

Métodos de Evaluación e Inspección

Curso Interfaces de usuario

Teresa Onorati

Ignacio Aedo

Paloma Díaz

Ana Tajadura-Jiménez

Javier Sanz Rodríguez

Imagen de [Free Photos](#) en [Pixabay](#)

Tests de usabilidad

- Los tests o pruebas de usabilidad son centrales para entender la interacción con las interfaces
- Los objetivos y las preguntas se centran en cómo de bien los usuarios realizan tareas con el producto.
 - Implica grabar el rendimiento de usuarios típicos realizando tareas típicas.
 - Se mide el tiempo en completar la tarea, el número y el tipo de errores.
- Es común comparar productos o prototipos.

Experimentos y Tests de Usabilidad

- Los experimentos prueban las hipótesis para descubrir nuevos conocimientos al investigar la relación entre dos o más variables.
- Un test de usabilidad es una experimentación aplicada.
- Los desarrolladores verifican que el sistema pueda ser utilizado por la población de usuarios prevista para sus tareas.
- Se pueden llevar a cabo experimentos dentro del test de usabilidad

Tests de usabilidad

- Configuraciones controladas.
- Se observa a los usuarios. Se mide el tiempo de ejecución de tareas.
- Datos recopilados por video y registro de interacción (p.ej. pulsaciones de teclas).
- Los datos se utilizan para calcular los tiempos de ejecución e identificar y explicar los errores.
- Los cuestionarios y entrevistas de satisfacción de los usuarios proporcionan datos sobre las opiniones de los usuarios.
- Las observaciones de campo sirven para entender el contexto.

Condiciones de los tests

- Laboratorio de usabilidad u otro espacio controlado.
- Énfasis en:
 - seleccionar usuarios representativos;
 - desarrollar tareas representativas.
- Tareas generalmente de alrededor de 30 minutos.
- Las condiciones del test o prueba son las mismas para cada participante.
- El formulario de consentimiento informado explica los procedimientos y trata los problemas éticos relacionados con el test y la protección de los datos personales.

Tipos de datos

- Tiempo para completar una tarea.
- Tiempo para completar una tarea después de un tiempo específico alejado del producto.
- Número y tipo de errores por tarea.
- Número de errores por unidad de tiempo.
- Número de veces que se accede a la ayuda en línea y los manuales.
- Número de usuarios que hacen un error.
- Número de usuarios que completaron con éxito una tarea.

¿Cuántos participantes?

- El número es un problema práctico. Depende de:
 - calendario de las pruebas.
 - disponibilidad de participantes.
 - coste de ejecución de las pruebas.
- Típicamente 5-10 participantes.
- Algunos expertos sostienen que las pruebas deberían continuar hasta que no se obtengan nuevos conocimientos.

¿Cuántos evaluadores?

- Nielsen sugiere que, en promedio, 5 evaluadores identifican el 75-80% de los problemas de usabilidad.
- Cockton y Woolrych (2001) señalan que el número de usuarios necesarios para encontrar el 75-80% de los problemas de usabilidad depende del contexto y la naturaleza de los problemas.
- www.nngroup.com/articles/how-many-test-users/?lm=how-to-conduct-usability-studies&pt=report

Un método de 7 pasos

1. Determine las tareas de usuario más importantes.
2. Descubra qué aspectos del sistema son los más preocupantes.
3. Agrupe los elementos de 1 y 2, luego ordene los problemas por orden de importancia para los usuarios y la organización.
4. Para cada problema, formule una declaración de problema.
5. Para cada enunciado del problema, enumere las metas de investigación.
6. Para cada objetivo de investigación, enumere las actividades y los comportamientos de los participantes.
7. Para cada grupo de objetivos, escriba escenarios de usuario.

<https://www.nngroup.com/articles/ux-research-goals-to-scenarios/?lm=how-to-conduct-usability-studies&pt=report>

Tests de usabilidad

Ejemplo - Prueba de usabilidad del iPad

- 7 participantes con más de 3 meses de experiencia con iPhones.
- Firmaron un formulario de consentimiento informado explicando:
 - qué se le pediría al participante que hiciera;
 - el tiempo necesario para el estudio;
 - la compensación que se ofrecería por participar;
 - el derecho del participante a retirarse del estudio en cualquier momento;
 - una promesa de que la identidad de la persona no sería revelada; y
 - un acuerdo de que los datos recopilados serían confidenciales y estarían disponibles solo para los evaluadores
- A continuación se les pidió que exploraran el iPad.
- Por último se les pidió que realizaran tareas con ciertas especificaciones, asignadas de forma aleatoria.

Tests de usabilidad

- Problemas detectados:
 - El acceso a la web fue difícil.
 - La falta de "affordance" (capacidad de sugerir una acción) y retroalimentación.
 - Perderse.
 - No saber dónde tocar.
- Acciones de los evaluadores:
 - Informar a los desarrolladores.
 - Informe disponible para el público en nngroup.com.
- La accesibilidad para todos los usuarios es importante.

Diseño experimental

- Probar la hipótesis.
- Predecir la relación entre dos o más variables.
- La variable independiente es manipulada por el investigador.
 - Ejemplos: diferencias individuales entre los usuarios (género, edad, ...), tipo de información que se muestra a los usuarios, dispositivos que se usan, ...
- La variable dependiente es influenciada por la variable independiente.
 - Ejemplos: % de personas que finalizan las tareas, rendimiento en tiempo, # de clics sobre un elemento, ...
- Los diseños experimentales típicos tienen una o dos variables independientes.
- Diseño validado estadísticamente y replicable.

Diseño experimental

- **Diferentes participantes** (Diseño entre-sujeto): un solo grupo de participantes se asigna aleatoriamente a las condiciones experimentales.
- **Mismos participantes** (Diseño intra-sujeto): todos los participantes aparecen en todas las condiciones.
- **Participantes emparejados**: los participantes se emparejan de dos en dos, por ejemplo, según su experiencia, género, etc

Diseño experimental

Diseño	Ventajas	Desventajas
Diferentes (Entre-sujeto)	No hay efectos de orden	Muchos sujetos y las diferencias individuales son un problema
Mismos (Intra-sujeto)	Pocos sujetos, no hay que tener en cuenta diferencias individuales	Contra-balanceo necesario para evitar los efectos de orden
Emparejados	Igual que entre-sujeto, pero las diferencias individuales se reducen	No se puede asegurar un emparejamiento perfecto en todas las diferencias

Estudios de campo

- Los estudios de campo se realizan en entornos/contextos naturales.
- "In the wild" es un término para los prototipos que se usan libremente en entornos naturales.
- El objetivo es entender qué hacen los usuarios de forma natural y cómo la tecnología los afecta.
- Los estudios de campo se utilizan en el diseño del producto para:
 - identificar oportunidades para nueva tecnología;
 - determinar los requisitos de diseño;
 - decidir cómo presentar mejor la nueva tecnología;
 - evaluar la tecnología en uso.

Recogida y análisis de datos

- Observación y entrevistas:
 - Notas, imágenes, grabaciones
 - Vídeo
 - "Logging"
- Análisis:
 - Categorizado
 - Las categorías se forman de acuerdo a teorías:
 - Teoría fundamentada ("Grounded theory")
 - Teoría de la actividad

Presentación de datos

- El objetivo es mostrar cómo se están apropiando e integrando los productos en su entorno.
- Las formas de presentación típicas incluyen:
 - viñetas;
 - extractos;
 - incidentes críticos;
 - patrones y narrativas.

Inspecciones

- Varios tipos.
- Los expertos usan su conocimiento de los usuarios y la tecnología para revisar la usabilidad del software.
- Las críticas de los expertos pueden ser formales o informales.
- La **evaluación heurística** es una revisión guiada por un conjunto de heurísticas.
- Los **walkthroughs** (recorridos) implican pasar por un escenario planificado previamente que señala problemas potenciales.

Evaluación heurística

- Desarrollada por Jacob Nielsen a principios de la década de 1990.
- Basado en las heurísticas extraídas de un análisis empírico de 249 problemas de usabilidad.
- Estas heurísticas han sido revisadas para la tecnología actual por Nielsen y otros para:
 - dispositivos móviles,
 - "wearables" (dispositivos portátiles),
 - mundos virtuales, etc.
- Las pautas de diseño forman la base para el desarrollo de las heurísticas.

Evaluación heurística

- La evaluación heurística de los sitios web se centra en criterios clave (Budd, 2007)
 - Claridad
 - Minimizar la complejidad innecesaria y la carga cognitiva
 - Proporcionar a los usuarios un contexto
 - Promover una experiencia de usuario positiva y agradable

Heurísticas de Nielsen

1. Visibilidad del estado del sistema
2. Coincidencia entre el sistema y el mundo real
3. Control del usuario y libertad
4. Consistencia y estandarización
5. Prevención de errores
6. Reconocimiento antes que recuerdo
7. Flexibilidad y eficiencia de uso
8. Estética y diseño minimalista
9. Ayudar a reconocer, diagnosticar y recuperar la situación cuando se produce un error
10. Ayuda y documentación

Evaluación heurística

- Establece una lista apropiada de heurísticas.
- Selecciona tus evaluadores (diferentes de los usuarios finales, por lo general expertos en usabilidad y preferiblemente con experiencia en el dominio).
- Informa a tus evaluadores sobre qué hacer y cubrir.
- Primera fase de evaluación (alrededor de 2 horas para usar el producto libremente).
- Segunda fase de evaluación (aplicando las heurísticas individualmente).
- Regista los problemas (tú o los evaluadores).
- Sesión informativa entre los evaluadores para enumerar los problemas y sugerir posibles soluciones.

https://www.interaction-design.org/literature/article/heuristic-evaluation-how-to-conduct-a-heuristic-evaluation?utm_source=facebook&utm_medium=sm

Ventajas y problemas

- Pocos problemas éticos y prácticos que considerar porque los usuarios no están involucrados.
- Puede ser difícil y costoso encontrar expertos.
- Los mejores expertos conocen el dominio y los usuarios de la aplicación.
- Los mayores problemas:
 - Algunos problemas importantes pueden perderse;
 - A menudo se identifican muchos problemas triviales;
 - Los expertos tienen prejuicios o "biases".

Recorridos cognitivos

- Enfoque en la **facilidad de aprendizaje**.
- El diseñador presenta un aspecto de los escenarios de diseño y uso.
- El experto recibe las suposiciones sobre la población de usuarios, el contexto de uso y los detalles de la tarea.
- Uno o más expertos recorren el diseño del prototipo con el escenario.
- Los expertos son guiados por 4 preguntas.

Las 4 preguntas

- ¿Intentará y logrará el usuario el resultado correcto?
- ¿Se dará cuenta el usuario de que la acción correcta está disponible?
- ¿Asociará el usuario la acción correcta con el resultado que espera lograr?
 - Steve Krug, dice en su libro "No me hagas pensar" - "No importa cuántas veces tenga que hacer clic, siempre que cada clic sea una elección inconsciente y sin ambigüedades".
- Si se realiza la acción correcta; ¿verá el usuario que está progresando hacia su resultado previsto?

Recorridos cognitivos

- Ejemplo: Proceso de inicio de sesión en un sitio web
- Tareas
 - Abrir el navegador
 - Navegar al sitio
 - Hacer clic en el botón de "login"
 - Introducir el nombre de usuario en el campo "nombre de usuario"
 - Introducir la contraseña en el campo "contraseña"
 - Hacer clic en el botón de inicio

“Walkthrough” pluralístico

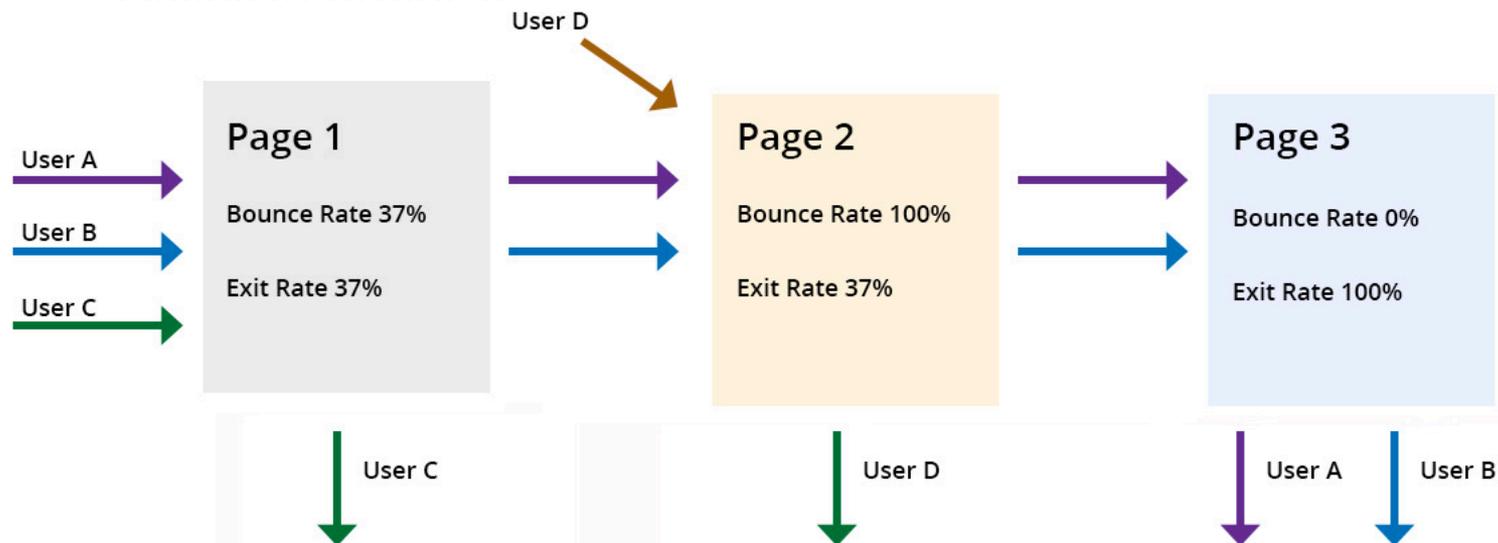
- Variación en el tema del recorrido cognitivo.
- Realizado por un equipo cuidadosamente administrado.
- El panel de expertos comienza trabajando por separado.
- Luego hay una discusión administrada que conduce a decisiones acordadas.
- El enfoque se presta al diseño participativo.
- También otras adaptaciones de recorridos cognitivos básicos.

Evaluación usando analítica

- Un método para evaluar el tráfico del usuario a través de un sistema o parte de un sistema.
- Muchos ejemplos: Google Analytics, Visistat, Learning Analytics, etc
- Registrar horas del día y direcciones IP del visitante.

Evaluación usando analítica

- Ejemplo: Google Analytics registra estos datos
 - **Tasa de rebote** = cuántos usuarios “nuevos” abandonan el sitio web después de visitarlo parcialmente (rebotes/entradas)
 - **Tasa de salida** = cuántos usuarios abandonan el sitio web (salidas/entradas)



Evaluación usando analítica

- Ejemplo: Google Analytics registra estos datos
 - Promedio de tiempo que pasan los usuarios que visitan el sitio web

Primary Dimension: **Page** Page Title Other

Secondary dimension Sort Type: Default

Pageviews advanced

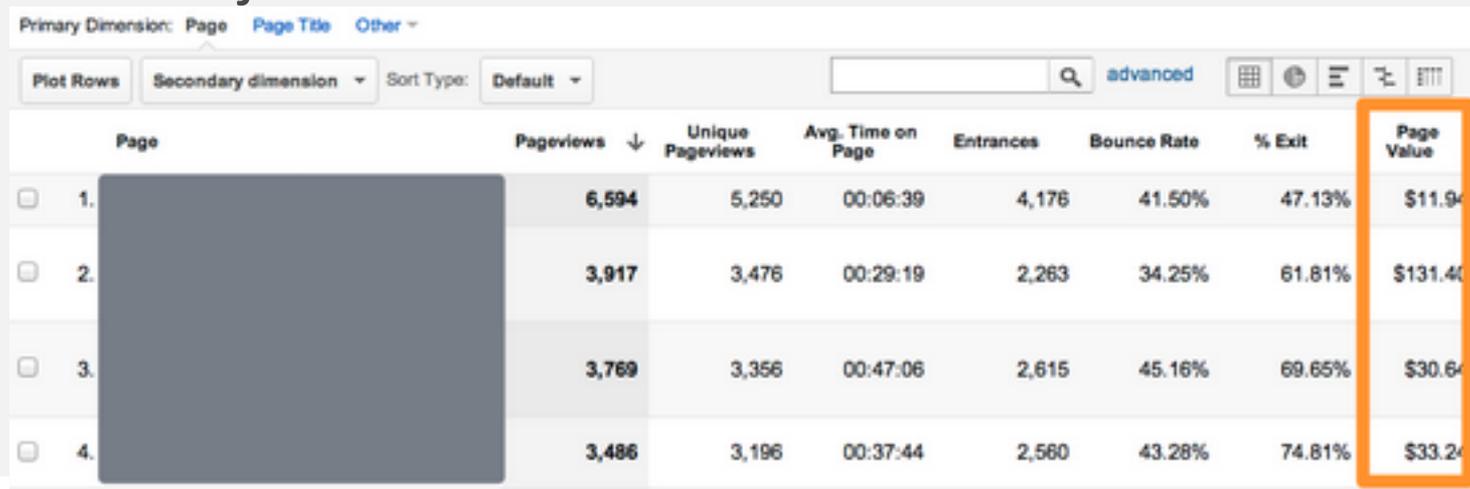
Avg. Time on Page (compared to site average)

Avg for View: 00:16:11 (0.00%)

Page	Pageviews	Avg. Time on Page
1. /	45,029 % of Total: 100.00% (45,029)	00:16:11
2. /5-logo-infographics-will-help-crack-c ompany-branding/	8,129	-60.82%
3. /web-design-trends-prediction-for-20 16-by-prince-pal/	3,120	56.40%
4. /web-design-portfolio/	2,481	196.76%
	2,157	-63.64%

Evaluación usando analítica

- Ejemplo: Google Analytics registra estos datos
 - Valor de página = precio dado a una página web en particular y que sirve para calcular la usabilidad. En el caso de los sitios web de comercio electrónico, el objetivo es el retorno de transacciones



Primary Dimension: Page Page Title Other

Plot Rows Secondary dimension Sort Type: Default

Page	Pageviews ↓	Unique Pageviews	Avg. Time on Page	Entrances	Bounce Rate	% Exit	Page Value
1.	6,594	5,250	00:06:39	4,176	41.50%	47.13%	\$11.94
2.	3,917	3,476	00:29:19	2,263	34.25%	61.81%	\$131.40
3.	3,769	3,356	00:47:06	2,615	45.16%	69.65%	\$30.64
4.	3,486	3,196	00:37:44	2,560	43.28%	74.81%	\$33.24