2006).

Sin embargo, todo este entramado teórico-práctico no ha estado exento de críticas y de algunas insuficiencias sustanciales en cuanto al uso de herramientas e instrumentos que favorezcan realmente procesos auténticos evaluativos que mejoren tanto el desarrollo profesional del docente como la formación y desarrollo integral del alumno. Las nuevas metodologías surgidas en el siglo XXI, a través de los procesos de globalización, han modificado radicalmente la operatividad de las funciones legítimas tanto de los profesores como las de los estudiantes, y ambos están condenados a entenderse en esta nueva evaluación, más democrática y auténtica. El proceso de convergencia hacia el Espacio Europeo de Educación superior (EES) implica una serie de cambios en la forma de entender, organizar y llevar a cabo la docencia universitaria. En la mayor parte de los casos, la clave parece estar en que la formación universitaria cambia su centro de atención de los procesos de enseñanza (impartidos por el profesor) a los procesos de aprendizaje (desarrollados por el alumnado). Lógicamente, este tipo de cambios afecta a todos los elementos curriculares en mayor o menor medida. La evaluación es uno de los elementos que se ve más seriamente afectado y que más debería evolucionar para ajustarse a esta forma de entender y desarrollar la docencia universitaria. Por otra parte, es conveniente tener siempre presente que ninguna innovación curricular será efectiva

si no va acompañada de innovaciones en el modo de concebir la evaluación (Bonsón y Benito, 2005).

Este modo de evaluación contextualizada a través del planteamiento de situaciones relevantes en la vida real y problemas de naturaleza compleja se ha llamado “*evaluación auténtica*” (Torres y Perera, 2010; Diaz y otros, 2009; Gómez y otros, 2009; Archbald y Newman, 1988; Airasian, 2001). Y esta evaluación es la que desarrolla sus principios a través de la evaluación formativa ya que es la que acompaña en todo momento al alumno y la que mejora la comprensión del proceso por parte del maestro, aunque algunas veces dificulte su desarrollo el excesivo número de alumnos por grupo, algo que se contradice con el nuevo ajuste y espíritu del Plan Bolonia, a través de sus índices o coeficientes de experimentalidad, que no es otra cosa que adecuar porcentualmente la razón más justa para el desarrollo de buenas prácticas entre número de alumnos y profesor con el fin de ofrecer una formación de calidad para los nuevos profesionales. La elección de una modalidad organizativa es inherente a la libertad de cátedra de un profesor, pero tiene sus limitaciones, en función de la naturaleza de la materia y los recursos con los que se cuenta. Dicha reflexión debería contemplar, entre otras cuestiones, el grado de experimentalidad o practicidad profesional de cada título y materia; así como su coeficiente de agrupamiento; es decir, en qué medida se considera necesario trabajar con grupos de alumnos de tamaño relativamente pequeño. (Montanero, M.; Mateos, V.L; Gómez, V. y Alejo, R., 2006). De igual forma lo expresan Cebrián, Raposo y Accino (2007), en estudios llevados en relación al aprendizaje en entornos virtuales como plataformas apropiadas para el desarrollo de competencias profesionales en grupos reducidos. Tanto la enseñanza a distancia requiere de espacios y momentos presenciales (véase la red de centros de la UNED repartidos por el territorio en España) como la enseñanza presencial cuando utiliza espacios virtuales (caso de las plataformas virtuales y el tiempo de crédito asignado a tareas de estudiantes) requiere nueva consideración en la ratio. La enseñanza presencial mediada con virtualidad debería poseer una ratio docente menor, y esto nos permitiría seguir con más personalización el aprendizaje de los estudiantes.

No obstante, esta evaluación la encontramos hoy ampliamente desarrollada a través de la herramienta más operativa y funcional del mercado como es la erúbrica o matriz de valoración, que va a ser el instrumento básico en el desarrollo de este trabajo ya que avalan tanto la satisfacción por parte de los usuarios como su usabilidad (Serrano Angulo & Cebrián Robles 2012; 2014; Gallego-Arrufat & Raposo-Rivas 2014). Usabilidad y satisfacción de los usuarios que están muy estrechamente relacionados con nuestro objetivo según los autores anteriores, como es demostrar que el empleo de la erúbrica facilita en los estudiantes un proceso de autoevaluación y evaluación.

Una de las vías para lograr una evaluación de mayor calidad, de carácter más cualitativo y procesual, es el empleo de protocolos de evaluación, o rúbricas. Estos protocolos, no sólo favorecen una evaluación más sistematizada por parte del docente, sino que son una herramienta de extraordinario valor para el desarrollo de competencias de monitorización, autoevaluación y evaluación entre pares, contribuyendo a un mayor entendimiento del propio proceso de aprendizaje y, en definitiva, a una mayor autonomía y autorregulación del aprendizaje (Stevens y Levi, 2005).

Dos procedimientos metodológicos propios de la evaluación formativa que están imbricados en el desarrollo de este trabajo son: por un lado, *la evaluación entre pares* y *la autoevaluación*.

La evaluación entre iguales puede entenderse como una forma específica de aprendizaje colaborativo (Brown & Glasner 2003; López-Pastor 2009; Blanco 2009) en el que los aprendices realizan una valoración sobre el proceso o producto de aprendizaje de todos o de algún estudiante o grupo de estudiantes. Así pues, la evaluación entre iguales se puede utilizar con éxito en cualquier disciplina, área y nivel. Siguiendo los estudios llevados a cabo por Flecha (2012), en su proyecto « INCLUD-ED », podemos afirmar también, que la érubrica, como las estrategias de evaluación entre pares y la autoevaluación, conjuntamente con la formación de grupos interactivos y heterogéneos en clase, se les puede dar también el calificativo de actuaciones de éxito en distintos contextos y situaciones relevantes de enseñanza-aprendizaje, donde la apuesta firme se sitúa en el aprendizaje de calidad de todos, sin excluir a nadie (Prins, Sluijsmans, Kirschner & Strijbos, 2005; Panadero & Jonsson 2013; Cebrián de la Serna, Serrano Angulo & Ruiz Torres 2014). Así pues, para plantear y ejecutar con éxito en el aula experiencias de evaluación entre iguales, la metodología docente debe ser flexible y estar abierta a la colaboración y la participación de los estudiantes (Bretones, 2008). Por tanto, a su vez será conveniente favorecer el desarrollo de ambientes de aprendizaje auténticos donde se compartan significados y se construya conjuntamente el conocimiento, ofreciendo, además, amplias oportunidades de interacción social (Keppell y otros, 2006).

La autoevaluación, por su parte, es primordial en el desarrollo y adquisición de las nuevas competencias según la configuración del Nuevo Espacio de Educación Superior Europeo. La implicación del alumno en todo el proceso de evaluación (desde el principio hasta el final) permite que el alumno conozca los estándares que propone el profesor, su grado de dificultad como los medios que hay que proponer para alcanzar con éxito los resultados previstos. Por tanto, la evaluación también les debe ayudar a ellos a ser más conscientes sobre cuál es o debe ser el nivel de adquisición de competencias que van alcanzando, qué puntos fuertes tienen y qué puntos débiles deben corregir (Boekaerts, Pintrich y Zeidner, 2000). Este aspecto de la evaluación pone su énfasis en la autoevaluación, en la reflexión del estudiante sobre su proceso de formación continua, lo que le permite autorregular dicho proceso hacia aquellos elementos más deficitarios (Díaz y otros, 2009).

De esta forma, la érubrica es una de las formas más interesantes de especificar los criterios de calidad de la evaluación de competencias que facilitan la calificación del trabajo del estudiante especialmente cuando se trata de tareas complejas, imprecisas y subjetivas. Dado que las érubrica bien hechas obligan a definir evidencias más objetivas (Cebrián de la Serna, Serrano Angulo & Ruiz Torres 2014).

La rúbrica puede definirse como un conjunto de criterios específicos que permiten valorar los aprendizajes, los conocimientos y las competencias logradas por el estudiante en un trabajo, exposición, diseño colaborativo, tareas o asignaturas. Existiendo estudios específicos en contexto de educación inicial de enseñantes donde la autoevaluación, la autorregulación y las érubricas están relacionadas con las tareas de enseñanza bajo guiones bien explícitos (Panadero, Alonso-Tapia & Reche 2013). La e-rúbrica es una herramienta que ofrece a los estudiantes información acerca de las competencias que se esperan de ellos, junto con los “indicadores”, criterios o evidencias que le informan de lo que tienen que hacer para lograr estas competencias (Cebrián M., Raposo, M Y Accino, J.; 2007). De esta forma, los estudiantes antes del inicio del curso, de las tareas, de los proyectos, etc. disponen de una información más precisa y detallada.

La matriz o rúbrica de evaluación sirve para averiguar cómo está aprendiendo el alumno, pudiendo considerarse como una herramienta de evaluación formativa. (Andrade, 2008; Conde y Pozuelo, 2007 y Hafner y Hafner, 2003). Así pues, y siguiendo a Stevens y Levi (2005), podemos afirmar que una de las vías para lograr una evaluación de mayor calidad, de carácter más cualitativo y procesual, es el empleo de protocolos de evaluación o rúbricas. Estos protocolos, no solo favorecen una evaluación más sistematizada por parte del docente, sino que son una herramienta de extraordinario valor para el desarrollo de competencias de monitorización, autoevaluación y evaluación entre pares, contribuyendo a un mayor entendimiento del propio proceso de aprendizaje y, en definitiva, a una mayor autonomía y autorregulación del estudiante. Por tanto, consideramos cómo las nuevas líneas de investigación centradas en los procesos de aprendizaje de los alumnos deben estar orientados a través de herramientas como la erúbrica que potencia la autorregulación de los alumnos en su proceso de aprendizaje y le facilita compartir estas experiencias a través de grupos heterogéneos e interactivos en aras de una enseñanza de calidad para todos y a lo largo de toda la vida; también a través de entornos virtuales o de entornos semipresenciales o learning blended. Así lo demuestran diferentes estudios que afirman el beneficio de esta herramienta, la evaluación entre pares y la evaluación entre iguales, en la consecución de actuaciones de éxito en contextos diferentes (Falchikov & Goldfinch 2000; Andrade & Du 2005; Marcos, Tamez & Lozano 2009).

La experiencia

En el presente documento se va a tratar de exponer la experiencia metodológica llevada a cabo en la Escuela de Magisterio “María Inmaculada” de Antequera, escuela adscrita a la Universidad de Málaga. Siendo un proyecto a dos cursos académicos fueron empleados soportes tecnológicos diferentes, en una primera fase se utilizó rúbricas con el software Limesurvey, mientras se diseñaba otra herramienta específica para albergar estas rúbricas con un software propio del proyecto de investigación en el que se inscribe esta experiencia¹.

Para comenzar a trabajar la evaluación a través de la herramienta de la e-rúbrica, se seleccionaron dos asignaturas troncales pertenecientes al 1º Grado de Educación Infantil. Tales materias, con 6 créditos de carga lectiva eran: “Teoría y Política de la Educación Infantil”, en el primer semestre, y “Hacia una escuela inclusiva: modelos y prácticas” en el segundo semestre del curso académico 2011-2012.

En ambos casos se decidió trabajar con dos grupos (control y experimental), de forma que la variable independiente era la erúbrica utilizada y la dependiente los resultados académicos. Por su parte, el grupo de control estaba formado por 5 alumnos y 48 alumnas, con edades comprendidas entre los 18 y 20 años. Éste fue el grupo no sometido a la evaluación mediante e-rúbricas, la evaluación de este grupo se caracterizaba por la exposición de los proyectos sin ninguna guía inicial por parte del profesor tanto en el diseño, ejecución y puesta en escena del contenido de los mismos: buscaban la forma, el método y las estrategias necesarias sin saber nada de las evidencias significativas

¹ Este proyecto se realizó bajo el amparo del proyecto I+D+I titulado “Servicio Federado de eRúbrica para la evaluación de aprendizajes universitarios, Proyecto del Plan Nacional I+D+I 2010-2013. EDU 2010-1543. Resolución de 31 de diciembre de 2009 (-BOE 31/12/2009-. SECRETARÍA DE ESTADO DE INVESTIGACIÓN DIRECCIÓN GENERAL DE INVESTIGACIÓN Y GESTIÓN DEL PLAN NACIONAL DE I+D+I

que servirían para una evaluación más total y comprometida con sus competencias profesionales.

El grupo experimental, estaba formado por 53 alumnos, de edades comprendidas entre los 18 y 20 años, procedentes de bachillerato, módulos profesionales y otras titulaciones. Este grupo estaba formado por 3 alumnos y 50 alumnas.

Para el trabajo de exposición de proyectos (donde se utilizaba la rúbrica) se empleó una metodología consistente en un aprendizaje basado en proyectos, de modo que permitiera el intercambio de experiencias, el debate y el flujo de opiniones, la búsqueda de objetivos comunes... todo ello con evaluación entre pares y autoevaluación mediante rúbricas.

En su conjunto, la e-rúbrica de exposición de proyectos estaba basada en 6 competencias instrumentales que, a su vez, versaban sobre evidencias del siguiente tipo: Habilidades y gestión de recursos tecnológicos; Expresión oral y uso de diferentes técnicas; Utilización de sistemas de aprendizaje a través del uso de idiomas de la Comunidad Europea; Producción de materiales y recursos tic; Desarrollo de capacidades de colaboración y Participación activa y liderazgo en el trabajo de equipo. Todo ello en la idea de responder a las competencias genéricas y su evaluación (Blanco 2009), especialmente las del trabajo en equipo. También, por su parte, existían una serie de evidencias más centradas en el análisis de proyectos educativos: - Presentación del modelo didáctico/caso o trastorno; Dominio del contenido; Contextualización y Transferencia del conocimiento.

En primer lugar, y antes de presentar todo el trabajo, se pasó a la explicación y simulación de la tarea de forma global así como la formalización de los grupos. Se organizaron 7 grupos con componentes comprendidos entre 7-8 alumnos, iguales para los dos semestres, aunque hubo un cambio de un componente de un grupo a otro en el segundo semestre, en el grupo A. Tanto en el primer semestre como en el segundo, nos centramos en el trabajo y presentación al grupo de un modelo didáctico y un caso/trastorno de diversidad funcional que se trabaja en Educación Infantil de la siguiente forma:

- Elección de los modelos didácticos y casos/trastornos de diversidad funcional
- Estudio detallado de las ideas básicas del proyecto
- Directrices generales de organización y gestión de la información
- Reflexión de grupo

Una vez realizado este primer acercamiento al modelo y por ende al caso/trastorno, y haber estado varias sesiones profundizando en la temática de los proyectos, ideas básicas y contenidos de los textos, por grupo clase, se llevó a cabo la explicación del guión-trabajo a través de la exposición de toda la e-rúbrica en clase, detallando y explicando cada una de las competencias así como las evidencias y estándares de evaluación donde íbamos a poner el acento e importancia en las exposiciones. La e-rúbrica es una herramienta que ofrece a los estudiantes información acerca de las competencias que se esperan de ellos, junto con los "indicadores", criterios o evidencias que le informan de lo que tienen que hacer para lograr estas competencias (Cebrián M., Raposo, M. y Accino, J.; 2007). De esta forma, los estudiantes antes del inicio del curso, de las tareas, de los proyectos, etc. disponen de una información más precisa y detallada. A la vez, y durante

el proceso (sobre todo cuando es una e-rúbrica) pueden disponer de una comunicación más inmediata sobre sus avances.

En el segundo semestre, todo el procedimiento fue similar al llevado a cabo en el 1º semestre, aunque con algunas diferencias como:

- Los alumnos expusieron, como se ha ido comentando, en lugar de un modelo didáctico un caso/trastorno de diversidad funcional en educación infantil, ya que se trataba de la materia “Hacia una escuela inclusiva: modelos y prácticas”, siguiendo todo el protocolo del primer semestre. El cambio de contenidos responde a un nuevo semestre donde se introduce una nueva materia, como se ha dicho, pero se sigue el mismo protocolo de intervención que en el anterior: idéntica población, grupo experimental y control, idénticas competencias... no interfiriendo en los resultados académicos los saberes culturales de cada asignatura, ya que aunque eran conceptos diferentes las temáticas no interfirieron en el proceso, sino más bien el conocimiento de los estándares y criterios dados con anticipación en la erúbrica, como queda reflejado en el estudio.
- La herramienta que se utilizó fue la plataforma eRúbrica de Gtea², que consistió en asignar componentes a cada uno de los grupos formalizados a través de una contraseña con el fin de que los alumnos se iniciaran en todo el proceso, siendo partícipes directos en la composición de los datos y evaluaran directamente a los grupos a través de la herramienta y también hicieran su autoevaluación.

La pretensión de este trabajo consistió en lo siguiente:

- Comparar la evaluación del grupo A (Experimental) y grupo B (Control) a través de las herramientas “LimeSurvey Y Gtea” conociendo unos los estándares y criterios de evaluación y otros no, en la exposición y asimilación de contenidos, su presentación...
- Comparar la evaluación por proyectos hecha por el profesor y la realizada por los alumnos---ACERCAMIENTO-DISPERSIÓN
- Comparar la evaluación de los grupos A y B del 1º y 2º semestre a través de las notas de proyectos y notas finales
- Comparar la evaluación entre pares y la realizada por uno mismo (AUTOEVALUACIÓN)

Datos

Llegados a este punto se ha de recordar que los datos con los que se trabajan son pertenecientes a una población, formada por dos grupos de la misma Titulación (1º de Infantil, A y B), con asignaturas distintas en el primer semestre y en el segundo (Primer semestre: Teoría y Política de la Educación Infantil y Segundo semestre: Hacia una escuela inclusiva: modelos y prácticas. El grupo A será el grupo control y el B el grupo experimental, tanto en el primer semestre como en el segundo, y ambos grupos de n=53.

² ERúbrica del grupo Gtea (Grupo de investigación Globalización, tecnología, educación y aprendizaje) de la Universidad de Málaga. <http://gteavirtual.org/rubric>

	GC	GE
Nº Alumnos	53	53
Sexo femenino	50	48
Sexo masculino	3	5

Tabla 1. Participantes

En la materia “ Teoría y Política de la Educación”, en el grupo de control (A) y el grupo experimental (B) las notas otorgadas por el profesor hacen una media de 8.5 y 9 respectivamente.

Analizando la distribución de frecuencias obtenida, podemos observar que mientras que el grupo de control las notas oscilan entre 5 y 9,5; en el Grupo Experimental las notas se encuentran enmarcadas entre 8 y 9,5.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 5,0	1	1,9	1,9	1,9
7,5	13	24,5	24,5	26,4
8,0	1	1,9	1,9	28,3
8,5	14	26,4	26,4	54,7
9,0	16	30,2	30,2	84,9
9,5	8	15,1	15,1	100,0
Total	53	100,0	100,0	

Tabla 2. Distribución de Frecuencias sobre Notas Obtenidas en Teoría y Política de la Educación Infantil. Grupo de Control.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 8,0	8	15,1	15,1	15,1
8,5	14	26,4	26,4	41,5
9,0	24	45,3	45,3	86,8
9,5	7	13,2	13,2	100,0
Total	53	100,0	100,0	

Tabla 3. Distribución de Frecuencias. Notas Obtenidas en Teoría y Política de la Educación Infantil. Grupo Experimental.

Hasta aquí, podríamos afirmar que la diferencia de medias entre las notas otorgadas por el profesor de la materia tanto en el grupo experimental como en el de control no es significativa: 8.5 en el Grupo de Control y 9 en el Grupo Experimental.

No obstante, analizando el abanico de notas que se abre en un grupo y en otro, podemos observar que el grupo en donde las evidencias de evaluación eran conocidas (grupo experimental) las notas oscilaban entre el 8 y el 9,5. No ocurre lo mismo con el grupo sometido a una evaluación de tipo ordinaria en donde el profesor ha sido el encargado de llevar a cabo la evaluación (grupo de control) en este caso las notas oscilan entre el 5 y el 9,5. Por tanto, sí podemos realizar una primera valoración sobre el hecho de que el grupo en el que las partes de la evaluación, como sus indicadores eran conocidos, las notas son más altas que en el grupo en donde esto no ocurre.

En el caso de la materia “Hacia una escuela inclusiva: modelos y prácticas” ocurre que con el Grupo de Control las notas oscilan entre el 5 y el 9,5; siendo la media obtenida 7.5. En el Grupo Experimental, las notas van desde 6.5 a 9 y la media coincide con la obtenida en el Grupo de Control.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 5,0	1	1,9	1,9	1,9
6,0	1	1,9	1,9	3,8
6,5	2	3,8	3,8	7,5
7,0	7	13,2	13,2	20,8
7,5	17	32,1	32,1	52,8
8,0	12	22,6	22,6	75,5
8,5	4	7,5	7,5	83,0
9,0	4	7,5	7,5	90,6
9,5	5	9,4	9,4	100,0
Total	53	100,0	100,0	

Tabla 4. Notas Obtenidas en “Hacia una escuela inclusiva: modelos y prácticas”. Grupo de Control

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	6,5	1	1,9	1,9	1,9
	7,0	3	5,7	5,7	7,5
	7,5	26	49,1	49,1	56,6
	8,0	8	15,1	15,1	71,7
	8,5	14	26,4	26,4	98,1
	9,0	1	1,9	1,9	100,0
Total		53	100,0	100,0	

Tabla 5. Notas Obtenidas en "Hacia una escuela inclusiva: modelos y prácticas". Grupo Experimental.

En el caso de esta materia, ocurre lo mismo pero la media de la nota otorgada por el profesor tanto en grupo de control como en grupo experimental es la misma y, habiendo diferencia entre las notas más bajas y más altas que se obtienen en cada uno de estos grupos, tal diferencia no es significativa. Incluso, en este caso, la nota más alta es otorgada en el grupo de control, y no en el experimental.

Una cuestión importante en este hecho, sería poder analizar las características de este alumnado, si alguno conocía la materia de antes, la había cursado ya, etc.

Como se ha citado en el apartado sobre metodología, el alumnado fue agrupado en 7 grupos de trabajo diferentes, cada grupo evaluó, mediante e-rúbrica y por el procedimiento de evaluación de pares, diferentes aspectos de la exposición de proyectos de los componentes de los demás grupos.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	40	16,8	16,8	16,8
	2	38	16,0	16,0	32,8
	3	33	13,9	13,9	46,6
	4	34	14,3	14,3	60,9
	5	31	13,0	13,0	73,9
	6	34	14,3	14,3	88,2
	7	28	11,8	11,8	100,0
Total		238	100,0	100,0	

Tabla 6. Equipo al que se evalúa

En la tabla anterior, se muestra el número de evaluaciones que ha obtenido cada grupo por parte de sus compañeros, de tal manera que, el grupo 1 obtiene 40 evaluaciones, el grupo 2, 38...

Los alumnos evalúan en esta e-rúbrica mediante una escala de 5 valores diferentes que van desde "No se muestra" a "excelente"

1	No se muestra
2	Aceptable
3	Mejorable
4	Considerable
5	Excelente

Tabla 7. Escala de valores para la evaluación

En la sección de la e-rúbrica relativa a la competencia sobre **Adquirir las habilidades necesarias para una adecuada gestión de la información con fines educativos...** se puede desglosar los siguientes aspectos. Esta competencia está compuesta por 5 evidencias diferentes que se muestra en la tabla siguiente.

EVIDENCIA	RESULTADO	
	Valor	%
"La exposición es ramificada o interactiva"	Considerable	46.2
"En la exposición utiliza un único recurso como apoyo a la misma"	Considerable	29.8
"En la exposición utiliza diferentes recursos tecnológicos"	Considerable	27.8
"Utiliza otros recursos en la exposición no necesariamente tecnológicos"	Excelente	39.1
"Gestiona el tiempo de forma coordinada con los recursos tecnológicos"	Considerable	46.6

Tabla 8. Evidencias de la Competencia "Adquirir las habilidades necesarias para una adecuada gestión de la información con fines educativos..."

Con esta primera competencia, ocurre que las medias de cada una de las evidencias se encuentran entre el valor 3 y el 4.

La segunda competencia que analiza la e-rúbrica sobre Exposición de Proyectos es llamada "Expresarse oralmente y dominar el uso de diferentes técnicas de expresión".

Esta competencia está desglosada en 9 evidencias diferentes que vamos a ver a continuación.

EVIDENCIA	RESULTADO	
	Valor	%
“El discurso mantiene despierta la curva de atención en todo momento”	Considerable	41.2
“Expone la información del proyecto de forma organizada, estructurada y ordenada”	Considerable	44.5
“La información se presenta de forma clara y lógica pudiéndose seguir con facilidad”	Considerable	45.4
“El volumen de voz es alto, claro y demuestra seguridad”	Considerable	42.9
“Muestra dominio en su postura corporal, lenguaje gestual y movimiento en el espacio”	Considerable	45.8
“El lenguaje es académico y profesional”	Considerable	49.2
“Gestiona el tiempo según lo acordado”	Excelente	38.2
“Para el discurso oral se apoya de textos escritos en la presentación”	Considerable	39.9
“El discurso oral se realiza con total naturalidad sin apoyo escrito en papel”	Considerable	30.9

Tabla 9. Evidencias de la Competencia “Expresarse oralmente y dominar el uso de diferente técnicas de expresión”

“Abordar con eficacia situaciones de aprendizaje de lenguas en contextos multiculturales y multilingües” es la siguiente competencia que, a su vez, está compuesta por un total de tres evidencias.

- “Se utiliza en parte un idioma de la Comunidad Europea pudiéndose seguir con facilidad”. En esta evidencia el máximo valor es el 5 que se da un total de 72 veces, suponiendo un total del 30.3% de las evaluaciones obtenidas.
- “Se utiliza totalmente un idioma de la Comunidad Europea pudiéndose seguir con facilidad”. El valor más repetido es el 1 (no se muestra) suponiendo un total del 32.2% de las evaluaciones obtenidas.
- “En la exposición se utilizan abundantes recursos y ejemplificaciones desde internet en ese idioma”. Ante esta evidencia el valor más repetido es el 1 (No se muestra), suponiendo un 33% del total de las valoraciones obtenidas.

		Se utiliza en parte un idioma de la Comunidad Europea	Se utiliza totalmente un idioma de la Comunidad Europea	En la exposición se utilizan abundantes recursos y ejemplificaciones desde internet en ese idioma.
N	Valid	238	236	233
	Missing	0	2	5
	Median	4,00	3,00	3,00
	Minimum	1	1	1
	Maximum	5	5	5

Tabla 10. Valores en la Competencia “Abordar con eficacia situaciones de aprendizaje de lenguas en contextos multiculturales y multilingües”

La cuarta competencia que encontramos en la e-rubrica para la evaluación de proyectos es “Afrontar la producción de materiales y recursos TIC de forma creativa e innovadora ajustándose a las necesidades del proyecto”. Tal competencia está compuesta por un total de 4 evidencias diferentes.

		Sabe presentar las líneas básicas de la experiencia educativa	Desarrolla la estructura científica del contenido del proyecto educativo	Adquiere metodológica mente la intención y finalidad de la propuesta educativa	Traslada esta experiencia a contextos diferentes de interpretación y la aplica correctamente
N	Valid	238	87	87	86
	Missing	0	151	151	152
	Median	4,00	4,00	4,00	4,00
	Minimum	2	2	2	2
	Maximum	5	5	5	5

Tabla 11. Valores de la Competencia “Afrontar la producción de materiales y recursos TIC de forma creativa e innovadora ajustándose a las necesidades del proyecto”.

La quinta competencia, “Desarrollar las capacidades de colaboración, participación activa y trabajo en equipo”, está formada por un total de 12 evidencias diferentes. Estas evidencias han sido valoradas mediante la misma escala de valoración. De esta última competencia no se destacan datos significativos que aporten relevancia al objeto de investigación. A nivel de categoría, se puede observar que, el valor más positivo (5, excelente) aparece con una frecuencia de 357, mientras que el más negativo (1, no se muestra) con una frecuencia de 215.

Resultados

No perdiendo de vista los objetivos de investigación que nos atañe, en este primer análisis, podemos inferir dos conclusiones demostrables:

1. En las dos materias en las que se ha empleado la herramienta e-rúbrica para la evaluación, y teniendo en cuenta los datos de los grupos de control y los experimentales, se manifiesta que:

En los grupos de control (donde las evidencias de evaluación no eran conocidas) el intervalo de notas obtenidas es más amplio. Es decir, las notas obtenidas en el grupo de control en la materia "Teoría y Práctica en la Educación Infantil" van desde el 5 hasta el 9,5; mientras que en el grupo experimental de la misma materia las notas oscilan entre el 8 y el 9,5. En el caso de la materia "Hacia una escuela inclusiva: modelos y prácticas" ocurre que con el Grupo de Control las notas oscilan entre el 5 y el 9,5; siendo la media obtenida 7.5. En el Grupo Experimental, las notas van desde 6.5 a 9 y la media coincide con la obtenida en el Grupo de Control.

2. En la medida en que las "evidencias" son más claras para el alumnado y menos subjetivas, las valoraciones a las mismas suelen estar repartidas entre todos los valores (De 1 a 5). Cuando la evidencia es menos inteligible para el alumno, éste suele otorgarle valores intermedios.

Conclusiones

Podemos concluir diciendo:

1. El hecho de que en el grupo experimental, las notas fueran más altas que en el grupo de control, nos hace realizar la valoración de que la e-rúbrica es un poderoso instrumento de cohesión e inclusión social de los alumnos tanto en la formalización de tareas, como en el desarrollo de la misma. Conocer los aspectos que se valoran hace que el alumno estructure y focalice mejor el objeto de estudio.
2. Entendemos que esta herramienta se puede utilizar en diferentes contextos de enseñanza-aprendizaje, ya que es una poderosa técnica en el diseño de actuaciones de éxito y en la formalización, no solo nacional sino internacional, de experiencias que llevan inevitablemente a resultados académicos exitosos en el momento que su empleo y evaluación no solo están sujetas al propio maestro sino que se centra en el uso y buen manejo de estándares y competencias por parte de los propios alumnos.
3. Que hay una gran variabilidad en cuanto a los resultados obtenidos por los grupos experimentales (con mayor éxito global) con respecto a los de control, no solo en el área disciplinar sino también en las demás áreas transversales.

Futuras líneas de trabajo.

En la actualidad, y dado los resultados obtenidos hasta el momento, nos enfrentamos a centrarnos en analizar cada una de las evidencias que componen la e-rúbrica, así como su claridad y nivel de objetividad para el alumnado que va a someterse a este tipo de evaluación.

También, y a raíz de los resultados de que disponemos, se van a analizar los resultados obtenidos en el cuestionario de usabilidad y satisfacción pasado al alumnado sometido a este tipo de evaluación.

De la misma forma, tendremos que tener en cuenta para futuros trabajos, ya que es posible que se pueda derivar de este estudio, pero que habrá que someterlo a revisión y análisis, que cada vez que los alumnos manejan con más frecuencia esta herramienta la dispersión entre las notas puestas por el maestro y las dadas por los alumnos se aproximan considerablemente.

Al igual que lo dicho anteriormente, se ha de considerar alguna limitación que se ha encontrado en el trabajo como ha sido la de poder analizar las características de este alumnado previamente, si alguno conocía la materia de antes, la había cursado ya, etc.

Referencias Bibliográficas

- Airasian, P. (2001). "The classroom as an assessment.", *Classroom Assessment: Concepts and Applications*. MacGraw Hill, pp. 1-26
- Archbald, D.A. y Newman, F.M. (1988). *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School*. National Association of Secondary School Principals. University of Wisconsin-Madison.
- Andrade, H. and Du, Y. (2005). "Student perspectives on rubric referenced assessment". *Practical, Assessment, Research & Evaluation*. V. 10, nº 3, April.
- Andrade, H. (2008). "Self-Assessment Through Rubrics". *Educational Leadership*, 65 (4), pp. 60-63.
- Bevan, N. (1998). "What is usability" In In: *Human Aspects in Computing: Design and Use of Interactive Systems with Terminals*. Citesser.
- Blanco, A. (2009). *Desarrollo y evaluación de competencias en educación superior*. Madrid: Narcea.
- Brown, S. y Glasner A. (2003). *Evaluar en la Universidad*. Madrid: Narcea.
- Boekaerts, M., Prinrich, P. y Zeidner (eds.) (2000). *Handbook of self-regulation*. Nueva York: Academic Press.
- Bonson, M. Y Benito (2005). "Evaluación y Aprendizaje". En Benito & Cruz (2005). *Nuevas claves para la docencia universitaria en el Espacio Europeo de Educación Superior*. Madrid: Narcea.
- Bretones, A. (2008). "Participación del alumnado de Educación Superior en su evaluación". *La atención a la diversidad ha legitimado la desigualdad*. *Revista de Educación*, 347, pp. 181-202

- Cebrián, M., Raposo, M Y Accino, J. (2007). *Formative evaluation Tools within European Space of Higher Education (ESHE): e-Portafolio and e-rubric*. EUNIS Conference Grenoble (Francia) Junio, 2007.
- Cebrián, M.; Serrano, J. y Ruiz, m.. (2014). *Las eRúbricas en la evaluación cooperativa del aprendizaje en la Universidad Comunicar*, 43. DOI: 10.3916/C43-2014-15. DOI <http://dx.doi.org/10.3916/C43-2014-15>
- Conde, A. y Pozuelo, F. (2007). "Las plantillas de evaluación (rúbrica) como instrumento para la evaluación. Un estudio de caso en el marco de la reforma de la enseñanza universitaria en el EEES". *Investigación en la escuela*, 63, pp. 77-90.
- Díaz Barriga, F. y otros (2009). "Evaluación auténtica de competencias docentes: una experiencia de construcción de sistemas de rúbricas en un centro virtual". Proyecto PAPIMEPE 303207. [http:// Giddetunam.org](http://Giddetunam.org).
- Falchikov, N. y Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer & Teacher Marks. *Review of Educational Research*, 70, 3, 287-322.
- Flecha, R. (2012). "La atención a la diversidad ha legitimado la desigualdad. En *Cuadernos de Pedagogía*, nº 429, Diciembre 2012, pp. 20-27
- Gallego-Arrufat, M.J. y Raposo-Rivas, M.(2014). "Compromiso del estudiante y percepción del proceso evaluador basado en rúbricas". *Redu. Revista de docencia universitaria*. V.12 (1). pp. 197-215.
- Hafner, JC: and Hafner, P. M. (2003). "Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating". *International Journal of Science Education* 25(12), pp. 1509-1528.
- Keppell, M. Y otros (2006). "Peer learning and learning-oriented assessment in technology-enhanced environments". *Assessment & Evaluation in Higher Education*, 31 (4), 453-464.
- López-Pastor, V. (2009). *Evaluación formativa y compartida en educación superior*. Madrid: Narcea.
- Marcos, L. Tamez, R. y Lozano, A. (2009). "Aprendizaje móvil y desarrollo de habilidades en foros asíncronos de comunicación". *Comunicar*, nº 33, v. XVII, pág. 93-110.
- Panadero, E. Alonso-Tapia, J. y Reche, E. (2013). "Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers". *Studies in Educational Evaluation*, 39(3), 125-132. (DOI: <http://dx.doi.org/10.1016/j.stueduc.2013.04.001>).
- Prins, F.J., Sluijsmans, D.M.A., Kirschner, P.A. & Strijbos, J.W. (2005). "Formative Peer Assessment in a CSCL Environment: A Case Study, Assessment and Evaluation in Higher Education", 30, 417-444. (DOI: <http://hdl.handle.net/1820/916>).
- Simona Rychen, D. y Hersh Salganik, L. (2006). *Las competencias clave para el bienestar personal, social y económico*. Aljibe: Málaga.
- Stenhouse, L. (1984). *Investigación y desarrollo del currículo*. Morata: Madrid.
- Santos Guerra, M. A. (1999). *La evaluación: un proceso de diálogo, comprensión y mejora*. Aljibe: Málaga.

- Serrano, J. y Cebrian, D. (2012). "Usabilidad de la e-rúbrica mediante cuestionarios online con limesurvey". En C. Leite y M. Zabalza (coords.), *Ensino Superior. Inovação e igualdade na docencia* (467-486). CIE: Porto.
- Serrano, J. y Cebrian, D. (2014). "Usabilidad y satisfacción de la erúbrica. *Redu. Revista de docencia universitaria*. V. 12 (1). pp. 177-195.
- Stevens, D.D. Y Levi, A.J. (2005). *Introduction to Rubrics*. Canadá.
- Torres Gordillo, J.J. y Perea Rodríguez, V.H. (2010). "La rúbrica como instrumento pedagógico para la autorización y evaluación de los aprendizajes en el Foro online en Educación Superior". *Píxel-Bit. Revista de Medios y Educación*, 36. pp. 141-149.
- Weinert, B. (1985). *Manual de Psicología de la Organización*. Barcelona: Herder.

Fuentes Electrónicas

- Andreu Barrachina, LL.; Sanz Torrent, M. y Serrat Sellabona, E. (2009). Una propuesta de renovación metodológica en el marco del Espacio Europeo de Enseñanza Superior: los pequeños grupos de investigación cooperativos. *REIFOP*, 12 (3), 111-126. [Consultado 2/08/2012] <http://www.aufop.com>
- Díaz y otros (2009). Evaluación auténtica de competencias docentes: Una experiencia de construcción de sistemas de rúbricas en un entorno virtual, *Proyecto PAPIME-303207*. [Consultado 29/08/2012] <http://www.giddetunam.org>
- Gómez Patiño, R.G.; Tablas López, V. y López López, M.A. (2009). La evaluación con rúbricas en el desempeño por competencias profesionales. *X Congreso nacional de Investigación Educativa*. [Consultado 16/08/2012] <http://www.comie.org.mx/swgc/vz/sitio/sitio.php>
- Montanero, M.; Mateos, V.; Gómez, V. Y Alejo, R. (2006). Orientaciones para la elaboración del Plan Docente de una asignatura. [Consultado 21/09/2012] <http://www.unex.es/unex/organización>

Pérez Galán, R.; Cebrián Robles, D.; Rueda Galiano, A.B. (2014). Evaluación de pares y autoevaluación con erúbricas: caso de estudio en el grado de Educación Primaria. *REDU - Revista de Docencia Universitaria*, 12 (4), 337-456.

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

Rafael Pérez Galán

Universidad de Málaga
Departamento de Teoría e Historia de la Educación

Mail: rpg@uma.es

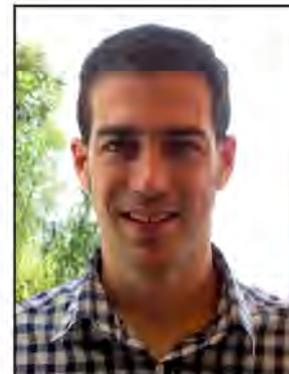


Doctor en Ciencias de la Educación y diplomado en Profesor de EGB. En la actualidad ejerce como profesor asociado en el departamento de Teoría e Historia de la Educación de la Universidad de Málaga y en el Centro de Magisterio adscrito a la Universidad de Málaga, María Inmaculada, de Antequera. Coordinador del practicum de grado en los nuevos planes de estudio de formación de maestros y miembro del equipo directivo. Sus temas de interés y líneas de investigación son la escuela comprensiva y las nuevas tecnologías y las herramientas digitales como recursos de inclusión en los contextos de formación del profesorado.

Daniel Cebrián Robles

Universidad de Málaga
Departamento de Métodos e Innovación

Mail: dcebrian@uma.es



Doctor por la Universidad de Málaga. Ingeniero Industrial. 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”. 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 [<https://gteavirtual.org/rubric>] y desarrollador de Multimedia Annotations [<http://openvideoannotation.org>] para Harvard University y su integración en el MOOC de edX.

Ana B. Rueda Galiano

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

Mail: anarueda@uma.es



Profesora asociada al departamento de Métodos de Investigación e Innovación Educativa hasta 2012, en la universidad de Málaga. Licenciada en Psicopedagogía (Junio 2007). Diploma de Estudios Avanzados en Investigación e Innovación Educativa, obtenido el 4 de Diciembre de 2009 en la Facultad de Ciencias de la Educación de la Universidad de Málaga, como culminación a los cursos de doctorado del Bienio 2007/2009 en Investigación e Innovación Educativa, ofertado por el Departamento de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Ciencias de la Educación de la Universidad de Málaga. Diplomada en Magisterio. Especialidad en Educación Musical. Máster Universitario en Atención Temprana “Dr Miguel de Linares Pezzi” (Junio 2009).

Artículo 4: Usability and Satisfaction in Multimedia Annotation Tools for MOOCs



Usability and Satisfaction in Multimedia Annotation Tools for MOOCs

Usabilidad y satisfacción en herramientas de anotaciones multimedia para los MOOC

- ✉ Dr. Juan-José Monedero-Moya is Associate Professor in the Faculty of Educational Sciences at the University of Malaga (Spain) (jmonedero@uma.es).
- ✉ Daniel Cebrián-Robles is Industrial Engineer from the Higher Technical School of Industrial Engineering, University of Malaga (Spain) (dcebrian@uma.es).
- ✉ Dr. Philip Desenne is Consultant in the HarvardX Instructional Technology for Annotation at Harvard University, Boston (USA) (desenne@fas.harvard.edu).

ABSTRACT

The worldwide boom in digital video may be one of the reasons behind the exponential growth of MOOCs. The evaluation of a MOOC requires a great degree of multimedia and collaborative interaction. Given that videos are one of the main elements in these courses, it would be interesting to work on innovations that would allow users to interact with multimedia and collaborative activities within the videos. This paper is part of a collaboration project whose main objective is «to design and develop multimedia annotation tools to improve user interaction with contents». This paper will discuss the assessment of two tools: Collaborative Annotation Tool (CaTool) and Open Video Annotation (OVA). The latter was developed by the aforementioned project and integrated into the edX MOOC. The project spanned two academic years (2012-2014) and the assessment tools were tested on different groups in the Faculty of Education, with responses from a total of 180 students. Data obtained from both tools were compared by using average contrasts. Results showed significant differences in favour of the second tool (OVA). The project concludes with a useful video annotation tool, whose design was approved by users, and which is also a quick and user-friendly instrument to evaluate any software or MOOC. A comprehensive review of video annotation tools was also carried out at the end of the project.

RESUMEN

El auge del vídeo digital a nivel mundial puede ser una de las causas del crecimiento exponencial de los MOOC. Las evaluaciones de los MOOC recomiendan una mayor interacción multimedia y colaborativa. Siendo los vídeos unos de los elementos destacados en estos cursos, será interesante trabajar en innovaciones que permitan una mayor capacidad a los usuarios para interactuar con anotaciones multimedia y colaborativas dentro de los vídeos. El presente artículo es parte del proyecto de colaboración, cuyo objetivo principal fue «El diseño y creación de herramientas de anotaciones multimedia para mejorar la interactividad de los usuarios con los contenidos». En este artículo mostraremos la evaluación de dos herramientas como fueron Collaborative Annotation Tool (CaTool) y Open Video Annotation (OVA) esta última desarrollada por el proyecto e integrada en el MOOC de edX. El proyecto abarcó dos cursos académicos (2012-14) y se aplicó un instrumento de evaluación en diferentes grupos de la Facultad de Educación a un total de 180 estudiantes. Se compararon los datos obtenidos entre ambas herramientas con contrastes de media, resultando diferencias significativas a favor de la segunda herramienta. Al concluir el proyecto se dispone de una herramienta de anotaciones de vídeo con diseño validado por los usuarios; además de un instrumento sencillo y rápido de aplicar para evaluar cualquier software y MOOC. Se realizó también una revisión amplia sobre herramientas de anotaciones de vídeos.

KEYWORDS | PALABRAS CLAVE

Usability, satisfaction, design tools, evaluation software, multimedia annotations, educational software, MOOC, university education. Usabilidad, satisfacción, diseño de herramientas, evaluación de software, anotaciones multimedia, software educativo, MOOC, enseñanza universitaria.



1. Introduction

The development of digital video has allowed users greater accessibility; it has made its way into our homes and lives, turning consumer and retail services such as YouTube into a sociological phenomenon. YouTube viewings currently account for an average of 6 million video hours per month¹. Clearly much has changed since the Lumière brothers invented cinema (Díaz-Arias, 2009: 64). This development has provided the gateway for developing technologies that allow users to share and collaborate (Computer Supported Collaborative Learning: CSCL). Such technologies also include collaborative video annotation technologies (Yang, Zhang, Su & Tsai, 2011), which have led to the emergence of innovative social projects where video annotation tools are collectively used (Angehrn, Luccini & Maxwell, 2009). The digitization of videos (Bartolomé, 2003) opened up new interactive possibilities in education, along with hypermedia (García-Valcarcel, 2008), and has represented a breakthrough for learning and teaching by leaving behind the passive reading of videos (Colasante, 2011). There is a long history of experimental studies on how to apply videos in education (Ferrés, 1992; Cebrián, 1994; Bartolomé, 1997; Cabero, 2004; Area Moreira, 2005; Aguaded and Sánchez, 2008; Salinas, 2013). In the field of teacher training, there are examples related to the concept of microteaching, which has been questioned due to its reductionist approach to teacher initial training. Nevertheless, it was such an effort to come up with a rather rigorous idea of teaching. Leaving aside the theoretical starting point of this paper, there are some recent studies and developments of video annotation tools that, supported by other conceptions of teaching (Schön, 1998; Giroux, 2001), have shown efficacy in meta-evaluations for initial training (Hattie, 2009). The application contexts of the above studies are many and varied, and address processes such as reflection, shared evaluation and collective analysis of classroom situations. Therefore, they have proven to be effective tools for teachers and teacher trainees to collectively analyse everyday teacher practice (Rich & Hannafin, 2009a; Hosack, 2010; Rich & Trip, 2011; Picci, Calvani & Bonaiuti, 2012; Etscheidt & Curran, 2012; Ingram, 2014). In relation to initial training and the development of reflective skills, Orland-Barak & Rachamim (2009) carried out an interesting review and study by comparing different models of reflection using videos as a support. Rich and Hannafin (2009b) conducted another significant review of technological solutions and the potential of video annotation tools for teaching. They conducted a comparative analysis of

these tools based on the following criteria: how to use, note style, collaboration, safety, online-offline, format, resource import vs. export, learning curve and cost (free/hiring research teams). We then found an even more extensive review (Rich & Trip, 2011), shown in table 1, which was completed by solutions, presented in the last international workshop on multimedia notes 'iAnnotate14'².

2. Integrating collaborative annotation tools in MOOC

Video and other related emerging technologies (analysis of big data, ontologies, semantic web, geo-location, multimedia notes, rubric-based assessment, federation technologies, etc.) quickly gained prominence in MOOCs, shaping the core structure of these courses. The appealing and widespread use of videos may have played a role in the boom of MOOCs, prompting a search for new interactive ways to read videos and general contents. It was only recently that MOOCs have incorporated previous experiences and developments on the features of collaborative multimedia annotations; allowing for a more interactive, multimedia learning process, and sharing users' views on these platforms. This has also provided the gateway for a new model of learning community within the MOOC, which can manage a significant flow of meanings extracted from reading contents and from annotations in different codes, namely: video, text, image and sound notes, as well as hyperlinks and eRubrics (Cebrián-de-la-Serna & Bergman, 2014; Cebrián-de-la-Serna & Monedero Moya, 2014).

These notes can be made in different formats and codes showing contents, such as: annotations in videos, texts, images, maps, charts, etc. as well as annotations created by users. The above possibilities open up a whole new line of new technological developments and research on the dynamic narrative of messages, given the speed with which MOOC platforms and courses are being implemented worldwide. Therefore, we need to innovate in the design and content of video tools based on their new interactive possibilities, in order not to replicate mistakes from the past, when, in the early stage of a new technology, the narrative models of preceding technologies would be incorporated without exploring the interactive potential of the new formats. Something similar happened during the transition from radio messages to television messages, as pointed out by Guo, Kim & Rubin (2014), who conducted a study on the video sessions of four edX courses. They checked the different formats used and concluded that recording cannot be extrapolated to

MOOC, because students do not pay enough attention. As a consequence, they suggested a list of recommendations that can be summarized as follows: more interactive and easy-to-edit videos, shorter (6 minutes), and easy-to-share notes. The development of educational software and the possibilities offered by free software have generated a community of developers who share their experience. The fact that these products get feedback from users also constitutes a model of software production; as communities of practice emerge around tools, services and specific platforms such as GitHub³.

The symbiotic relationship between developers and communities of practice has allowed MOOCs to evolve from structured approaches (xMOOCs) to communicative and collaborative approaches (cMOOCs) in their platforms and courses. However, both approaches require new interactive features in the videos. An

example of such features is the project here presented, which has been led by the HarvardX team for integration into the edX MOOC, and whose objectives are as follows: on the one hand, designing high-capacity multimedia annotation tools to create multimedia meaning and sharing it with users; and on the other, competence assessment, self-assessment and peer assessment through eRubrics. In order to quickly introduce these changes of great impact, we must count on assessment strategies for end-users to evaluate tools while they are being developed. Tools must be quick and easy to use, in order to collect data that will guide production (technical and content production), even before the beta version emerges. This is why our GTEA group carries out a design, test and evaluation line for educational software, which aims to find a balance between educational innovation and technological innovation, i.e. between generating new envi-

Table 1. A Comparison of Annotation Tools, by Rich and Hannafin (2009b); Rich & Trip, 2011 and iAnnote14

Tool Name	How to Use	Note Style	Collaboration	Cost	Website
VAST	Independent application	Users select a point in the video and text areas are offered	Not possible	Free	http://www.professional-vision.org
VITAL	Based on website	Users create video clips and reference videos as hyperlinks in a written document	Not possible	Contract / Free	http://vital.ccnmtl.columbia.edu
VideoTraces	Independent application	Users select an excerpt in the video and narrate audio comments	Different users can note down, respond and create a discussion thread	Contract	http://depts.washington.edu/pett/projects/videotraces.html
Video Paper	Independent application; it can be exported to web	Users select an excerpt in the video and associate text to it, by using subtitles with a timed transcript	Not possible	Free	http://vpb.concord.org/
MediaNotes	Independent application	Users select and mark titles and comments on the video and associate default clips to it	Different users can make notes on the same video	\$\$	http://www.bluemangolearning.com/products/medianotes
Studiocode	Independent application	Users create and apply a set of codes to a video sequence	Different users can share and compare notes in a list	\$\$\$	http://www.studiocodegroup.com
Iris	Based on a website	Text, live coding Includes statistical analysis tools	Different users can make notes on the same video	\$\$\$	http://www.therenow.net
Video Ant	Based on a website	Text notes. Tagging not possible	Not possible	Free	http://ant.umn.edu
Viddler	Based on a website	No text nor voice in the video excerpt	Different users can make notes on the same video	Free	http://www.viddler.com
Factlink	Based on a website	Plugin to edit text in any visible page in Chrome	Accessed and shared via Twitter and Facebook	Free	https://factlink.com/in-your-browser
Remark	Based on a website	Video annotations on a «frame»	Different users can edit and share notes	\$\$	https://remarkhq.com
OVA	Web-based and integrated into platforms like MOOC	Multimedia notes on video sequences, with text editors, rubrics, etc.	Different users can make notes on one or more files	Free	http://openvideoannotation.org

ronments and users' usability and satisfaction. The ultimate aim is for new interactive methodologies such as multimedia annotation tools for MOOCs, to be validated by end-users. To do so, we need to create a parallel line of research and evaluation instruments that are reliable and valid for decision-taking when designing educational software. We must take into account all possible elements for software evaluation from the users' perspective (satisfaction, usability, cost, portability, productivity, accessibility, safety, etc.), in order to examine their ease of use (aka usability), regardless of their context, personal differences, different supports (tablets, mobile phones, computers, etc.).

This paper uses the following definition of usability: 'the extent to which a product can be used by certain users to achieve specific goals with effectiveness, efficiency and satisfaction in a particular context of use' (Bevan, 1997). Satisfaction is often seen as a construct within usability studies and instruments, although we believe it is rather the opposite. The ease of use of a tool or service is an element that belongs to the overall user satisfaction. The satisfaction of technological tools and services can even be considered as a sub-category within user satisfaction studies, as shown by studies on students' satisfaction of university life (Blázquez, Chamizo, Cano & Gutiérrez, 2013). This is a live debate, given the massive presence of technological services and resources, and the digitisation that most communication, teaching, research and administration processes have recently gone through within universities. Both usability and user satisfaction are measured by questionnaires completed by users. We can find usability questionnaires in websites and systems (Bangor, Kortum & Miller, 2008; 2009; Kirakowski & Corbett, 1988; Molich, Ede, Kaasgaard & Karyukin, 2004; Sauro, 2011), satisfaction questionnaires, and questionnaires on both usability and satisfaction (Bargas-Avila, Lötscher, Orsini & Opwis, 2009; McNamara & Kirakowski, 2011).

3. Methodology

The present project started from the mutual interest shared by our team and HarvardX Annotation Management in creating tools to facilitate meaning processes based on collective multimedia annotations. The general aim of the project was to create a new tool for multimedia annotations specifically designed to



Graph 1. eRúbrica tool integrated into CaTool annotations.

respond to the new features of technological progress (e.g. semantic web, annotation ontology, etc.), as well as to the social practices that are currently being developed by users on the Internet (learning in communities of practice, using mobile devices, collaborative work, communication in social networks, creating eRubrics, etc.). The tool is currently integrated into the edX MOOC, and has been in use since January 2014 in the courses offered by HarvardX⁴. The technological development started from scratch, although it was based on the progress that had been made in the field of multimedia annotations on the Open Annotation Community Group, and taking into account the aforementioned literature as well as other developments by Harvard University. The results presented here are part of a collaborative project and show users' opinions on the usability and user satisfaction in relation to an instrument designed to assess web tools. Such data is often required to design and improve tools. This is why the methodology used in this paper contrasted end-users' usability and satisfaction in the Collaborative Annotation Tool (CaTool) (created by Harvard University, 2012), against the added features of the new tool created by the Open Video Annotation project (OVA).

For methodological purposes, the new added features of video annotation were considered as the independent variable. The development had a dual purpose: to serve as a collective multimedia annotation service, and to integrate the new features into the edX MOOC. The present paper will only show the results of assessing the video annotation features that had been added to the edX MOOC. However, this platform hosted the full-featured OVA video annotation.

text, sound and quality image (the last two in experimental stages).

The study was divided into two parts: a) The first stage during the 2012-13 academic year, where the Collaborative Annotation Tool (CaTool) was trialled on groups of different subjects in the Faculty of Educational Sciences at the University of Malaga (Spain). The usability and user satisfaction instrument that we had already created for other tools was also tested during this stage. b) The second stage during the 2013-14 academic year, where the usability and user satisfaction instrument designed during the first stage was improved and applied to two groups from the Degree of Education that shared the same teacher, methods and tasks; we compared two different annotation tools: CaTool and a beta tool that only had the OVA video annotation feature. In the first stage (2012-13) the Collaborative Video Annotation tool was tested in the class within the Education department and on different types of subjects within the degree programme (core subjects, elective subjects, internships, etc.). The tool was federated by our team, and its combination with other tools, such as eRubric and federation technology, had provided interesting features in practice (see Image no.1). The state of the art in relation to the design, creation and assessment of previous video annotation tools was also collected at this stage.

At the second stage, during the second half of 2013, a new Open Video Annotation (OVA)⁵ was created (image 2), which responded to an interactive and communicative teaching model in the MOOC. The creation and design of this tool was guided by the HarvardX annotation manager, and included the following features: a) Editing entries could be done in a multimedia format (video, text, image, etc.). b) Multimedia annotations could be added within the resource

itself (in the video, image, etc.). c) Annotations could be shared and discussed by a large number of users, so that when someone received a message with a note on it, a simple click would take them to that particular note within the resource. d) Editing tags in a database of ontological annotations was possible. As an option, each entry also had the possibility of geolocating. f) Annotations could easily be shared on social networks. eRubrics could be created when editing annotations.

During the 2013-14 academic year, CaTool and OVA were tested. The test involved the same teacher, methodology, class lab and all the student groups (180 in total) of the mandatory second year technological resources course within the degree of Education in the Faculty of Educational Sciences at the University of Malaga. After this, the enhanced instrument of usability and user satisfaction from stage 1 was used. The first experiment was performed on the CaTool, and the second on a beta tool (a month later); but only on the OVA video annotation feature, and with some limitations (it could only be used with the Chrome browser).

4. Analysis and results

The participant sample consisted of all the students from the aforementioned mandatory course in the Faculty of Educational Sciences who got to work with these tools for the first time. Once they performed the task set by the teacher, they were asked to answer a questionnaire on usability and user satisfaction. The questionnaire consisted of a series of descriptive questions (age, gender, user level, etc.), followed by 26 sentences to be rated on a Likert rating scale of 1 to 5. There were direct sentences (1 = the worst; 5 = the best) as well as indirect sentences (1 = the best; 5 = the worst). As for usability, there were 17 sentences: 5 direct and 12 indirect. For user satisfaction there were 9: 7 direct and 2 indirect. The order of the sentences in the questionnaire was random, in order to avoid answering without reading. There was an open question at the end, for students to write free comments. The average response time was 4 minutes. The questionnaire was filled out online by using LimeSurvey, while data was analyzed by using the SPSS (version 20). For analysis purposes, we ensured answers had to be thought through, and sentences could not be rated by simply filling out the questionnaire. To this end, we detec-



Graph 2: Multimedia annotation tool.

ted 16 answers that marked similar values in blocks corresponding to direct and indirect sentences, so they were therefore considered as non-valid answers. We carried out the $y=6-x$ transformation in the values of indirect sentences, so that calculations could not be counteracted.

Significant differences were found in favour of OVA among the means of the questionnaire. When analysing the questionnaire by blocks, significant differences were also found in the usability blocks, but not in the user satisfaction blocks (table 2).

The contrast of the usability and satisfaction instrument between the two tools throws up significant differences in favour of OVA in the following items: 'I found the application to be pleasant', 'I found the application exhausting to use', 'The application does not need explaining to be used', 'I needed help to access the application', 'I ran into technical problems', 'It requires expert help', 'The response time in the interaction is slow'.

Graph 1 shows the histograms of the total scores for each tool. It shows that, from the 105 score onwards, there are more ratings for OVA than for CaTool, while the opposite goes for scores under 105. According to their comments, respondents support the questionnaire results: they consider these tools to be easy, useful and innovative. The negative aspects were mainly attributed to technical issues: Internet access, slow server or browser limitations in the beta version.

5. Discussion and conclusions

The potential of the video digitalizing process has been foreseen for a long time, along with new teaching processes at universities (Aguaded & Macías,

	N	Mean	Standard Deviation	Standard Error	Confidence interval for the mean at 95%	
					Lower Limit	Upper Limit
CaTool	80	94.5375	13.08680	1.46315	91.6252	97.4498
OVA	84	99.9643	14.54900	1.58743	96.8070	103.1216
Total	164	97.3171	14.07812	1.08932	95.1463	99.4878

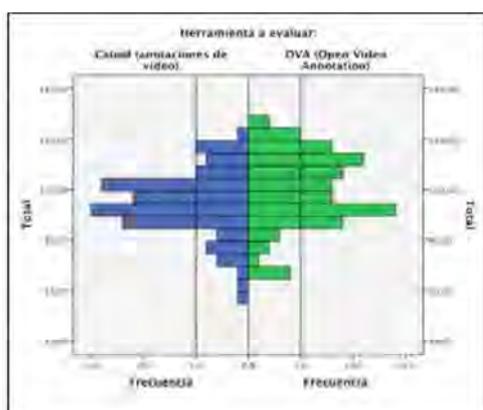
2008: 687), except that nowadays we look forward to even further possibilities that go beyond past predictions. Socialization and distribution of information, free access to premium content, networks and learning communities to share and generate new ways of learning, the technological development of the Internet

	Levene's test for equal variances		T-test for equality of means			
	F	Sig.	t	df	Sig. (bilateral)	Mean Difference
Total Scores	2.510	.115	-2.507	162	.013	-5.42679
Total Usability (direct)	.008	.927	-2.072	162	.040	-.99107
Total Usability (indirect)	.150	.699	-3.249	162	.001	-3.90000
Total Satisfaction (direct)	.518	.473	-.500	162	.518	-.42262
Total Satisfaction (indirect)	.000	.994	-.408	162	.684	-.11310

(augmented reality, mobile technology, wearable, network capacity, etc.) are forcing universities to respond to new challenges.

MOOC platforms are not immune to these changes, and will soon incorporate experiences and developments in the area of collective multimedia annotations. Innovations find in these massive platforms an ideal setting for developing, testing and experimenting with educational research. Certainly, we consider this new environment as an ideal setting for conducting new experiments, studies and educational projects such as the one put forward here. The present project has shown that collective multimedia annotations are generally highly-rated by students when they are easy to use (as observed in the aforementioned mean differences), and when displaying certain features that are fashionable amongst the young. For instance, features related to mobility, social networks, collective interaction and broadcast of shared meanings, as could be observed in the best rated features and in the open essay answers when the two tools were compared. These features were added to the new Open Video Annotation (OVA) tool, which aims to be in line with university students' symbolic and communicative competence. Students should be therefore more critical and prepared for what Castell (2012: 23-24) defines as mass self-communication. He considers this to be vital in symbolic construction, as it mainly depends on «the created frameworks, i.e. the fact that the transformation of the communication environment directly affects the way in which meaning is constructed».

We believe that collective multimedia annotation has many educational possibilities in university teaching. Some of these possibilities go beyond the existing format, reaching the aforementioned 'created framework' nowadays represented by MOOCs. Their application and research can be interesting in further educational settings beyond those studied in this project, such as: a) Blended learning models currently developed at universities, which use materials and resources to sup-



Graph 3: Histograms of total scores on the two tools.

port teaching; b) Learning objects with multimedia annotations and semantic web (García-Barricanal, Sicilia, Sánchez-Alonso & Lytras, 2011); c) Supervision during the Practicum (Miller & Carney, 2009) with ePortfolios (electronic portfolios), filled with multimedia proof of learning and where the meanings given to annotations can be shared. d. Dissemination of scientific knowledge, as suggested by Vázquez-Cano (2013: 90), by combining the written format with the video-article and the scientific pill. Such combination would provide scientific production with more visibility, broadcast and flow of exchange. All the above contexts and experiences are innovative and consistent with the practice that we wish to widely implement in universities, thus representing a strong leadership in the knowledge society.

Support and notes

¹ The collaborative project was entitled Open Video Annotation Project (2012-2014) (<http://goo.gl/51W37d>) and was made possible through the joint funding of institutions such as: Talenta scholarships and Grea Group (<http://grea.uma.es>) PAI SEJ-462 Andalusian Regional Government, University of Malaga and Center for Hellenic Studies -CHS- (Harvard University) (<http://chs.harvard.edu>) (09-07-2014).

² YouTube Statistics (<http://goo.gl/AYrCL>) (09-07-2014).

³ International Workshop on Multimedia Annotations 'Annotate14', San Francisco, California (USA), April 3-6, 2014 <http://fannote.org> (09-07-2014).

⁴ Open Source Platform <http://github.com>.

⁵ The first course using OVA was 'Poetry in America: Whitman', in edX Harvard University <http://goo.gl/19bupN> (09-07-2014).

⁶ OVA Tool (<http://openvideoannotation.org>) (09-07-2014).

References

- AGUADÉ, J. & MACÍAS, Y. (2008). Televisión universitaria y servicio público. *Comunicar*, 31(XVI), 681-689. (<https://doi.org/od4fkw>).
- AGUADÉ, I. & SÁNCHEZ, J. (2008). Niños adolescentes tras el visor de la cámara: experiencias de alfabetización audiovisual. *Estudios sobre el Mensaje Periodístico*, 14, 293-308.
- ANGEHRN, A., LUCCINI, A. & MAXWELL, K. (2009). InnoTube: A Video-based Connection Tool Supporting Collaborative Innovation. *Interactive Learning Environments*, 17, 3, 205-220. (<http://doi.org/bw48vv>).
- AREA, M. (2005). Los criterios de calidad en el diseño y desarrollo de materiales didácticos para la www. *Comunicación y Pedagogía*, 204, 66-72.
- BANGOR, A., KORTUM, P.T. & MILLER, J.T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574-594.
- BANGOR, A., KORTUM, P.T. & MILLER, J.T. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3), 114-123.
- BARGAS, J.A., LÖTSCHER, J., ORSINI, S. & OPWIS, K. (2009). Intranet Satisfaction Questionnaire: Development and Validation of a Questionnaire to Measure User Satisfaction with the Intranet. *Computers in Human Behavior*, 25, 1241-1250. (<http://doi.org/b39md8>).
- BARTOLOMÉ, A. (1997). Uso interactivo del video. In J. FERRÉS & P. MARQUÉS (Coord.), *Comunicación educativa y nuevas tecnologías*. Barcelona: Praxis, 320 (1-13).
- BARTOLOMÉ, A. (2003). Video digital. *Comunicar*, 21, 39-47. (<http://goo.gl/MDeYQd>) (29-04-2014).
- BEVAN, N. (1997). Quality and Usability: A New Framework. In VAN-VEENENDAAL, E. & McMULLAN, J. (Eds.), *Achieving Software Product Quality*. Netherlands: Turen Nollothenius, 25-34.
- BLÁZQUEZ, J.J., CHAMIZO, J., CÁNO, E. & GUTIÉRREZ, S. (2013). Calidad de vida universitaria: Identificación de los principales indicadores de satisfacción estudiantil. *Revista de Educación*, 362, 458-484. (<http://doi.org/tp5>).
- CABERO, J. (2004). El diseño de videos didácticos. In J. SALINAS, J., CABERO & I. AGUADÉ (Coords.), *Tecnologías para la educación: diseño, producción y evaluación de medios para la formación docente* (pp. 141-156). Madrid: Alianza.
- CASTELLIS, M. (2012). *Redes de indignación y esperanza*. Madrid: Alianza.
- CEBRÍAN-DE-LA-SERNA, M. (1994). Los videos didácticos: claves para su producción y evaluación. *Pixel-Bit*, 1, 31-42. (<http://goo.gl/v3Ay6>).
- CEBRÍAN-DE-LA-SERNA, M. & BERGMAN, M. (2014). Formative Assessment with eRubrics: an Approach to the State of the Art. *Revista de Docencia Universitaria*, 12, 1, 23-29. (<http://goo.gl/A4cpaa>).
- CEBRÍAN-DE-LA-SERNA, M. & MONEDERO, J.J. (2011). Evolución en el diseño y funcionalidad de las rúbricas: desde las rúbricas «cuadradas» a las rúbricas Federadas. *Revista de Docencia Universitaria*, 12, 1, 81-98. (<http://goo.gl/xNhnqR>).
- COLASANTE, M. (2011). Using Video Annotation to Reflect on and Evaluate Physical Education Pre-service Teaching Practice. *Australian Journal of Educational Technology*, 27(1), 66-88. (<http://goo.gl/f2HfZB>).
- DÍAZ-ARIAS, R. (2009). El video en el ciberespacio: usos y lenguaje. *Comunicar*, 33, 17, 63-71. (<http://doi.org/ft5qr>).
- ETSCHIEDT, S. & CURRAN, C. (2012). Promoting Reflection in Teacher Preparation Programs: A Multilevel Model. *Teacher Education and Special Education* 35(1) 7-26. (<http://doi.org/dk53x2>).
- FERRÉS, J. (1992). Video y educación. Barcelona: Paidós.
- GARCÍA-BARRICANAL, E., SICILIA, M.A., SÁNCHEZ-ALONSO, S. & LYTRAS, M. (2009). Semantic Annotation of Video Fragments as Learning Objects: A Case Study with YouTube Videos and the Gene Ontology. *Interactive Learning Environments*, 19, 1, 25-44. (<http://doi.org/b2pkpf>).
- GARCÍA-VALCARCEL, A. (2008). El hipervideo y su potencialidad

- pedagógica. *Revista Latinoamericana de Tecnología Educativa (RELATEC)*, 7, 2, 69-79.
- GIROUX, H.A. (2001). *Cultura, política y práctica educativa*. Barcelona: Graó.
- GUO, P., KIM, H. & RUBIN, R. (2014). How Video Production Affects Student Engagement: An Empirical Study of MOOCs Videos. *Proceedings of the First ACM Conference on Learning @ scale Conference* (pp. 41-50). March 4-5, Atlanta, Georgia, USA. (<http://doi.org/tp6>).
- HATTIE, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York: Routledge.
- HOSACK, B. (2010). VideoANT: Extending Online Video Annotation Beyond Content Delivery. *TechTrends*, 54, 3, 45-49.
- INGRAM, J. (2014). Supporting Student Teachers in Developing and Applying Professional Knowledge with Videoed Events. *European Journal of Teacher Education*, 37(1), 51-62. (<http://doi.org/tp7>).
- KIRAKOWSKI, J. & CORBETT, M. (1988). Measuring User Satisfaction. *4th Conference of the British Computer Society Human-Computer Interaction Specialist Group*, 329-338.
- MCNAMARA, N. & KIRAKOWSKI, J. (2011). Measuring User-satisfaction with Electronic Consumer Products: The Consumer Products Questionnaire. *International Journal Human-Computer Studies*, 69, 375-386. (<http://doi.org/d5xzqn>).
- MILLER, M. & CARNEY, J. (2009). Lost in Translation: Using Video Annotation Software to Examine How a Clinical Supervisor Interprets and Applies a State-mandated Teacher Assessment Instrument. *The Teacher Educator*, 44(4), 217-231. (<http://doi.org/dhj2bv>).
- MOLICH, R., EDE, M.R., KAASGAARD, K. & KARYUKIN, B. (2004). Comparative Usability Evaluation. *Behaviour & Information Technology*, 23(1), 65-74.
- ORLAND-BARAK, L. & RACHAMIM, M. (2009). Simultaneous Reflections by Video in a Second-order Action Research-mentoring Model: Lessons for the Mentor and the Mentee. *Reflective Practice*, 10, 5, 601-613. (<http://doi.org/db82mr>).
- PICCI, P., CALVANI, A. & BONAIUTI, G. (2012). The Use of Digital Video Annotation in Teacher Training: The Teachers' Perspective. *Procedia, Social and Behavioral Sciences*, 69, 600-613. (<http://doi.org/tp8>).
- RICH, P. & TRIP, T. (2011). Ten Essential Questions Educators Should Ask When Using Video Annotation Tools. *TechTrends*, 55, 6, 16-24.
- RICH, P. J., & HANNAFIN, M. (2009a). Scaffolded Video Self-analysis: Discrepancies between Preservice Teachers' Perceived and Actual Instructional Decisions. *Journal of Computing in Higher Education*, 21(2), 128-145.
- RICH, P. J. & HANNAFIN, M. (2009b). Video Annotation Tools. Technologies to Scaffold, Structure, and Transform Teacher Reflection. *Journal of Teacher Education*, 60, 1, 52-67. (<http://doi.org/dzdv4n>).
- SALINAS, J. (2013). Audio y vídeo Podcast para el aprendizaje de lenguas extranjeras en la formación docente inicial. *IV Jornadas Internacionales de Campus Virtuales*. 14-15 Febrero. Universidad de las Islas Baleares. (<http://goo.gl/EHq2Jo>) (29-04-2014).
- SAURO, J. (2011). *Measuring Usability with the System Usability Scale (SUS)* (<http://goo.gl/63krpp>) (29-04-2014).
- SCHÖN, D.A. (1998). *El profesional reflexivo: ¿cómo piensan los profesionales cuando actúan?* Barcelona: Paidós.
- VÁZQUEZ-CANO, E. (2013). El videoartículo: nuevo formato de divulgación en revistas científicas y su integración en MOOC. *Comunicar*, 41(XXI), 83-91. (<http://doi.org/tnk>).
- YANG, S., ZHANG, J., SU, A. & TSAI, J. (2011). A collaborative multimedia annotation tool for enhancing knowledge sharing in CSCL. *Interactive Learning Environments* 19, 1, 45-62. (<http://doi.org/c-dxd7>).

Resumen

1. Resultados

La tesis ha concluido con diversos y diferentes resultados dependiendo de su naturaleza y dimensiones, como han sido:

1. *Se han obtenido resultados en el desarrollo tecnológico y la experimentación o prueba de concepto de las tecnologías de federación, gracias al enfoque de evaluar centrado en el usuario, donde el instrumento de usabilidad y satisfacción creado y utilizado ha servido de estrategia para mejorar estos servicios y herramientas, tanto para comparar las opiniones de los usuarios cuando utilizan herramientas diferentes, como cuando son herramientas con similares funciones.*

La primera aportación de los artículos compendiados aborda esta dimensión conceptual de los sistemas federados de la eRúbrica, como también las diferentes funcionalidades de la herramienta. Es decir, se ha desarrollado un servicio tecnológico basado en web que ha sido una prueba de concepto y modelo de herramienta federada. Entre otras ventajas aspira a un avance y mejora en la facilidad o usabilidad del acceso de los servicios federados por los usuarios. La tesis ha contribuido a medir y validar la usabilidad y satisfacción de los usuarios cuando utilizaban esta tecnología de federación. Al mismo tiempo, el trabajo de la tesis ha permitido mejorar las funcionalidades de la propia herramienta eRúbrica de GTEA, gracias a las continuas evaluaciones sobre usabilidad y satisfacción pasadas a los distintos grupos de usuarios que la utilizaron. Al final de este trabajo, y junto con otros trabajos previos en la tesis donde se validaba el instrumento de medida con el instrumento SUS, podemos concluir que se ha conseguido un cuestionario para medir la usabilidad y satisfacción por los usuarios de servicios y herramientas web educativas.

Al mismo tiempo, fruto de esta prueba de concepto y primer resultado ha sido la figura 1 pp.577 dentro del primer artículo, que muestra el modelo definitivo logrado por el proyecto para generar un “Entorno Federado para la Colaboración”. Esto supone poder desarrollar modelos de prácticas y proyectos de colaboración entre estudiantes de diferentes instituciones universitarias. La facilidad en el acceso federado -con un solo clic- con su propia clave institucional, los estudiantes de distintas universidades pueden trabajar en modelos de evaluación colaborativa y de pares, compartir proyectos, contenidos, recursos y la propia tecnología de la eRúbrica, entre otras herramientas. Por lo que, el principal alcance de la tesis en este aspecto ha consistido en hacer realidad en la práctica de aula y con un modelo de evaluación formativa mediante eRúbrica, el paso del modelo teórico a una evaluación de la herramienta con el instrumento de medida.

Desde una dimensión tecnológica, la tesis ha permitido exponer y discutir el concepto de eRúbrica en un sistema federado de instituciones de educación superior con base en el proyecto de investigación I+D+i EDU2010-15432. Tomando como uno de los factores importantes de medida para los usuarios finales la usabilidad y satisfacción en el uso de la herramienta y en los mecanismos del acceso federado a la propia herramienta, y los propios entornos personales de los estudiantes o usuarios finales de un contexto de prácticas tan específicos como se dan en la enseñanza universitaria.

2. *En la tesis se ha logrado un producto final como es un instrumento de medida validado y fiable para evaluar la usabilidad y satisfacción de los usuarios de herramientas o servicios de eRúbrica federada.*

Los resultados han dado un coeficiente de fiabilidad alfa de Cronbach de 0.889, esta es una fiabilidad aceptablemente alta considerando que sólo se tienen 22 enunciados. En el instrumento se ha obtenido una media total de 80.66 y una desviación típica de 12.031. Este instrumento se compone de algunas variables descriptivas sobre los usuarios adaptable a cada situación o interés seguida de 22 ítems en una escala Likert del 1 al 5. 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. También este resultado nos proporcionó un estudio del instrumento de evaluación de la usabilidad y satisfacción de cualquier herramienta en la Web.

Dada la cantidad de herramientas y servicios que existen en internet para un mismo tipo de servicio y objetivo, todas ellas se hacen competitivas y sobreviven gracias a la valoración por los usuarios finales. Por lo tanto, este *instrumento de medida* que se ha obtenido es importante para que los gestores de herramientas conozcan la opinión de los usuarios, de modo que les permita de forma muy rápida medir la facilidad de acceso, uso y satisfacción de los servicios y herramientas para contextos específicos como son los educativos. Esta aportación está especialmente recogida en el segundo artículo presentado, con los resultados en un grupo y contexto específico para la enseñanza y aprendizaje universitario. Igualmente, y dentro de este segundo artículo, también se recomendaban mejoras de la eRúbrica que fueron realizadas dentro del proyecto I+D+i EDU2010-15432, como han sido: el diseño gráfico, las condiciones técnicas así como las opciones de guardar y ver, las incompatibilidades surgidas, y la necesidad de añadir una breve explicación en línea sobre el funcionamiento de la herramienta.

3. *Otro resultado de la tesis es la aplicación del instrumento de medida para evaluar la usabilidad y satisfacción de otras herramientas web educativas diferentes a la inicialmente estudiada de la eRúbrica.*

En el segundo artículo se presenta los resultados de la aplicación del instrumento de usabilidad y satisfacción a otras herramientas. En el que se compara los datos entre las herramientas Catool, Mahara, Moodle, eRúbrica y Webquest. La muestra fue de 198 personas con 84 respuestas completas, donde el 78,6% eran 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. De entre los resultados, 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. El soporte empleado mayoritariamente es el ordenador. Aunque hay que señalar que todos los que dicen haber utilizado el Teléfono móvil han sido en el caso de la e-Rúbrica. El navegador más utilizado es el Google Chrome, seguido de Mozilla Firefox. Estas respuestas han sido obtenidas en el contexto de la práctica, dado que siempre se evaluaba en un proceso de enseñanza en laboratorios, donde los estudiantes disponían de equipos informáticos. Entendemos que posiblemente fuera de este contexto los resultados serían diferentes.

Un análisis de conglomerados nos muestra que existen dos grupos de estudiantes: 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 la mayoría manifiestan una mayor usabilidad y satisfacción. Para una mejor interpretación de los resultados se puede expresar las puntuaciones directas en una escala típica derivada de 0 al 100. 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 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.

4. El tercer artículo se presentan los resultados de la aplicación de la eRúbrica federada en un diseño de evaluación formativa y contexto de enseñanza universitaria, como prueba de concepto para la autorregulación de los aprendizajes mediante la autoevaluación y evaluación por pares.

Se obtuvo los resultados de comparar la evaluación de un grupo A (Experimental) y un grupo B (Control) donde el grupo experimental conocía los estándares y criterios de evaluación mediante eRúbrica, y el otro grupo no. La experiencia se realizó en la exposición y asimilación de contenidos en la presentación de proyecto I+D+I. Donde el docentes y los estudiantes evaluaron y compararon sus calificaciones, obteniendo los índice de acercamiento y dispersión. A la vez que se obtuvo los datos de las autoevaluaciones de los propios estudiantes.

Por un lado, en los grupos de control los intervalos de las calificaciones obtenidas fueron más amplios, oscilando desde el 5 hasta el 9,5; mientras que en el grupo experimental de la misma materia las notas oscilan entre el 8 y el 9,5. En el caso de la materia “Hacia una escuela inclusiva: modelos y prácticas” en el Grupo de Control las notas oscilan entre el 5 y el 9,5; siendo la media obtenida 7.5. En el Grupo Experimental, las notas van desde 6.5 a 9 y la media coincide con la obtenida en el Grupo de Control. Y por otro lado, y en la medida en que las “evidencias” son más claras para el alumnado y menos subjetivas, las valoraciones a las mismas suelen estar repartidas entre todos los valores (de 1 a 5). Cuando la evidencia es menos inteligible para el estudiante, éste suele otorgarle valores intermedios. Con lo cual se pudo comprobar que la herramienta ha tenido un efecto positivo en los resultados académicos y principalmente en la autoevaluación de los aprendizajes, al conocer mejor los criterios y las evidencias de aprendizaje han permitido mayor autodirigir sus aprendizajes.

5. En el cuarto artículo, el instrumento de medida construido permitió la evaluación y las orientaciones de mejora en el diseño de la herramienta Open Video Annotation -OVA-.

Los resultados de este trabajo muestran la evaluación de las funciones de anotaciones de

vídeo añadidas al Mooc de edX; si bien, en esta plataforma se instalaron todas las funciones de Open Video Annotation –OVA-: anotaciones de vídeo, texto, sonido e imagen de calidad (estas dos últimas en fases experimental). Durante el curso 2013-14 se experimentó las dos herramientas – CaTool y OVA- con el mismo docente, metodología, laboratorio de clase y grupo de estudiantes (180 en total) de segundo curso del Grado de Pedagogía.

En este estudio comprobamos que las respuestas eran reflexivas y pensadas, y no se habían contestado por el simple hecho de rellenar el cuestionario. Detectando y eliminando del análisis estadístico 16 respuestas que con valores semejantes en los bloques correspondientes a sentencias directas y en los de sentencias indirectas, considerándose como respuestas no válidas. Para el análisis de los datos se realizó la transformación $y=6-x$ en los valores correspondientes a las sentencias indirectas. Los resultados fueron concluyentes, existiendo diferencias entre ambas herramientas. Donde, al total de las puntuaciones para cada una de las herramientas, se observa que a partir de un total de 105 puntos hay más valoraciones para OVA que para CaTool, y ocurre lo contrario para las valoraciones inferiores a 105 puntos. Estudiándolo por factores también hay diferencias significativas en el factor de usabilidad, pero no en el de satisfacción.

De las observaciones escritas por los encuestados apoyan los resultados obtenidos en el análisis del instrumento pudiéndose resumir los siguientes resultados: En general, todos valoran estas herramientas como fáciles, útiles e innovadoras. Los aspectos negativos se debieron principalmente a problemas de índole técnica: acceso a Internet, lentitud del servidor en el que se alojan o limitaciones por el navegador en la versión beta.

La comparación del mismo instrumento entre las dos herramientas nos dieron diferencias significativas a favor de OVA especialmente en los ítems siguientes:

- La aplicación me resultó agradable
- Fue agotador utilizar la aplicación
- Se puede usar sin necesidad de explicaciones previas
- He necesitado ayuda para acceder
- Me encontré con problemas técnicos
- Requiere ayuda de un experto
- El tiempo de respuesta en la interacción es lento

2. Conclusiones

Siguiendo la estructura propia de los objetivos planteado en la tesis, vamos a presentar las conclusiones finales del proyecto de tesis.

1. Se ha diseñado una herramienta web federada centrada en el usuario, como nos recomienda (Bevan, 1998). Este diseño centrado en el usuario nos permitió crear un servicio web federado en GTEA con acceso para todos los usuarios de todas las universidades españolas asociadas a RedIris. Además, se amplió el objetivo inicial a otras universidades extranjeras con acceso mediante EduGain. Animados principalmente por la presencia de extranjeros en movilidad Erasmus en los grupos de estudiantes estudiado (Objetivo 1). Resultados expuesto principalmente en la primera publicación del compendio. El sistema de acceso federado rebasó las previsiones establecidas en los objetivos llegando a otras universidades extranjeras, especialmente latinoamericanas por este acceso federado de eduGain y uno propio del grupo GTEA. Hecho que ha permitido una mayor expectativa en la aplicación del instrumento de usabilidad y satisfacción aplicado a otros contextos universitarios y herramientas en Latinoamérica, por cuanto que se ha creado un sistema federado en el Sistema Nacional de Educación a Distancia –Sined- México (2012) y más recientemente en la Universidad de Casa Grande en Ecuador (2015).
2. Se ha obtenido un instrumento de alta fiabilidad y validez para medir la usabilidad y satisfacción de servicios de eRúbrica del proyecto (Objetivo 2), como puede comprobarse en los diferentes trabajos presentados en este compendio, especialmente el segundo artículo. A la vez que, se ha aplicado a otras herramientas web educativas diferentes a la eRúbrica, (Mahara, Moodle, Prezzi, Webquest y foro); por lo que se concluye que este instrumento de medida sirve para evaluar cualquier otra herramienta web de Internet (Objetivo 3), con una alta fiabilidad del 0.889, siendo la media total de la escala 80.66 con una desviación típica de 12.031.

Los resultados concluyen un aspecto bastante evidente y medible, como ha sido comprobar que una herramienta complicada de usar será menos utilizada que otra que sea más fácil; asimismo ocurre con la satisfacción.

3. En Internet existen muchos servicios y herramientas para un mismo tipo de servicio o funcionalidad, por lo que, se hace necesario evaluar la usabilidad de las herramientas y la satisfacción de los usuarios. Este hecho hace relevante el trabajo presentado aquí por cuanto que es un instrumento fácil y rápido de aplicar, no importa el soporte con el que se responda (móvil, tableta, computadora de mesa...).
4. La aplicación del instrumento ha mejorado el servicio y funcionalidad de la herramienta eRúbrica, para la evaluación formativa y por competencia mediante eRúbrica, especialmente para que los docentes puedan desarrollar metodologías de evaluación de pares,

autoevaluación y evaluación de grupos en contextos universitarios (Objetivo 4), como puede observarse en la tercera publicación. Obteniendo un impacto positivo en los aprendizajes para los grupos de experimentación con eRúbrica frente a los de control.

5. Se concluye igualmente que el instrumento de medida permite ser un modelo y estrategia valiosa para comparar servicios y herramientas web educativas con similares funciones. Como ha sido OVA y Catool que presentan sus resultados en el artículo cuarto del compendio (Objetivo 5). Este resultado es muy interesante por tres razones:

Primero, porque existen una gran cantidad de herramientas en Internet, en muchos de los casos con similares funciones. Por lo que, el instrumento sirve como modelo para comparar dos herramientas con funciones similares, y ofrece información relevante y rápida a los docentes para tomar decisiones para elegir la herramienta que mejor valoran los usuarios.

Segundo, dado que es una herramienta implementada dentro de la plataforma MOOC de edX, el *instrumento de medida* creado puede servir para evaluar otros servicios más amplios que una herramienta o servicio exclusivo y único, como sería otras herramientas implementadas en una plataforma.

En tercer y último lugar, la herramienta OVA es un servicio de vídeo, muy empleado en las plataformas MOOC como edX. Plataformas que últimamente han sido muy valoradas y a la vez, cuestionadas por el alto índice de abandonos y bajo número de acreditaciones. Un instrumento de medida como el creado en esta tesis que nos permita tomar evaluaciones rápidas de las opiniones de los usuarios, es necesario para tomar decisiones igualmente rápidas en estas plataformas con gran número de usuarios.

6. La tesis y el diseño de instrumentos de medida para la usabilidad y satisfacción han permitido conocer la respuesta de los usuarios finales en un contexto más amplio que una plataforma, como ha sido, considerar Internet como la auténtica plataforma. El acceso y federación de identidades han sido utilizados y evaluados para desarrollar un nuevo modelo de ambientes de aprendizaje más personalizados. Esta situación se da en el caso en el que docentes y estudiantes son capaces de elegir, por ejemplo la eRúbrica de una lista de posibles herramientas que, o bien pertenecen a su institución, o pueden pertenecer a otra instituciones dentro del sistema federado. Al centrarse en la opinión y valoración por los usuarios o grupos, también estamos enfocando el proceso de enseñanza en el aprendizaje más que en la tecnología.

En suma, mediante herramientas federadas es mayor la facilidad de acceso a los servicios (usabilidad al utilizar un solo “clic”) y mayor satisfacción a la hora de construir sus propios entornos personales, y con ello, generar un mejor aprendizaje. Esto nos reafirma en la necesidad de que las universidades ofrezcan sus servicios federados, y que todas sus herramientas deberían realizar al menos un estudio de usabilidad y satisfacción por los usuarios finales como el desarrollado y propuesto en esta tesis.

Igualmente, y una vez alcanzados los diferentes objetivos específicos, podemos concluir que se ha logrado el objetivo general de la tesis con “La elaboración y validación de un nuevo

instrumento de medida para evaluar la usabilidad y satisfacción de herramientas y servicios web educativos”. Este trabajo tiene una especial vinculación y complementariedad al proyecto I+D+i desarrollado al trabajar con eRúbricas y llegar a la conclusión de que son unas herramientas que pueden facilitar los procesos de autoevaluación y autorregulación en el aprendizaje de los estudiantes. Además está de acuerdo a las recomendaciones actuales del Espacio Europeo de Enseñanza Superior en cuanto a la evaluación de las competencias. Donde el componente tecnológico es muy importante en estos procesos, por lo que la evaluación y opinión de estos servicios tecnológicos por medio de instrumentos de medida como el creado en esta tesis, es una aportación relevante que ayuda a la autorregulación y autoevaluación de los aprendizajes, dado que estos procesos y nuevos planteamientos pedagógicos están centrados y dependientes de los propios usuarios.

3. Nuevas perspectivas o Investigaciones futuras

Como posibles perspectivas futuras recogemos las siguientes:

- Este trabajo puede servir para proyectos futuros ampliando el estudio a otros contextos y poblaciones más extensas.
- En un momento del proceso de creación y mejora del instrumento de medida en la tesis, se elaboró preguntas enunciadas de forma directa e inversa, con el objeto de evitar unas respuestas irreflexivas de los usuarios, sin lectura de los enunciados y sin sentido, pudiendo descartar posteriormente las respuestas que distorsionaban los resultados. Como se obtuvo un resultado positivo al utilizar esta estrategia, se pensó que sería interesante aplicar y adaptarla a la eRúbrica, sobre todo cuando se pretende un modelo de evaluación de pares con esta herramienta. Por lo que, se propone como línea de trabajo futuro introducir evidencias con carácter positivo y negativo en la eRúbrica.
- El instrumento se ha aplicado a una herramienta muy específica en educación como es la evaluación formativa. Sería interesante estudiar otras variables asociadas al tipo de funcionalidad y naturaleza que ofrece esta herramienta como es la evaluación pero en contextos educativos y metodologías diferentes. Ponemos, por caso, un aprendizaje basado en proyecto dentro de herramientas de “juegos serios”, donde la evaluación entra también en juego con pares y autoevaluaciones por logros. También aplicar dicho instrumento para diferenciar desde grupos de estudiantes de diferentes áreas con las mismas metodologías, en la idea de conocer las diferencias encontradas por el perfil de los usuarios en las distintas titulaciones universitarias.
- La última versión del instrumento comparó dos nuevas herramientas como fueron Catool y Ova, que dieron lugar a la publicación del cuarto artículo presentado en esta tesis. Quedan otros datos aún sin explotar y que pueden generar otras publicaciones desde el enfoque de estudio comparativo de herramientas distribuidas frente a herramientas integradas. Tal es el caso de las herramientas analizadas con nuestro instrumento como la eRúbrica, Webquest, Catool, Foro, Prezzi y OVA frente a otras plataformas LMS que disponen de herramientas integradas caso de Mahara y Moodle.
- Hay un reciente auge muy importante de los cursos masivos abiertos en línea llamados MOOC (Massive open online course). Estos cursos están implementando herramientas como OVA y nuevos diseño de cursos con el objeto de hacerlos más interactivos y atractivos. Entendemos que sería una interesante línea el estudio de la utilidad, la satisfacción y la usabilidad de dichos cursos desde el punto de vista tecnológico y pedagógico. Desde el punto de vista tecnológico, dado que aún no están del todo definidas otras herramientas que se están integrando en dicha plataforma, como también, otros formatos diferentes que surgen como mini-MOOC, la integración del formato MOOC en plataformas LMS convencionales, etc. Desde el punto de vista pedagógico, hay un reto importante en los MOOC como es el gran número de abandono y el bajo número de

acreditaciones. Igualmente, la heterogeneidad de los perfiles y usuarios en los MOOC es todo un reto para la evaluación de la satisfacción y usabilidad de los usuarios.

- Sería interesante ampliar el estudio a otras titulaciones universitarias, incluso a otros niveles educativos. Igualmente, como la formación es una estrategia importante para el desarrollo de todos los ciudadanos, en la idea de formación para toda la vida, no importa la edad y su nivel profesional; entendemos que también podría ser muy interesante la aplicación de este instrumento para el estudio de usuarios genéricos en el uso de nuevos servicios tecnológicos educativos.

Referencias bibliográficas

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.
- Bevan, N. (1998). European usability support centres: support for a more usable information society. In *Proceedings of TAP Annual Concertation Meeting*. Retrieved from <http://nigelbevan.com/papers/tapus98.pdf>
- Bevan, N. (2009). Criteria for Selecting Methods in User Centred Design. In *I-USED*. Retrieved from http://www.nigelbevan.com/papers/Criteria_for_selecting_methods_in_user_centred_design.pdf
- Bevan, N., Kirakowski, J., & Maissel J. (1991). What is usability? In *Proceedings of the 4th International Conference on Human Computer Interaction, Stuttgart*. Elsevier. Retrieved from <http://www.nigelbevan.com/papers/whatis92.pdf>
- Brooke, J. (1996). *SUS: A quick and dirty usability scale, pages 189–194. Usability Evaluation in Industry*. Taylor and Francis, London.
- Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 213–218). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=57203>
- Coll, C., & Engel, A. (2014). Introduction: Personal Learning Environments in the context of formal education / Introducción: los Entornos Personales de Aprendizaje en contextos de educación formal. *Cultura Y Educación*, 26(4), 617–630.
<http://doi.org/10.1080/11356405.2014.985947>
- DIS, I. (1998). 9241-11 (1998): Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11-Guidance on usability. *International Organisation for Standardisation*.
- Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. *MIS*

- Quarterly*, 259–274.
- EduGain. (2015). <http://www.geant.net/service/eduGAIN/Pages/home.aspx>. Retrieved from <http://www.geant.net/service/eduGAIN/Pages/home.aspx>
- Estadísticas de YouTube. (2015). <http://www.youtube.com/yt/press/es/statistics.html>. Retrieved from <https://gteavirtual.org/rubric/>
- European Commission. (2008). *The use of ICT to support innovation and lifelong learning for all - A report on progress* (Commission staff working document No. SEC(2008) 2629). Brussels: European Commission. Retrieved from : <http://ec.europa.eu/education/lifelong-learning-programme/doc/sec2629.pdf>
- Ferre, X., Bevan, N., & Escobar, T. A. (2010). UCD method selection with usability planner. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 829–830). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=1869041>
- Gteavirtual. (2015). <https://gteavirtual.org>. Retrieved from <http://www.rediris.es/>
- Harper, B. D., Slaughter, L. A., & Norman, K. L. (1997). Questionnaire Administration Via the WWW: A Validation & Reliability Study for a User Satisfaction Questionnaire. In *WebNet* (Vol. 97, pp. 1–4).
- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79–102.
- Johannessen, G. H. J., & Hornbæk, K. (2014). Must evaluation methods be about usability? Devising and assessing the utility inspection method. *Behaviour & Information Technology*, 33(2), 195–206. <http://doi.org/10.1080/0144929X.2012.751708>
- Kelly, M. (1994). *MUSiC Final Report Parts 1 and 2: the MUSiC Project*. Hampshire, UK.
- Kirakowski, J. (1994). Background notes on the SUMI questionnaire. *Human Factors Research Group University College Cork, Ireland. Originally.*
- Kirakowski, J., & Cierlik, B. (1998). Measuring the usability of web sites. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, pp. 424–428). SAGE Publications. Retrieved from <http://pro.sagepub.com/content/42/4/424.short>
- Kirkup, G., & Kirkwood, A. (2005). Information and communications technologies (ICT) in higher education teaching—a tale of gradualism rather than revolution. *Learning, Media and*

- Technology*, 30(2), 185–199. <http://doi.org/10.1080/17439880500093810>
- Lewis, J. R. (1992). *Psychometric evaluation of the computer system usability questionnaire: The CSUQ*. Tech. Rep.
- Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In *Human Centered Design* (pp. 94–103). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-02806-9_12
- Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, 8(2), 3–6.
- McLaughlin, J., & Skinner, D. (2000). Developing Usability and Utility: A Comparative Study of the Users of New IT. *Technology Analysis & Strategic Management*, 12(3), 413–423. <http://doi.org/10.1080/09537320050130633>
- Molich, R., Ede, M. R., Kaasgaard, K., & Karyukin, B. (2004). Comparative usability evaluation. *Behaviour & Information Technology*, 23(1), 65–74.
- Proyecto I+D+i EDU2010-15432. (2010). http://erubrica.uma.es/INDEX.PHP?page_id=434. Retrieved from http://erubrica.uma.es/INDEX.PHP?page_id=434
- RedIris. (2015). <http://www.rediris.es/>. Retrieved from <http://www.rediris.es/>

Anexo A. El cuestionario

La última versión del instrumento es:

Cuestionario de usabilidad y satisfacción para la evaluación de herramientas web

Cuestionario

*** Herramienta a evaluar:**
Seleccione una de las siguientes opciones

Por favor escoja... ▼

*** Sexo:**

Femenino Masculino

*** Edad:**

Sólo se aceptan números en este campo

*** Experiencia que tienes con la aplicación:**
Seleccione una de las siguientes opciones

Acabo de conocerla
 Usuario básico
 Usuario medio
 Usuario avanzado

*** ¿Qué experiencia tienes con otras aplicaciones similares?:**
Seleccione una de las siguientes opciones

Ninguna
 Usuario básico
 Usuario medio
 Usuario avanzado

*** Soporte empleado con la herramienta:**

Seleccione una de las siguientes opciones

- Ordenador Windows
- Ordenador Mac
- Ordenador Linux
- Ordenador Otros
- Tableta Android
- Tableta Apple
- Tableta Otros
- Móvil Android
- Móvil Apple
- Móvil Otros
- Otro:

*** Navegador web usado:**

Seleccione una de las siguientes opciones

- Google Chrome
- Mozilla Firefox
- Internet Explorer
- Safari
- Otras
- Otro:

*** Fue fácil de usar:**

- 1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

*** Aparecieron incompatibilidades que me dificultaron el manejo:**

- 1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* La aplicación me resultó agradable:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Fue agotador utilizar la aplicación:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Se puede usar sin necesidad de explicaciones previas:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Me resultó complicada la edición:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Me resultó motivador trabajar con la aplicación:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Hubiera preferido usar otra aplicación conocida en lugar de ésta:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Me desorienté en algún momento con la aplicación:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Las opciones de los menús son claras:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Necesité poco tiempo para manejarla:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* He necesitado ayuda para acceder:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Me encontré con problemas técnicos:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Las opciones de la aplicación fueron las esperadas:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* Me resultó muy laborioso hacer algo con la aplicación:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* En algún momento tuve pánico:

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

* En caso de necesidad volvería a usar la aplicación en el futuro:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* Encontré opciones difíciles de interpretar:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* Requiere ayuda de un experto:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* El diseño gráfico es pobre:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* Recomendaría la aplicación a otras personas:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* El tiempo de respuesta en la interacción es lento:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

* Las respuesta dada es difícil de entender:

1 2 3 4 5

(1 Muy desacuerdo y 5 Muy de acuerdo)

*** Las ayudas para comprender la aplicación fueron útiles:**

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

*** La edición resulta muy flexible:**

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

*** En general, estoy satisfecho con la aplicación:**

1 2 3 4 5

? (1 Muy desacuerdo y 5 Muy de acuerdo)

Escribe las observaciones que consideres (Optativos):

Anexo B. Justificación de la contribución del doctorando en los artículos

En el primer artículo, nuestra aportación se centró en todo lo relacionado con la explicación en la creación y desarrollo de la herramienta eRúbrica de Gtea, sus funciones y alcances de la misma; como la explicación y justificación en esta herramienta del modelo tecnológico que la sustenta, como son las tecnologías de federación y los mecanismos de autenticación. Innovación esta última que facilita la usabilidad y por tanto la satisfacción de los usuarios al no tener que autenticarse una y otra vez en el acceso de múltiples servicios en internet, y en la creación de sus entornos personales de aprendizaje mediante herramientas y servicios web. Dejando para los dos otros autores la presentación, objetivos y explicación del diseño de investigación exclusivo en el proyecto I+D+i donde se enmarca nuestra tesis; así como, el estado del arte en las metodología de evaluación formativa y por competencia mediante eRúbrica.

En el segundo artículo trabajamos en la búsqueda de referencias bibliográficas que constituían el marco teórico y el estado del arte de nuestra investigación. Realizamos una recogida y revisión literaria que nos permitió conocer los avances hasta la fecha sobre la elaboración e instrumentos de medida para la usabilidad y satisfacción de usuarios. Al mismo tiempo, exponemos la elaboración de nuestro instrumento que partió desde otros encontrados en la literatura, así como los pasos y procedimientos para su validación. Fue tarea nuestra también la creación del instrumentos online, la recogida de estos datos para su análisis y la representación gráfica de resultados. Siendo los docentes del proyecto I+D+i quienes indicaron a sus estudiantes a que cumplimentaran los cuestionario online. Nuestro director y primer autor del artículo nos orientó en la creación del instrumento como en su validación. Trabajamos conjuntamente con el director en la redacción del análisis de los resultados de los datos obtenidos y las conclusiones.

En el tercer artículo por nuestra parte elaboramos el instrumento online, el primer autor y docente invitó a sus estudiantes a participar en el instrumento online. Nosotros recogimos los datos, se trataron y analizaron conjuntamente con la tercera autora, y elaboramos un informe estadístico y gráfico sobre sus resultados. La discusión de los resultados fue compartida en un trabajo en equipo por los tres autores.

En el cuarto artículo creamos e integramos las herramientas Collaborative Annotation Tool (CaTool) y Open Video Annotation -OVA en el portal de Gtea para que los docentes y primer autor lo utilizará con una metodología en su aula. Por nuestra parte, realizamos también la elaboración de instrumentos de recogida de datos online, recogimos los datos una vez el primer autor y docentes participantes invitaron a los estudiantes a contestar la encuesta. Fue nuestro trabajo el análisis y redacción de informes estadístico; en cambio, fue en colaboración con los demás autores la discusión de resultados, la redacción y revisión editorial, como la revisión del estado del arte.

Anexo C. Índice de impacto de las revistas

1. Revista European Educational Research Journal

Desde el análisis de Scimago encontramos los siguientes datos:

Publisher: Symposium Journals Ltd. Publication type: Journals. ISSN: 14749041

Está indexada en la base de Scopus estando en el Q2 desde Scimago Lab. H Index: 12. SRJ 0,42. total citas 89 Total References 1.859

The European Educational Research Journal (EERJ) is an international peer-reviewed journal devoted to Europeanization in educational research, and is the official journal of the European Educational Research Association (EERA). It publishes aspects of educational research which illuminate the cases and contents of the emerging borderless space of European educational research. A Policy Statement is available.

Policy-makers, administrators and practitioners with an interest in European issues are now invited to subscribe. The EERJ publishes peer reviewed articles, essay reviews and research reports (forms of research intelligence across Europe)

2. Revista de Docencia Universitaria. REDU.

Desde sus página podemos encontrar la siguiente información:

REDU. Revista de docencia universitaria, es una publicación científica que se dirige a un publico interesado por la educación superior en general y la docencia universitaria en particular: gestores, docentes e investigadores sobre este ámbito. Indexada en las siguientes bases de datos:

Academic Search Complete (EBSCO)

Base de datos ISOC del Consejo Superior de Investigaciones Científicas.

Sumarios ISOC - Ciencias Sociales y Humanidades. Producida por el Centro de Ciencias Humanas

CIRC: Clasificación Integradra de Revistas Científicas.

ERA: Educational Research Abstracts

ERIH Plus (European Reference Index for the Humanities and Social Science)

FUENTE ACADÉMICA PREMIER (EBSCO Publishing)

IRESIE: Índice de Revistas de Educación Superior e Investigación Educativa

DICE: Difusión y Calidad Editorial de las Revistas Españolas de Humanidades y Ciencias Sociales

INDEX COPERNICUS INTERNATIONAL

IN-RECS:

REDU. Revista de Docencia Universitaria ocupa en 2011 el puesto nº 38 de un total de 162 revistas del ámbito de educación. Se sitúa en el 2º cuartil con un índice de impacto de 0.198

REDU. Revista de Docencia Universitaria ocupa en 2010 el puesto nº 31 de un total de 164 revistas del ámbito de la educación. Se sitúa en el 2º cuartil con un índice de impacto de 0.095

MIAR: Matriu d'Informació per a l'Avaluació de revistes

RESH Universal Impact Factor Impact factor 2012: 0.9602

REDU se encuentra en el ranking de las 100 revistas de lengua española más citadas de todas las áreas según Google Scholar.

REDU ocupa el puesto número 9 de un total de 52 revistas científicas españolas. del área de Educación, según el índice H de Google Scholar Metrics (2009-2013).

3. Revista Comunicar.

Desde el análisis de Scimago y su propia página web encontramos los siguientes datos:

Publisher: Grupo Comunicar Ediciones. Publication type: Journals. ISSN: 11343478, 19883293

Es una revista indexada en Journal Citation Reports (JCR)

Comunicación 2013: 0,350 (2012), 56 de 72. 5 Year Impact Factor 0,293. Inmediacy Index: 0,100; Eigenfactor Score: 0,00013; Article Influence Score: 0,034.

Educación 2013: 0,350 (2012), 173 de 216. 5 Year Impact Factor 0,293. Inmediacy Index: 0,100; Eigenfactor Score: 0,00013; Article Influence Score: 0,034

En RECYT (FECYT-MEC)

Sello de Excelencia de Calidad FECYT 2013-2016 (12 indicadores) (BOE 12-06-2013)

Revista científica bilingüe en español e inglés en todos sus artículos, y abstracts en chino.

Decidida vocación internacional y latinoamericana en sus temáticas, lectores y autores.

22 años de edición y 1618 artículos publicados de investigaciones y estudios.

Presencia en 268 bases de datos internacionales, plataformas de evaluación de revistas, directorios selectivos, portales especializados, catálogos hemerográficos...

Riguroso y transparente sistema ciego de evaluación de manuscritos, auditado en RECYT; Consejo Científico Internacional y una red pública de revisores científicos de 365 investigadores de 27 países de todo el mundo.

Gestión profesional de manuscritos a través de la Plataforma OJS, de la Fundación de Ciencia y Tecnología, con compromisos éticos publicados para la comunidad científica de transparencia y puntualidad, antiplagio (CrossCheck), sistemas de revisión...

Alto nivel de visibilización con múltiples sistemas de búsqueda, DOIs, ORCID, pdfs dinámicos, EPUB..., con conexión a gestores documentales como Mendeley, RefWorks, EndNote y redes sociales científicas como Academia.Edu.

Especializada en educomunicación: comunicación y educación, TIC, audiencias, nuevos lenguajes...; monográficos especializados en temas de máxima actualidad.

Doble formato: impreso y on-line; Digitalmente, accesible a texto completo, de forma gratuita, para toda la comunidad científica y lectores de todo el mundo.

Coediciones impresas en España para Europa, y Ecuador y Chile, para América; Editada por Comunicar, asociación profesional no lucrativa, veterana en España (26 años) en educomunicación, que colabora con múltiples centros y Universidades internacionales.

En indexaciones (2014), «Comunicar» es la única revista española en JCR en Comunicación y la segunda española en Educación. En Scopus es Q2 en «Cultural Studies» y Q3 en Educación y en Comunicación. Es Revista de Excelencia RECYT 2013/16 y está inserta en ERIH+. Es la primera revista de Comunicación y de Educación en Google Scholar Metrics y la 48 del ranking 2014 (h19) de 423 de Comunicación del mundo (primera española).

Anexo D. Summary and Conclusions in English

1. Summary

At the beginning of this work, an original instrument was developed to measure the usability and satisfaction adapted to eRúbrica that was validated using as criteria the SUS questionnaire (Brooke, 1996). Afterwards the instrument was modified so that it was able to obtain a specific measure to evaluate the two factors, the usability and satisfaction, on service of educational web tools.

Among the questionnaires to measure usability, usefulness and satisfaction are highlighted the most interesting of the bibliography for the purpose this work below:

SUS questionnaire (System Usability Scale) (Bevan, N. et al., 1991; Brooke, 1996). It consists of 10 questions and is very simple to complete, to use and is easily comparable with other instruments. It is also free to use. It is applied once the users have worked with the application or tool that is going to be evaluated, in our case the eRubric.

The questionnaire QUIS (Chin et al., 1988; Harper et al., 1997), (Questionnaire for User Interface Satisfaction). It was developed in the late 80s, and currently are developing improved versions. In the following website you can find more detailed information and updates thereof: <http://lap.umd.edu/quis/>. It is an instrument of 27 questions that allow the evaluation mainly the user satisfaction. Plus, it also addresses the issue of usability. It is of subjective valuation, but unlike the SUS questionnaire is applied while it is working with the software, instrument ...

The USE questionnaire (Lund, 2001), (Usefulness, Satisfaction, and Ease of Use) is still for reviewing. It is very similar to SUS, but this questionnaire not only measures the usability, but also measures the usefulness, learning and user satisfaction. It consists of 30 questions and is one of the most comprehensive in measuring satisfaction, usability and usefulness, it is also very simple to implement as SUS and is a very simple scale to apply.

The SUMI questionnaire (Kirakowski, 1994) (Software Usability Measurement Inventory, SUMI). The questionnaire is used for assessing the quality of a given software taking into account the usability mainly (is quite specific to that effect) from the viewpoint of the end user.

The WAMMI questionnaire (Kirakowski & Cierlik, 1998) (Website Analysis and Measurement Inventory). It is an extension of SUMI questionnaire, not only measures the usability of web tools, but also has questions of user satisfaction and usefulness. It consists of 20 questions.

There are more tools in the study of art, but it would be excessive discuss them all. However, we can not forget to mention them, such as: MUSiC, (Kelly, M., 1994), which considers usability as measured by quality in use, QEM tool, it is a quantitative assessment methodology quality in web applications, considering four premises, reliability, efficiency, usability and functionality. ISOMETRIC, is used for comparison of the competitive products based on ISO 9241-10, adapting

to the principles of ISO 9241-11. MAGICA was the development of a methodology for measuring the usability with the following sections: 1) measurement of user satisfaction, for it makes use of WAMMI 2) time measurement task completion using the RUE measure (relative Effective User), 3) Measurement of cognitive effort, with questionnaires as SMEQ and CSUQ, (Lewis, 1992), among many.

Of these, the SUS questionnaire was the most useful tool which eventually was used to validate by the different advantages:

- For free
- Easy
- Full
- Many authors have validated the SUS questionnaire (Bangor et al., 2008; Molich et al., 2004).
- Presents a reliability coefficient of 0.91 for the eight questions of usability and 0.7 for the two questions of learning (Lewis & Sauro, 2009).

In the following phases of the development of the instrument were improved the questions from the results obtained in the different versions of the instrument. For improving reliability, the understanding of the questionnaire items was drafted in a more comprehensive way, and was revised by different experts involved in R & D project that helped with the sentence syntax.

For studying the validity, reliability and scale of the first used instrument was taken into account the criterion validity, Cronbach's alpha for reliability and scale of typical score derived respectively. This first instrument has been improving during the project process and applying it to different contexts as shown in the articles that were published, reaching the 4th version.

The identifier of the question does not follow a logical order, because it was thought that the best choices was to mix satisfaction and usability questions so that there were questions of the SUS and usability questionnaires. That way prevents to not answer questions by random. The idea is to apply both tests simultaneously as if it was two tests in parallel.

The study and population context were all students of the Faculty of Education at the University of Málaga, which was the R & D project coordinator, and the center associated with this university in Antequera, ranging in time from the year 2010- 11 to 2013-14. The sample of students varied according to the different moments of the study, however, it was applied mostly in groups of core subjects of the Degree in Primary Education and Pedagogy with approximately 75 students per group, plus other optional Pedagogy subjects in the third and fourth course.

For the first validation of the instrument was applied to students of 5 primary school groups and in the core subject of new technologies applied to education with a total of 375 students.

In the second article was shown the information of how was developed the second and third version of the instrument and its improvement. Collecting and analyzing data from 2013, where 198 people accessed the questionnaire, of which only 84 fully filled the questionnaire, 77 students and 7

teachers. The aim of this paper was to evaluate the usability and satisfaction of the different tools available online to university education and especially the eRubric, among other tools.

The students participating in this study belonged to the first year of the degree of Primary Education, Faculty of Education, University of Malaga, in the course of New technologies applied to education, so most had not previously used the eRubric, except repeaters.

Continuing the process of improving of the instrument was applied to evaluate the usability and satisfaction of eRubric in contexts and practices raised by the R & D was used the self-assessment and peer assessment. This study was published in the third article that is presented as part of the thesis. The sample was taken from the population of the Faculty of Education Antequera, center affiliated with the University of Malaga, and throughout the academic year 2011-2012, with different core subjects (Theory and Policy toward Preschool and school inclusive : models and practices) and with identical groups.

For the latest study to improve the usability and satisfaction instrument was applied to other different tools to previous studies, as were -Ova- Open video annotation and annotation tool Collaborative -Catool- (see fourth article). Wherein the third and fourth versions of the instrument was used. In this case, the sample was 180 students in the core subjects of teaching and technological resources education in the second year of the Education Degree, and two optional in the fourth year of the same degree.

2. Conclusions

Following the structure of the objectives proposed in the thesis, we present the final conclusions of the thesis project.

1. It has designed a web tool federated user-centric, as it is recommended in (Bevan, 1998). This user-centered design allowed us to create a federated web service in GTEA for the access for all users of Spanish universities associated with RedIris. In addition, the initial goal was expanded to others foreign universities with access by eduGAIN. This service was created mainly thanks to the presence of foreigners in mobility of Erasmus student groups that were studied (Objective 1). The results were exposed mainly in the first publication of the compendium.

The federated access system exceeded the forecasts made in reaching the goals, because the connections were extended to others foreign universities, especially the Latin American universities for this federated access of eduGAIN and our own mechanism in GTEA. This fact has allowed a greater expectation in the application of the instrument of usability and satisfaction applied to other university contexts and tools in Latin America, so that it was created a federated system in the National System of Distance Education -Sined- Mexico (2012) and most recently at the University of Casa Grande in Ecuador (2015).

2. It has obtained an instrument of high reliability and validity for measuring usability and satisfaction of the services of the eRubric project (Objective 2), as can be seen in the various papers presented in this work, especially the second one. At the same time, it has been applied to other educational web tools different to eRubric (Mahara, Moodle, Prezzi, Webquest and forum); so it is concluded that this measuring instrument allows to evaluate any other Internet web tools (Objective 3), with high reliability of 0.889, being the total average of the scale 80.66 with a standard deviation of 12.031.

The results conclude a fairly obvious and measurable aspect, as has been to verify that a complicated tool to use is less used than another easier; it is the same conclusion for satisfaction.

3. On Internet there are a lot of services and tools for the same type of service or functionality, so it becomes necessary to evaluate the usability of the tools and user satisfaction. This fact makes relevant the work presented here because it is easy and quick to apply the instrument, no matter the support where it is applied (either mobile, tablet, desktop computer ...).
4. The instrument has improved service and functionality eRubric tool for formative assessment and competition by eRubric, especially so that professors can develop peer evaluation, self-assessment and group evaluation methodologies in university contexts (Objective 4), as can be seen in the third publication. It is obtained a positive impact on learning for groups of experimentation using eRubric in comparison to control group.
5. It is also concluded that the measuring instrument can be a model and valuable strategy for

comparing educational services and web tools with similar functions. As OVA and Catool present their results in the fourth article of the work (Objective 5). This result is very interesting for three reasons:

First, because there are a lot of tools on the Internet, in many cases with similar functions. So, the instrument serves as a model for comparing two tools with similar functions, and provides relevant and timely information to professors to make decisions to choose the tool that best are valued for users.

Second, since it is a tool implemented within the MOOC platform of edX, the measuring instrument created can be used to assess other services broader than a tool or unique and exclusive services, as would be other tools implemented in a platform.

Third and finally, the OVA tool is a video service, widely used in MOOC platforms like edX. Platforms that have been lately highly valued and simultaneously questioned by the high dropout rate and low number of accreditations. A measuring instrument such as the one created in this thesis that allows us to take quick assessments of user reviews, is also needed to make quick decisions on these platforms with large numbers of users.

6. The thesis and design of measure instruments for usability and satisfaction have allowed to know the answer of end users in a broader context than a platform, as it has been to consider Internet as the authentic platform. Access and identity federation have been used and evaluated to develop a new model of more personalized learning environments. This situation occurs in the case in which professors and students are able to choose, for instance eRubric from a list of possible tools that either belong to the institution or belong to other institutions within the federated system. By focusing on the review and assessment by users or groups, we are also focusing the teaching process on learning rather than technology.

In short, the ease of access to the services is greater by the use of federated tools (usability to use a single "click") and greater satisfaction in building their own personal environments and thereby generate better learning. This reaffirms us in the necessity that universities can offer their federated services, and that all their tools should conduct at least one study of usability and satisfaction for end users as the developed and proposed in this thesis. Similarly, once achieved the different specific objectives, we can conclude that it has achieved the main objective of the thesis "Development and validation of a new measurement instrument to evaluate the usability and satisfaction of educational tools and web services."

This work has a special relationship and complementarity to the R & D project developed by working with eRubrics and conclude that they are tools that can facilitate the process of self-evaluation and self-learning of students. Furthermore, it agrees to the current recommendations of the European Higher Education regarding the assessment of competencies. Where technological component is very important in these processes, so that the evaluation and review of these technological services through measuring instruments as the one has been created in this thesis. It is a significant contribution that helps self-regulation and self-assessment of learning, since these

processes and new pedagogical approaches are focused and dependent on the users.

