

# Usabilidad de los Entornos Virtuales de Enseñanza y Aprendizaje. Evaluación por capas

**Ariel Ferreira Szpiniak**

Departamento de Computación. Facultad de Ciencias Exactas, Físico-Químicas y Naturales.  
Universidad Nacional de Río Cuarto, Río Cuarto, Córdoba, Argentina.  
aferreira@exa.unrc.edu.ar

**Cecilia V. Sanz**

III LIDI, Facultad de Informática. Universidad Nacional de La Plata, La Plata, Buenos Aires,  
Argentina.  
csanz@lidi.info.unlp.edu.ar

## Resumen

En este trabajo presentamos nuevos avances en la construcción de un modelo de evaluación de Entornos Virtuales de Enseñanza y de Aprendizaje basado en la usabilidad. El modelo propuesto forma parte de una tesis correspondiente a la Maestría de Tecnología Informática Aplicada en Educación, perteneciente a la Facultad de Informática de la Universidad Nacional de La Plata.

Se presenta una síntesis de los lineamientos generales y se detallan los aspectos principales para llevar a cabo la segunda, tercera, y cuarta capa de evaluación.

El modelo de evaluación está siendo aplicado al Entorno Virtual de Enseñanza y de Aprendizaje desarrollado y utilizado por la Universidad Nacional de Río Cuarto, con el objetivo de validarlo y realizar los ajustes necesarios en base a los resultados obtenidos.

**Palabras claves:** EVEA, usabilidad, evaluación.

## Introducción

Los Entornos Virtuales de Enseñanza y Aprendizaje (EVEA) forman parte del conjunto de aplicaciones informáticas diseñadas para la utilización de Internet con fines educativos. Su principal característica es la interactividad, como estrategia para favorecer el contacto entre docentes, alumnos y materiales de aprendizaje. En términos generales, suelen ser versátiles para poder adecuarse a diferentes propuestas y procurar que el diseño tecnológico acompañe al modelo

pedagógico. En realidad, los EVEA coadyuvan a la concreción de objetivos educativos al proveer una serie de herramientas que facilitan la gestión de usuarios y cursos, y los procesos de comunicación, evaluación, colaboración y distribución de contenidos, entre otros.

Los EVEA presentan un conjunto de funcionalidades para lograr que los procesos de enseñanza y aprendizaje, donde juega algún rol la virtualidad, puedan desenvolverse de la mejor manera. Pero, ¿cómo podemos saber si ese objetivo se logra? Para averiguarlo es necesario realizar una evaluación.

La evaluación puede orientarse de diferentes maneras, según lo que se pretenda evaluar. En general, la mayoría de los modelos de evaluación existentes analizan los EVEA desde el punto de vista funcional. Sin embargo, el análisis de las funcionalidades ideales no tiene en cuenta la forma en que se puede poner en práctica todo el potencial del EVEA. En la actualidad, la mayoría de estos sistemas poseen un núcleo común de herramientas, con lo cual carece de sentido seguir analizándolos casi exclusivamente desde el punto de vista de sus características funcionales. Por tal motivo, creemos necesario incorporar una forma de evaluar donde no se pierdan de vista estas características, pero se prioricen otros aspectos centrados en el modo en que los usuarios finales toman contacto con la funcionalidad [Mon06]. En este sentido, la usabilidad es un concepto muy útil, ya que tiene estrecha relación con la utilidad funcional. La usabilidad busca analizar cuán bueno es un sistema como para satisfacer todas

las necesidades y requerimientos de los usuarios. Las diferentes miradas, definiciones, y normas internacionales [Fer01] [ISO98] [ISO98b] [ISO01] [Kei03] [Lau95] [Nie93] [Nie00] [Nie03] [Red99] [Sch98] [Sha91] fueron revisadas y analizadas en el trabajo que marcó el comienzo de la tesis en la que se enmarca el trabajo [Fer07].

El método de evaluación de usabilidad que proponemos está orientado a una evaluación de producto, es decir, sobre entornos virtuales que ya han sido desarrollados, y están en pleno funcionamiento. También, puede facilitar la evaluación de nuevas versiones de un mismo sistema o la comparación entre ellas. Las ideas generales para la búsqueda de este modelo, se basan en una estrategia de cuatro niveles o capas, que descompone la evaluación de acuerdo a ciertos criterios, donde intervienen diferentes métodos, técnicas, actores, y miradas. Para ello, se parte del estudio general de las características del EVEA, hasta llegar a lo particular, donde participa el usuario final dentro de un contexto de uso, que trata de asemejarse lo más posible a la realidad [Fer09]:

*Primera capa:* destinada a analizar gran parte de la aceptabilidad práctica del EVEA. Esta conjuga la utilidad (usefulness) con otras variables como costos, compatibilidad, confiabilidad, y soporte, entre los más importantes [Nie93] [Zur99]. La utilidad a su vez está determinada por la utilidad práctica o funcional (utility), la usabilidad y la accesibilidad.

*Segunda capa:* orientada a evaluar la forma en que el sistema interactúa con el usuario, la interfaz que presenta y el modo en que permite realizar las tareas básicas. Los métodos de inspección, como el recorrido cognitivo y la evaluación heurística resultan adecuados para este tipo de acciones [Lew93] [Po92] [Wha94] [Alv05] [Nie90] [Nie92] [Mac93] [Man03] [Cat00].

*Tercera capa:* incorpora la opinión de los usuarios finales bajo un ambiente controlado por un observador. Para esto, son de gran ayuda los métodos de test de usuario como los de expresión del usuario en base a preguntas

[Woo98]. Los usuarios participan de la evaluación realizando una o más tareas que forman parte de un escenario tipo. A medida que el usuario interactúa con el entorno debe expresar en voz alta sus pensamientos, sensaciones y opiniones, mientras el evaluador capta y registra las impresiones que éste manifiesta oralmente; además de realizar preguntas directas cuando lo considera conveniente.

*Cuarta capa:* destinada exclusivamente a que los futuros usuarios finales aporten su punto de vista, dentro de un contexto o ambiente real. Lógicamente, para realizar este tipo de análisis también son adecuados los test de usuarios. En particular los remotos son muy eficaces, rápidos y fáciles de realizar [Hom03] [Mon06]. Así la persona realiza el test en su propio medio, con lo cual es posible evaluar el contexto de uso. Los test remotos se basan principalmente en el uso de cuestionarios para recolectar la información. El evaluador sólo está encargado de elaborar, suministrar el cuestionario, y posteriormente, procesar los resultados.

En artículos anteriores, hemos abordado los fundamentos y características generales del modelo, presentado la primera capa de evaluación, y la definición de los escenarios de uso necesarios para las restantes capas [Fer07] [Fer09] [Fer10]. En el presente trabajo describiremos el núcleo del modelo, es decir, pondremos el foco sobre la segunda, tercera, y cuarta capa de evaluación.

## **Evaluación de la usabilidad**

Las evaluaciones de usabilidad son factibles de utilizar en diferentes tipos de sistemas y distintas fases del proceso de desarrollo de software. Muchos de los métodos o técnicas existentes, pueden aplicarse en más de una fase, sin embargo, la mayoría está destinada a capturar problemas de usabilidad en etapas tempranas del desarrollo. Otra alternativa para evaluarla es como un atributo del producto final, en lugar de tomarlo en cuenta durante el desarrollo. Posiblemente, el hecho de centrarse en el contexto de uso sea uno de los rasgos más fuertes de esta mirada, ya que pone el

acento en la utilización concreta que el usuario hace del sistema, analizando en ese contexto las características del mismo. El contexto de uso involucra a los tipos de usuarios, las tareas, el equipamiento, y el entorno físico y social dentro del cual será utilizado el sistema. Para poder evaluarlo es necesario contar con usuarios reales, situaciones reales de trabajo y condiciones específicas. Por lo tanto, el producto deberá estar concluido o en una etapa avanzada de desarrollo. El usuario y el entorno en el cual se utilizará el sistema, pasan a ser ejes centrales para analizar su usabilidad.

En este trabajo, nos interesa evaluar los EVEA actuales desde el punto de vista de su usabilidad. La propuesta de evaluación se apoya en escenarios reales de uso, teniendo especial consideración por los alumnos y docentes, los objetivos que se proponen, las tareas específicas que realizan dentro del entorno durante las actividades de enseñanza y aprendizaje, el modelo mental que utilizan, el equipamiento e infraestructura que disponen, el lugar físico donde habitualmente se desenvuelven, y el entorno social en el cual están insertos.

## **Capa 2 del modelo de evaluación**

En un trabajo anterior señalamos que la segunda capa requiere de la participación de expertos, y métodos de inspección basados en heurísticas y recorridos cognitivos, para evaluar la forma en que el sistema interactúa con el usuario, la interfaz que presenta y el modo en que permite realizar las tareas básicas [Fer10]. Ambos métodos poseen una serie de similitudes respecto al trabajo con expertos, escenarios para los distintos roles de usuarios, y preguntas que debe hacerse el evaluador durante la aplicación. Producto de estas coincidencias estimamos conveniente conjugarlos para efectuar la evaluación de la interfaz desde ambos puntos de vista, pero fundamentalmente desde la evaluación heurística. Para ello, el especialista debe transitar una serie de escenarios de uso adoptando distintos personajes, navegando por las interfaces respectivas a fin de determinar el grado en que son respetadas las heurísticas,

tanto desde el punto de vista de los propósitos como desde los objetivos de los usuarios de un EVEA.

También, advertimos sobre la necesidad de definir una serie de escenarios que permitieran describir una situación de uso real, en la cual el experto pueda tomar el lugar del usuario final. Se trabajó en base a 21 escenarios que fueron construidos sobre un contexto de uso determinado, donde los personajes principales, docentes y alumnos, realizan las tareas más importantes y frecuentes relacionadas con dicho contexto [Fer10]. Estos escenarios son la base para que el evaluador efectúe el recorrido cognitivo y la evaluación heurística.

### **Heurísticas**

Se describen aquí las heurísticas seleccionadas para la evaluación. Si bien existen principios generales, dependiendo del área de aplicación suele haber normas propias o convenciones que se ven reflejadas en la interfaz de sus sitios y en la forma de trabajo de sus usuarios. Por tal motivo, las heurísticas están organizadas en dos niveles de evaluación, uno general, orientado a la tarea, y otro particular, orientado al diseño [Vil03]:

*Evaluación de alto nivel orientada a la tarea:* examina el aspecto y comportamiento desde el punto de vista de las tareas y objetivos propuestos por los escenarios de uso. Las heurísticas involucradas son: Complejidad, Visibilidad, Intuitividad, Topografía natural.

*Evaluación en detalle orientada al diseño:* analiza los aspectos concretos del conjunto de interfaces provistas por el EVEA, para completar las tareas. Contempla las siguientes heurísticas: Productividad, Retroalimentación, Control por parte del usuario, Reversibilidad y Manejo del error, Diseño y organización, Consistencia, Ayuda y documentación, y Estándares.

En ambos casos, cada heurística está acompañada de una serie de sub-heurísticas específicas. Estas heurísticas han sido seleccionadas, luego de una revisión bibliográfica específica [Fer07].

### **Evaluación de alto nivel orientada a la tarea**

Las tareas pueden ser evaluadas por diferente cantidad de evaluadores de acuerdo a factores tales como tiempo, disponibilidad y presupuesto. Por este motivo se proponen dos alternativas, una sencilla donde interviene **un solo evaluador**, y otra donde pueden intervenir **3 o 5 evaluadores**.

#### **Un solo evaluador**

El experto seleccionado debe adoptar los personajes para luego recorrer los escenarios definidos en cada caso. Para que la sesión de evaluación no sea tan larga y agotadora es conveniente acotar la cantidad de personajes, roles y escenarios.

Dentro de los docentes involucrados en los escenarios de uso tenemos un responsable y dos tutores, siendo éstos últimos los que desempeñan el rol más activo e imprescindible dentro del aula virtual. Por tal motivo, de los dos tutores, el experto simulará, en este caso, el rol de aquel que posee un contexto de uso que hace posible una mirada general sobre las tareas a realizar.

En el caso de los alumnos, los tres desempeñan un rol similar y contextos diferentes. El alumno seleccionado se enmarca dentro del contexto de uso que posibilita una evaluación más integral de las tareas seleccionadas. De acuerdo a lo planteado anteriormente, se utilizan 7 escenarios, 4 correspondientes al docente tutor y 3 al alumno. Para facilitar la evaluación se hace una síntesis de las tareas involucradas en los escenarios correspondientes a cada personaje, buscando obtener una serie de tareas de mayor nivel de generalidad, importancia y frecuencia de uso. Así fueron seleccionadas 8 tareas para el docente tutor y 6 para el alumno.

#### **Tres o cinco evaluadores**

En caso de contar con 3 o 5 expertos pueden utilizarse los mismos criterios, realizando el análisis **simplificado**, a nivel de cada tarea, o **exhaustivo**, distribuyendo entre los expertos los personajes, roles, tareas y contextos de uso implicados en la evaluación.

**Simplificado:** En este caso los expertos deben realizar cada una de las tareas asignadas

tomando como base el contexto de uso, general y particular, definido tanto para los docentes tutores como para los alumnos. Durante dicho tránsito, deben considerar las heurísticas seleccionadas para la evaluación de alto nivel orientada a la tarea. Las tareas involucradas son 14, 8 asignadas al docente tutor y 6 al alumno. En el caso de tener tres evaluadores, uno tomará el rol de docente tutor, y los dos restantes de alumnos. En el caso de cinco evaluadores, dos tomarán el rol de docente tutor, y los tres restantes de alumnos.

**Exhaustivo:** en este otro caso, el uso de los escenarios, y la adopción de roles, es más real, ya que puede hacerse una suerte de "simulación" siguiendo la secuencia de escenarios planteada. Al contar con varios evaluadores es posible distribuir todos los personajes entre los expertos: docente responsable, docente tutor, y alumno. Inclusive puede otorgarse más de un personaje a cada evaluador para cubrir los diferentes roles.

Independientemente de la cantidad de evaluadores y el tipo de evaluación utilizado, simplificada o exhaustiva, los evaluadores deben transitar cada escenario realizando las tareas incluidas en ellos, y tomando como base el contexto uso definido para cada caso. Para ello debe adoptar el o los roles asignados. Luego de completar cada escenario, se debe registrar su opinión en un cuestionario que reúne las heurísticas orientadas a la tarea de acuerdo a una escala de valoración cualitativa.

#### **Evaluación en detalle orientada al diseño**

Esta evaluación complementa la anterior analizando los aspectos concretos de cada interfaz. Podríamos decir que se trata de una evaluación heurística clásica, en el sentido que hace hincapié sobre el núcleo básico de heurísticas planteadas.

Existen muchas propuestas para evaluar de manera heurística las interfaces Web, desde aquellas que lo hacen de una manera general con preguntas muy abarcativas hasta aquellas minuciosas que involucran gran cantidad de preguntas. Para lograr un equilibrio y hacer

factible la aplicación del método, hemos propuesto una evaluación heurística acotada a los aspectos más relevantes de la interfaz, entendiendo que los EVEA actuales proporcionan herramientas de producción de contenido que garantizan el respeto por las reglas básicas de diseño, y proveen una estructura básica coherente enriquecida por los actores educativos.

La cantidad de evaluadores que participan no altera la forma de evaluación. La distribución de roles y escenarios es similar a la realizada para la evaluación orientada a la tarea, de acuerdo a la cantidad de evaluadores elegidos. Durante el tránsito de cada escenario, los expertos también deben tener en cuenta las heurísticas seleccionadas para la evaluación orientada al diseño. Al finalizar con todos ellos deben, completar otro cuestionario donde se detalla el nivel de conformidad alcanzado para cada sub-heurística. Este nuevo cuestionario posee la lista de principios heurísticos, donde cada uno es acompañado por el nivel de conformidad alcanzado según el criterio del evaluador, más un comentario que lo justifique en caso de ser necesario. La conformidad es medida a través de una escala ordinal de 1 a 5 puntos, cuyo valor inferior (1) representa que la conformidad nunca fue alcanzada, y el valor superior (5) que siempre lo fue.

### **Capa 3 del modelo de evaluación**

Las dos primeras capas focalizan su atención en la mirada que los expertos pueden hacer sobre el EVEA. En esta tercera, el objetivo es nutrirse de la opinión del usuario final, pero dentro de un ambiente controlado, y con la participación de observadores que guíen y faciliten el proceso. Para ello, es imprescindible la participación de usuarios reales.

La forma elegida para llevar a cabo esta etapa de evaluación es mediante la instrumentación de test de usuarios [Woo98]. El objetivo es establecer en qué medida el software se adapta a los estilos de trabajo reales de los usuarios, en lugar de forzar a los usuarios a adaptar sus estilos de trabajo al software. Los test de usuarios son considerados como el método de

evaluación más importante, ya que proporciona información sobre el uso real del software, observando a los usuarios finales mientras lo utilizan para realizar determinadas tareas, siendo la única forma válida de recoger datos fiables de usabilidad [Nie99].

Dentro de las alternativas existentes para este tipo de métodos, se ha seleccionado una variante del test de pensamiento en voz alta denominada test de expresión del usuario en base a preguntas [Nie94]. La metodología consiste en brindarle al usuario un escenario tipo, y solicitarle que efectúe las tareas involucradas en dicho escenario, bajo la atenta mirada de un observador. A medida que el usuario interactúa con el EVEA, el observador debe realizar preguntas directas acerca del producto o la tarea que el usuario está realizando, mientras que éste debe expresar en voz alta sus pensamientos, sensaciones y opiniones. El observador debe tomar registro de todo ello para procesarlo una vez concluido el test. El test de expresión del usuario en base a preguntas, es uno de los test de usuario cara a cara que mejor resultado arrojan para extraer información sobre el comportamiento e intenciones del usuario. Al basarse en la observación presencial, permite que el observador capte impresiones y sensaciones, recurriendo a preguntas para provocar manifestaciones explícitas respecto del producto. Además, es más fácil de implementar que el test de usuario clásico, ya que no es necesario contar con un laboratorio de usabilidad.

Es importante que los usuarios involucrados en los test abarquen los diferentes roles en que puede interactuarse con el EVEA. Por otro lado, dentro de cada rol deben seleccionarse usuarios con diferente nivel de experiencia en el uso del servicio Web.

### **Evaluadores**

Esta capa es muy costosa en términos de tiempo, cantidad de participantes y análisis, puesto que requiere de observaciones directas de distinto tipo de usuarios trabajando sobre el entorno. Los usuarios evaluadores se dividen en dos tipos: docentes y alumnos. Para cada

tipo se toman diferentes niveles de experiencia y frecuencia en el uso de Internet: novato, medio y avanzado.

Alcanzar la mayor objetividad posible es una premisa a lo largo de todo el proceso de evaluación. Si bien no existe consenso sobre cuál es la cantidad óptima de evaluadores a incorporar, es necesario buscar un equilibrio para que sea factible de realizar, 10 como mínimo y 30 como máximo [Lan99] [Nie93]. En tal sentido, para obtener un abanico de usuarios que cubra gran parte del espectro se seleccionan 10 usuarios como mínimo, de los cuales 4 deben ser docentes y 6 alumnos, con el objetivo de cubrir los 2 roles de tutor y los 3 de alumno, con al menos dos usuarios diferentes cada uno.

Como varios usuarios toman el mismo rol, se deben organizar sesiones de trabajo diferentes para que éstos no se superpongan. Las sesiones de test deben armarse, de manera tal que sea posible generar un simulacro a partir de los escenarios propuestos. Estos escenarios tienen cierta correlación entre ellos, que debe respetarse para poder realizar las tareas previstas adecuadamente. Por tal motivo los usuarios, pueden agruparse en dos sesiones de 5 integrantes (2 docentes y 3 alumnos), coordinadas y monitoreadas por dos observadores cada una.

### **Forma de evaluar**

Se proponen utilizar los 21 escenarios definidos, de los cuales 4 pertenecen al rol de docente responsable, 4 a cada docente tutor, y 3 a cada alumno. El método de test elegido es de expresión del usuario en base a preguntas, dentro del cual distinguimos dos tipos de **evaluación: directa e indirecta**. La primera es realizada por los usuarios, y la segunda por los observadores.

### **Evaluación directa**

El usuario final debe situarse en cada uno de los escenarios e interactuar con el entorno realizando tareas como si estuviera participando de un curso, asignatura, o capacitación con modalidad virtual. Para ello utiliza un guión provisto por el observador.

Luego de concluida la sesión, debe completar un cuestionario.

Los criterios tomados para la evaluación directa, son idénticos a los definidos en la segunda capa para la evaluación de alto nivel orientada a la tarea, que realizan los expertos (complejidad, visibilidad, e intuitividad y topografía natural). Al compartir los criterios también se comparte el instrumento de evaluación.

### **Evaluación indirecta**

El observador debe estar presente durante toda la sesión y registrar lo que acontece mientras el usuario realiza las tareas. El observador debe realizar preguntas directas al usuario, y éste debe expresar en voz alta sus observaciones, sugerencias, u opiniones. El usuario también puede realizarle preguntas al observador. Durante la sesión, y al finalizar la misma, se debe tomar registro y completar un cuestionario por cada usuario.

La evaluación indirecta pretende recabar información sobre un conjunto de métricas de usabilidad que buscan medir el grado de efectividad, eficiencia y satisfacción del EVEA. Para ello, en primer lugar, el observador debe tomar nota del rol y escenarios que posee cada usuario, y las observaciones, sugerencias u opiniones que vierte el mismo durante y después de la sesión. En segundo lugar debe analizar y registrar las siguientes métricas:

*Eficiencia:* porcentaje de tareas completadas

*Efectividad:* tiempo total invertido por el usuario en las tareas completadas, tiempo total ocupado en errores, tiempo total de aprendizaje, frecuencia de uso de la ayuda o documentación.

*Satisfacción:* veces que el usuario expresa frustración o enojo, utilidad del producto, satisfacción con respecto a las funciones y características, percepción de que la tecnología da soporte a las tareas según las necesidades del usuario.

## **Capa 4 del modelo de evaluación**

Se trata de una evaluación del entorno bajo condiciones reales de uso, es decir, en un

contexto establecido y con usuarios reales. Recordemos que no hay usabilidad sin contexto y sin usuarios, la usabilidad se aplica a usuarios concretos en contextos determinados [Vil03].

Las dos primeras capas utilizan los expertos para evaluar el EVEA, mientras que la tercera y la cuarta hacen lo propio, pero con usuarios reales. En particular, la tercera capa involucra a los usuarios a partir de un ambiente controlado, con la participación de observadores que guían, facilitan y registran el proceso. Esta cuarta, y última capa, hace lo propio pero dentro de un contexto real de uso, sin laboratorios de usabilidad, ni observadores, es decir, en un ambiente totalmente natural donde intervienen otras variables como la ubicación geográfica, puntos de acceso, tipos de equipamiento y acceso a la red, tipos de software y versiones, además de las tenidas en cuenta en la tercera capa.

La intención de la cuarta capa es conocer la opinión de los usuarios finales, brindándoles la posibilidad de que aporten su punto de vista desde un contexto real. Para realizar este tipo de evaluación también son adecuados los **test de usuarios**, en particular los **test remotos** como ya se mencionó [Hom03] [Mon06]. La idea es que el usuario realice el test en su propio lugar de estudio o trabajo, de manera que sea posible evaluar el contexto de uso. Los test remotos se basan principalmente en el uso de cuestionarios para recolectar la información. También se utilizan herramientas que analizan los “rastros” dejados por los usuarios mientras interactúan con el sistema. El evaluador es el encargado de elaborar, y suministrar el cuestionario. Luego de concluida la tarea por parte de los usuarios, debe procesar los resultados. Para ello utiliza las respuestas recibidas, enriquecidas con el análisis del comportamiento del usuario dentro del entorno virtual durante el test.

Al igual que en las capas anteriores es importante que los usuarios involucrados en los test abarquen los diferentes roles y posean distintos niveles de experiencia en el uso del servicio Web.

### **Evaluación real o simulada**

Para brindar flexibilidad al modelo de evaluación es posible aplicar esta última capa de dos maneras diferentes. La más sencilla es hacerlo sobre el EVEA basándose en el uso de uno o varios cursos reales, por parte de alumnos y docentes, y que tengan cierta semejanza con los escenarios de uso definidos anteriormente. La otra alternativa es recrear en el EVEA los escenarios que hemos definido (generar espacios virtuales para contener cursos e invitar a participar a docentes y alumnos). En este caso, el curso no es real sino un simulacro, pero los participantes son alumnos y docentes que trabajarán en sus propios contextos.

La evaluación real es una alternativa simple y ayuda a acortar los tiempos. El EVEA a evaluar ya debe estar en funcionamiento dentro de la Institución y contar con cursos adecuados para realizar la evaluación sin necesidad de recrear una situación hipotética. Este tipo de evaluación presenta una serie de desventajas porque en cursos reales pueden no estar presentes algunas tareas. Además, la cantidad de evaluadores no es la misma que para una evaluación simulada, y tampoco existe garantía de que estén representados todos los tipos de usuarios. Para reducir estos riesgos deben seleccionarse cuidadosamente los cursos y asegurarse que la cantidad de usuarios sea representativa, tanto en cantidad como en experticia. En la simulada pueden controlarse más estas variables.

### **Evaluadores**

Como gran parte del trabajo recae sobre el usuario, y no requiere observaciones directas, los costos en término de tiempo, cantidad de participantes, y análisis, son prácticamente los mismos, tanto para pocos como para muchos usuarios. Esto, al contrario que en la capa anterior, hace posible realizar una evaluación donde intervenga una cantidad importante de evaluadores. Por tal motivo, la muestra de usuarios puede ser numerosa. Siguiendo la línea de las capas anteriores, distinguimos dos tipos de usuarios: docentes y alumnos. Para cada tipo se toman los mismos niveles de

experiencia y frecuencia en el uso de Internet utilizados en la tercera capa.

No existe consenso sobre cuál es la cantidad óptima de evaluadores a incorporar porque hay muchas variables en juego. Seguramente, a mayor cantidad de evaluadores, mayor confiabilidad en los resultados, pero ello puede ir en desmedro de la rapidez para obtener los juicios, así como también correr el riesgo de que en el momento de contar con la información, ésta haya dejado de ser útil. Por tanto, hay que buscar un equilibrio que posibilite garantizar operatividad y objetividad. Existen diferentes propuestas, como la de Spyridakis [Spy92], Virzi [Vir92] o la de Landeta [Lan99] o Nielsen [Nie93c], específicamente para la evaluación heurística. Se propone trabajar aquí con 15 usuarios docentes y 27 usuarios alumnos, distribuidos de acuerdo a su experticia (novato, medio, avanzado). Consideramos que puede ser un número suficiente de evaluadores para obtener una visión general del contexto. Claro está, que es una cota mínima, pero no impide que puedan incorporarse más. El proceso puede dirigirse a una muestra de usuarios mayor, fundamentalmente en el caso de los alumnos, con el objetivo de contemplar los diferentes niveles de experiencia y frecuencia en el uso de Internet.

En este caso es posible aplicar herramientas tecnológicas, tanto para el suministro del cuestionario como para el procesamiento de los resultados. Mediante las herramientas de análisis de comportamiento del usuario puede obtenerse información respecto a su procedencia geográfica, así como también conexión, software y versión utilizados. De esta manera, podemos decidir cuándo finalizar la evaluación (revisando si se ha tomado un espectro diverso y suficiente), ya que nunca tendremos respuesta del 100% de los usuarios muestreados.

### **Organización**

Dada la cantidad de usuarios involucrados debe buscarse alguna forma de distribuirlos en diferentes cursos, de forma representativa y heterogénea, es decir, donde tengan

posibilidad de interactuar usuarios con diferentes roles y niveles de experticia.

Además de la voluntad manifiesta por el usuario en participar de la evaluación, es necesario conocer sus características básicas para determinar su nivel de experticia (novato, medio, avanzado) y su contexto general. Así se le podrá asignar un rol adecuado a cada uno. Para ello, debe realizarse un proceso de selección que posibilite agruparlos en función de los escenarios planteados. La forma elegida para realizar esta tarea es mediante una encuesta auto-administrada de forma on-line.

Los tiempos de la evaluación seguramente serán más largos que en las capas anteriores porque los evaluadores se encuentran bajo una situación real (realizan su vida cotidiana). El período de evaluación debería tener así un plazo máximo, aunque no todos hayan terminado las tareas por falta de tiempo o compromiso. Se estima adecuado pautar en una semana el plazo máximo para realizar las tareas planteadas. Cumplido ese plazo, el evaluador puede comenzar al procesamiento de los datos.

### **Forma de evaluar**

Hemos elaborado, en este caso, un test de usuario remoto. El usuario final debe situarse en un escenario concreto e interactuar con el entorno realizando tareas como si estuviera participando de una instancia de formación en modalidad virtual. Además, debe completar un cuestionario, una vez concluidas las tareas de cada escenario, y otro al terminar con todas las tareas solicitadas.

El test de usuario remoto está compuesto por tres cuestionarios, dos destinados al usuario final, para la evaluación directa, y uno al observador, para la evaluación indirecta. En este sentido es similar a la tercera capa pero el observador nunca toma contacto directo con el usuario.

### **Evaluación directa**

Los criterios tomados para la evaluación directa son similares a los definidos en la capa 2 y 3. Así, se propone una evaluación de alto nivel orientada a la tarea y otra orientada a



analizar el diseño y atributos de usabilidad (eficiencia, efectividad y satisfacción). En el caso de la evaluación orientada a la tarea, las heurísticas y forma de evaluación son idénticas, mientras que las heurísticas orientadas al diseño se mantienen, pero reduciendo la cantidad de sub-heurísticas.

Algunos instrumentos de evaluación también son compartidos. El cuestionario orientado a la tarea, que fuera utilizado por los expertos en la capa 2, y los usuarios en la capa 3, también será utilizado en esta capa. El cuestionario orientado al diseño y atributos de usabilidad es muy similar al utilizado en las capas 2 y 3.

### **Evaluación indirecta**

La evaluación indirecta pretende recabar información sobre un conjunto de métricas de usabilidad que buscan medir el grado de efectividad, eficiencia y satisfacción del entorno virtual en un contexto de uso determinado. Para ello el observador procesa y sintetiza los cuestionarios completados por los evaluadores (docentes y alumnos), accede a los datos arrojados por el análisis de log, y navega las aulas virtuales, utilizadas durante el test, para determinar el grado de completitud de las tareas planteadas en los escenarios. Todo ello, lo vuelca en un cuestionario que resume la opinión de los usuarios finales, según su rol y experticia, sumado a las huellas que los mismos dejaron en el EVEA durante el proceso de realización de las tareas propuestas en cada escenario de uso.

### **Grado de usabilidad**

El objetivo del modelo de evaluación que acabamos de presentar es aportar información para la toma de decisiones. Los parámetros a tener en cuenta para determinar la usabilidad del EVEA los decide cada equipo de evaluación. En general, cuando se habla de medir la usabilidad de un producto se hace referencia al nivel o grado de usabilidad del mismo. Esta es una medida empírica y relativa. Habitualmente se basa en opiniones de expertos, pruebas de laboratorio, observaciones de campo, y pruebas de usuarios. El resultado depende de las metas planteadas para un contexto de uso

determinado, o de una comparación con otras versiones o productos similares.

En muchas ocasiones, se intenta medir este grado de usabilidad mediante escalas cualitativas o cuantitativas. En una pequeña aproximación a este aspecto podemos señalar que el avance de una capa a la siguiente debe estar respaldado por una evaluación favorable. Cada capa debería tener un resultado, por ejemplo: muy aceptable, aceptable, no aceptable. La evaluación desfavorable de una capa, seguramente tendrá consecuencias importantes sobre la evaluación total, ya que difícilmente en las capas posteriores se mejore el grado de usabilidad del EVEA. Sin embargo, la amplitud del tema escapa los alcances de nuestro trabajo, por lo cual no está planteado, y se lo considera como un aspecto abierto.

### **Aplicación a distintos contextos**

Las cuatro capas plantean ciertos grados de acercamiento al usuario, siendo el primer nivel el más alejado. La transición de un nivel a otro, como señalamos recientemente, debe darse por resultados aceptables en las evaluaciones. Es decir que, en el caso de productos en desarrollo, la usabilidad debería mejorar de un nivel a otro, ya que se supone que los errores y dificultades más importantes han sido detectados y corregidos en las etapas previas. Sin embargo, como mencionamos al comienzo, el modelo que se propone resulta más adecuado para evaluar productos terminados, o para comparar nuevas versiones de un entorno respecto de las anteriores.

Las cuatro capas persiguen objetivos diferentes. La primera es la encargada de realizar una evaluación del entorno en general, las tres capas restantes se sitúan en un contexto de uso particular, por ejemplo el aula virtual de un curso. Para ello, las capas más cercanas al usuario utilizan escenarios de uso que permiten guiar y sistematizar la evaluación. En cada caso, se utilizan métodos y técnicas diferentes que se complementan entre sí y posibilitan enriquecer la evaluación.

Las capas propuestas pueden ajustarse a distintos contextos dependiendo de la

necesidad, nivel de profundidad de la evaluación, recursos y tiempo disponible. El modelo puede adaptarse en dos sentidos, prescindiendo completamente de algunas de las capas (ejemplo la capa 1), o seleccionando diferentes alternativas dentro de la segunda o tercera capa (cantidad de evaluadores, cantidad de escenarios, etc.), ya especificadas en la descripción de éstas.

## Conclusiones y trabajos futuros

Hemos recorrido un camino en el desarrollo de este modelo de evaluación. Se han definido cada una de las capas con sus propósitos, formas de aplicación de la evaluación, cantidad y tipo de evaluadores, y posibles instrumentos. Actualmente, el modelo se está poniendo a prueba mediante su aplicación para evaluar un EVEA específico denominado SIAT (Sistema Informático de Apoyo a la Teleformación), en el contexto institucional de la Universidad Nacional de Río Cuarto (UNRC). SIAT viene siendo desarrollado por la UNRC desde el año 2001, es utilizado por las carreras a distancia de la universidad, y como apoyo a las asignaturas de carreras presenciales [Thül1]. Uno de los objetivos es validar el EVEA y realizar los ajustes que sean necesarios para su mejora, desde el punto de vista de la usabilidad. Por otro lado, analizar los resultados obtenidos, y realizar una propuesta de ajuste al modelo acorde a las observaciones aportadas por quienes cumplieron los distintos roles de evaluadores.

## Bibliografía

[Alv05] Alva Obeso, M. E. Metodología de Medición y Evaluación de la Usabilidad en Sitios Web Educativos. 2005.  
[Car02] Carvalho, A. Usability testing of educational software methods, techniques and evaluators. 3° Simposio Internacional de Informática Educativa. Portugal. 2002.  
[Cat00] Catalán Vega, M. Metodologías de evaluación de interfaces gráficas de usuario. 2000.  
[Coo03] Cooper, A. About Face 2.0. P.22. Wiley Publishing Inc. USA, 2003.

[DeB00] De Benito, B. Herramientas web por entorno de enseñanza-aprendizaje, en CABERO, J. y otros (coords): Medios audiovisuales y nuevas tecnologías para la formación en el siglo XXI, Murcia. 2000.  
[DeB06]. De Benito, B; Salinas, J. Situaciones didácticas en los entornos virtuales de enseñanza-aprendizaje (EVEA) en la enseñanza superior: elaboración de un instrumento de análisis. EDUTEC. 2006.  
[Fer01] Ferré, X; Juristo, N; Windl, H; Constantine, L. Usability Basics for Software Developers. IEEE Software, vol 18 (11). 2001.  
[Fer07] Ferreira Szpiniak, A; Sanz, C. Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje. La importancia de la usabilidad. CACIC. Argentina. 2007.  
[Fer09] Ferreira Szpiniak, A; Sanz, C. Un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje basado en la usabilidad. Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología. Nro. 4, pp. 10-21. 2009.  
[Fer10] Ferreira Szpiniak, A; Sanz, C. Propuesta de un Modelo de Evaluación de Entornos Virtuales de Enseñanza y Aprendizaje. Escenarios de uso. V Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET'10), El Calafate. 2010.  
[Gaf00] Gaffney, G. What is a Scenario? Information & Design. 2000  
[Gon01] González Rodríguez, J; Olsina, L. Hacia la Medición de Calidad en Uso Web. 2001.  
[Har90] Harasim, L. Online Education: An Environment for Collaboration and Intellectual Amplification. En Harasim, L. (Ed.): On line Education. Perspectives on a New Environment. Preaguer Pub., New York. 1990.  
[Hom03] Hom, J. The Usability Methods Toolbox. 2003.  
[ISO01] Software Engineering – Product Quality – Part 1: Quality Model. ISO/IEC 9126-1. 2001.  
[ISO98] Ergonomic requirements for office work with visual display terminals. Dialogue principles. ISO 9241-10. 1998.

- [ISO98b] Ergonomic requirements for office work with visual display terminals. Guidance on Usability. ISO 9241-11. 1998.
- [Kei03] Keinonen, T. Usabilidad de los productos interactivos. 2003.
- [Lan99] Landeta, J. El método Delphi: Una técnica de previsión para la incertidumbre. Ed. Ariel Practicum. 1999.
- [Lau95] Landauer, T. The Trouble with Computers: Usefulness, Usability and Productivity. Massachusetts: Massachusetts Institute of Technology. 1995.
- [Lew93] Lewis, C.; Rieman, J.: Task-centred user interface design QA Practical introduction. 1993.
- [Mac93] Mack, R.; Nielsen, J. Usability inspection methods. ACM SIGCHI Bulletin-June, 28-33. 1993.
- [Man02] Manchón, E. El enfoque “persona” y los “escenarios”. [www.ainda.info](http://www.ainda.info). 2002.
- [Man03] Manchón, E. Evaluación heurística (o por expertos) de la usabilidad. 2003.
- [Mar05] Marcos, M; Rovira, C. Evaluación de la usabilidad en sistemas de información web municipales: metodología de análisis y desarrollo. En: Gascon, J., Burguillo F. y Pons A. (ed.). La dimensión humana de la organización del conocimiento. Universitat de Barcelona, España. 2005.
- [Mon06] Usabilidad: ¿Qué? ¿Cómo? ¿Cuándo? ¿Dónde? ¿Para qué? Francisco Montero. Un. de Castilla-La Mancha. España. 2006.
- [Mor04] Mordecki, D. Pensar primero. Biblioteca Concreta. 2004.
- [Nie90] Nielsen, J; Molich, R. Heuristic evaluation of user interfaces. Proceedings of the CHI'90, ACM Press, pp 249-256. 1990.
- [Nie92] Nielsen, J. Finding usability problems through heuristic evaluation. Proceedings of the CHI'92, ACM Press, pp 373-380. 1992.
- [Nie93] Nielsen, J. Iterative User-Interface Design. IEEE Computer. 1993.
- [Nie93b] Nielsen, K.; Landauer, T.. A mathematical model of the finding of usability problems. En CHI '93: Proceedings of the INTERACT '93 y CHI '93. New York. 1993.
- [Nie93c] Nielsen, J. Usability Engineering. Academic Press, 1993.
- [Nie99] Nielsen, J. Voodoo Usability. Alertbox. 1999.
- [Nie00] Nielsen, J. Usabilidad. Diseño de sitios web. Prentice Hall PTR. 2001
- [Nie03] Nielsen, J. Usability 101: Introduction to Usability. Alertbox . 2003.
- [Nor98] Norman, D. Psicología de los objetos cotidianos. Ed.Nerea. ISBN: 84-89569-18-5. 1998.
- [Pol92] Polson, P.G; Lewis, C; Rieman, J; Wharton C. Cognitive walkthroughs: A method for theory- based evaluation of user interfaces. International Journal of Man-Machine Studies 36, 741-773. 1992.
- [Red99] Redish, J., Dumas, J. A Practical Guide to Usability Testing, 1999.
- [Sch98] Schneiderman, B. Designing the User Interface: Strategies for Effective Human-Computer Interaction. 3°edition, Addison Wesley Ib.1998.
- [Spy92] Spyridakis, J. H. Conducting Research in Technical Communication: the application of true experimental designs. Technical Communications, Fourth Quarter, 607-624. 1992.
- [Sha91] Shackel, B. Usability – context, framework, design and evaluation. En Human Factors for Informatics Usability. pp. 21-38. Cambridge University Press. 1991.
- [Thü11] Thüer, S., Ferreira Szpiniak, A. Entornos Virtuales de Aprendizaje: Diseño de experiencias de usuario para la web 2.0. Conferencia Internacional ICDE – UNQ 2011.
- [Vil03] Villa, L. Usabilidad sin usuarios: heurística. [ww.alzado.org](http://ww.alzado.org). 2003.
- [Vir92] Robert A. Virzi, R. A. Refining the test phase of usability evaluation: how many subjects is enough? Hum. Factors, 34(4):457–468, 1992.
- [Wha94] Wharton, C; Rieman, J; Lewis, C; Polson, P. The cognitive walkthrough method: A practitioner's guide. In usability inspection methods. Nielsen, J. and Mack, R.L. (Eds.), John Wiley & Sons, pp.105-141. 1994.
- [Woo98] Woodward, B. Evaluation methods in usability Testing. 1998.
- [Zur99] Zurita, G; Sánchez, J; Nussbaum, M. Taller Internacional de Software Educativo – TISE'99. Universidad de Chile. 1999.