

City University Interaction Lab

Tema 9. Evaluación de la interfaz

Interfaces Persona Computador

Depto. Sistemas Informáticos y Computación. UPV

DOCENCIA VIRTUAL

Finalidad:
Prestación del servicio Público de educación superior (art. 1 LOU)

Responsable:
Universitat Politècnica de València.

Derechos de acceso, rectificación, supresión, portabilidad, limitación u oposición al tratamiento conforme a políticas de privacidad:
<http://www.upv.es/contenidos/DPD/>

Propiedad intelectual:
Uso exclusivo en el entorno de aula virtual.
Queda prohibida la difusión, distribución o divulgación de la grabación de las clases y particularmente su compartición en redes sociales o servicios dedicados a compartir apuntes.
La infracción de esta prohibición puede generar responsabilidad disciplinaria, administrativa o civil



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Objetivos

- ▶ Estudiar las diferentes técnicas de evaluación de la usabilidad
- ▶ Entender las características de cada tipo de evaluación, y ser capaz de seleccionar la técnica adecuada para cada proyecto
- ▶ Conocer cómo organizar una sesión de evaluación con usuarios
- ▶ Conocer los diferentes tipos de técnicas de evaluación experta
- ▶ Entender los tipos de evaluación que se realizan a lo largo del ciclo de vida de un sistema



Índice

- ▶ Introducción
- ▶ Tipos de evaluación
- ▶ Estudios con usuarios
- ▶ Evaluación experta
- ▶ Otros estudios
- ▶ Pruebas de aceptación
- ▶ Evaluación durante el uso



Introducción

- ▶ Realizar una evaluación externa de nuestros interfaces es imprescindible
 - Ayuda a entender la experiencia del usuario con el sistema y, cuando hay dificultades, encontrar formas de resolverlas
- ▶ Dependiendo del proyecto, se debería reservar una cantidad adecuada de recursos (5-20%)
- ▶ Se deben realizar pruebas de usabilidad a lo largo del proyecto, no únicamente al final
 - p.e., en las etapas iniciales se pueden presentar varios diseños en papel, para su evaluación por usuarios e incorporar sus sugerencias!



Introducción

- ▶ En sistemas de alta disponibilidad, se deben probar situaciones de alta carga de trabajo, e incluso de fallos parciales
 - y, a veces, no se puede hacer en laboratorio, sino en su contexto
- ▶ Existen múltiples metodologías de evaluación
 - dependiendo de tipo de aplicación, tipo de interfaz, tipos de usuarios...
 - ¿cuántos usuarios se usan en el test? ¿qué tipo de test se realiza? ¿expertos en evaluación o usuarios reales? ¿cómo se pueden evaluar todas las tareas de la aplicación?...

Introducción

- ▶ Recuerda que la usabilidad no es un concepto abstracto. Se puede medir y evaluar
- ▶ La evaluación debería medir el nivel en el que un producto se ajusta a los requisitos de usabilidad
 - Requisitos de usabilidad cualitativos
 - Características deseables. Pueden ser subjetivas y a veces difíciles de medir. “Los usuarios deberían poder comprar un artículo fácilmente y sin ayuda”
 - Requisitos de usabilidad cuantitativos/Métricas de usabilidad
 - Los requisitos se expresan numéricamente. “Cualquier página del sitio debería cargar en menos de 10s en una conexión 3G”, “4 de cada 5 usuarios noveles deben evaluar el producto como ‘fácil de usar’ o ‘muy fácil de usar’ en una escala de 5 niveles”



Introducción

- ▶ Niveles para evaluar las métricas de usabilidad
 - Actual
 - Caso mejor
 - Planeado
 - Caso peor
- ▶ Por ejemplo
 - Los usuarios necesitan 4 minutos para completar la tarea con el sistema existente (nivel actual). Idealmente, nos gustaría reducir dicho tiempo a la mitad (caso mejor). Probablemente 3 minutos es aceptable (planeado). Cualquier duración mayor de 3 minutos y 50 segundos no es aceptable (caso peor)

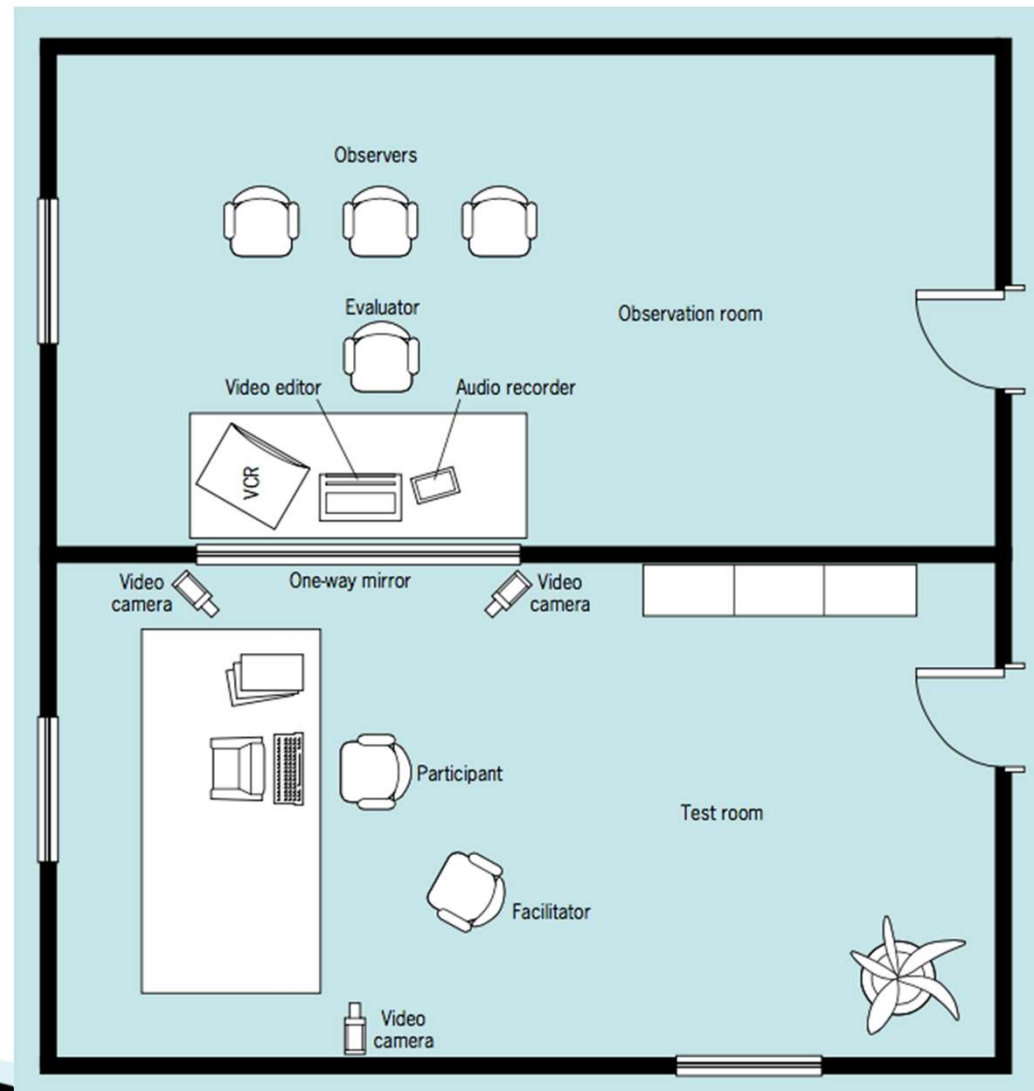
Tipos de evaluación

- ▶ Qué:
 - Encontrar tantos problemas de usabilidad como sea posible: evaluación diagnóstica
 - Medir el nivel de cumplimiento de los requisitos: evaluación por medidas
- ▶ Cuándo:
 - Durante el desarrollo: evaluación formativa
 - Al final: evaluación sumativa
- ▶ Cómo:
 - Temprana, con prototipos de baja fidelidad, informal: evaluación exploratoria
 - Al final, verificando el cumplimiento de los requisitos, experimento formal: evaluación de validación
 - Elegir una entre varias opciones, análisis estadístico: evaluación competitiva
- ▶ Quién:
 - Usuarios: Estudios con usuarios
 - Expertos: evaluación experta
 - Otros

Estudios con usuarios

- ▶ Selección de usuarios
 - ¿Cuántos? Normalmente 5 son suficientes
 - ¿Quiénes? Idealmente usuarios reales, un representante de un perfil de usuarios, un experto de usabilidad o un experto del dominio
- ▶ Dependiendo de dónde se realiza el estudio:
 - En el entorno del usuario: estudios de campo
 - En otros entornos: estudios controlados

Laboratorios de usabilidad



Laboratorios de usabilidad

- ▶ Laboratorio típico: espacio del participante



Espejo
unidireccional

Unidad de Investigación Acceso de la Universitat de València

Interfaces Persona Computador

Laboratorios de usabilidad

- ▶ Laboratorio típico: espacio de los observadores



<http://www.hq.nasa.gov/pao/portal/usability>

Laboratorios de usabilidad

Tipos de pruebas de usabilidad

- ▶ Pruebas de campo y laboratorios portátiles
 - Se lleva el laboratorio de usabilidad al lugar donde se usará el sistema final
 - Se deberá intentar capturar la mayor cantidad de información de cada ejecución
 - Otra opción es liberar versiones *beta* a un gran número de usuarios



noldus.com

Tipos de pruebas de usabilidad

▶ Pruebas de usabilidad remotas

- Un gran número de usuarios prueban el sistema desde el lugar donde usarán la aplicación
- Se puede buscar a los participantes en las bases de datos de clientes o en foros on-line
- Las pruebas pueden ser síncronas o asíncronas
- Ventajas: se puede acceder a un gran número de participantes, es barato, se prueban en el hardware del usuario
- Desventajas: el control sobre el usuario es menor y se pierden sus reacciones
- Algunos estudios dicen que este tipo de pruebas encuentran más problemas que las técnicas tradicionales

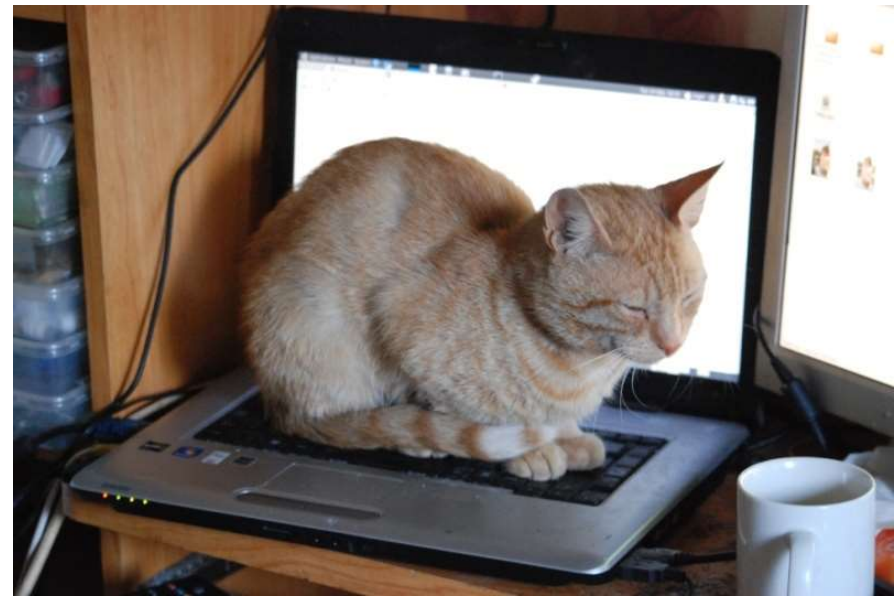
Tipos de pruebas de usabilidad

► Prueba-a-romperlo

- Originario de las pruebas de videojuegos, se le pide al usuario que intente “romper” el sistema
- Este tipo de pruebas de estrés permite robustecer las aplicaciones



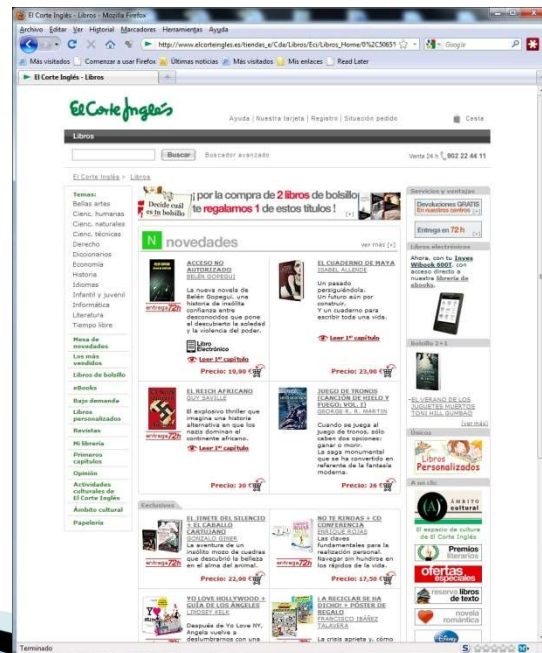
[flickr.com/photos/quazie/](https://www.flickr.com/photos/quazie/)



[flickr.com/photos/susan_g/](https://www.flickr.com/photos/susan_g/)

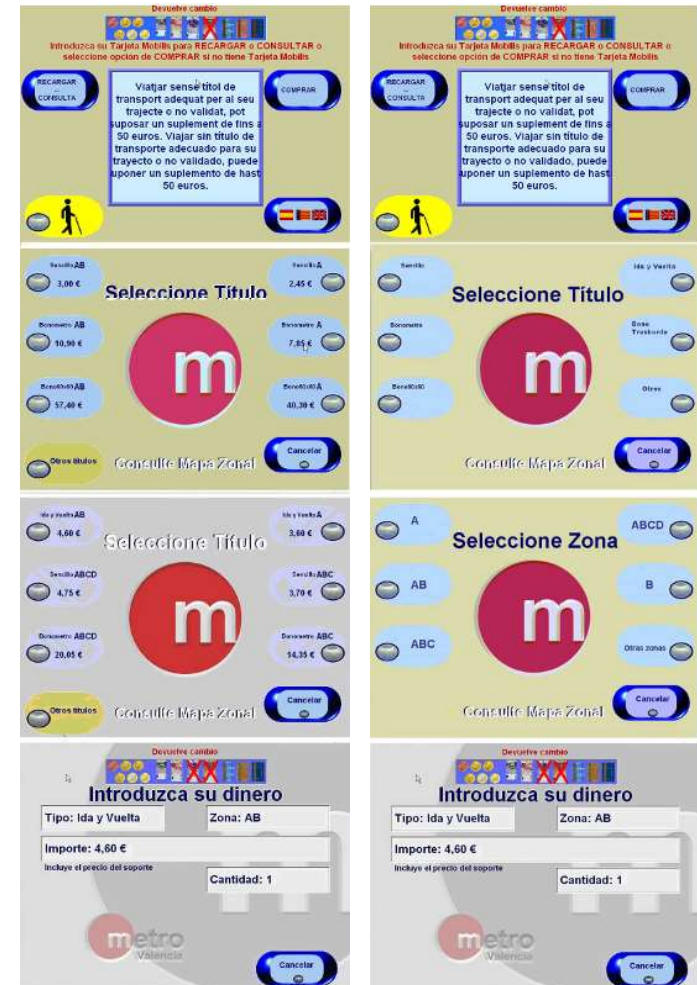
Tipos de pruebas de usabilidad

- ▶ Pruebas de usabilidad competitivas
 - Compara la nueva interfaz con una versión anterior, o con una interfaz de la competencia
 - Se compara el tiempo en realizar una tarea o la tasa de fallos en un sistema y en otro



Tipos de pruebas de usabilidad

- ▶ Pruebas de usabilidad competitivas
 - Metrovalencia:
http://www.metrovalencia.es/word/press/?page_id=298
 - Resultados Consulta Mayo 2012
 - *Pregunta 4: Si usted tuviera que comprar un billete de ida y vuelta de la zona AB, de las dos opciones (A y B) que le mostramos a continuación, ¿cuál es para usted la mejor forma de adquirir dicho billete? ¿Por qué?*



Opción A

Opción B

Tipos de pruebas de usabilidad

► Pruebas de usabilidad competitivas

Opción A



Opción B



Preparando el experimento

- ▶ Se debe realizar un plan detallado del experimento, incluyendo:
 - Qué se desea medir
 - Cantidad, tipos y fuentes de los participantes
 - Duración (30-90 min) y tareas que deben realizar los participantes
 - El contenido de los cuestionarios y entrevistas que se realizarán a los participantes
- ▶ Realizar una prueba piloto con un número reducido de participantes (1-3)
 - Nos aseguramos de que todos los detalles de la evaluación funcionan
 - Realizar la prueba en el mismo lugar y tan parecido a como se realizará el estudio real como sea posible
 - Analizar e interpretar los datos, para asegurarnos de que se ha recogido toda la información necesaria

Estructura de una sesión

- ▶ Bienvenida
- ▶ Firma del consentimiento y quizá un acuerdo de confidencialidad
- ▶ Encuesta filtro de selección
- ▶ Usar los escenarios de tarea (adaptados al lenguaje de los usuarios) para describir los pasos que debe seguir el usuario
 - Seleccionar los más importantes
- ▶ Entrevista después de la sesión
 - Revisar las grabaciones y preguntar al usuario sobre sus intenciones y acciones
 - Encuestas
- ▶ Incentivo

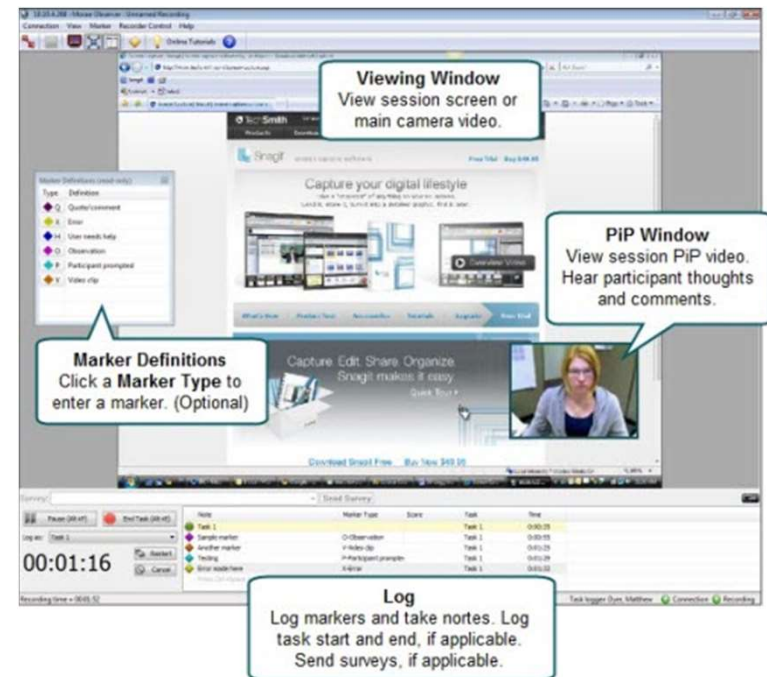
Estructura de una sesión

- ▶ Bienvenida
 - Los participantes deben ser tratados con respeto, y hacerles saber que no se les está evaluando
 - También deben ser informados de:
 - El objetivo del estudio
 - El procedimiento a seguir y la duración. Cómo pedir una pausa
 - Quién revisará las grabaciones y su uso después del estudio
 - Un compromiso de confidencialidad y anonimato
 - Un aviso de los peligros de la prueba
 - La posibilidad de abandonar la prueba sin penalización
 - Un modo de contacto posterior
- ▶ Filtro de selección: nos aseguramos de que el participante se ajusta al perfil buscado
 - su experiencia con los ordenadores, o con la tarea a realizar, motivación, nivel educativo, capacidad de comprensión
 - su agudeza visual, su mano dominante, edad, género

Durante la prueba

Midiendo tiempos

- ▶ Usa un cronómetro
- ▶ Acompaña cada comentario escrito durante la sesión con una marca de tiempo
 - O usa un programa de captura de pulsaciones de teclado para capturar toda la interacción del usuario. Problema: relacionar los comentarios escritos con cada evento
 - O usa programas especializados en evaluaciones de usabilidad (Ovo Solo, Noldus The Observer XT, TechSmith Morae...)



techsmith.com

Durante la prueba

- ▶ Una técnica de prueba de usabilidad consiste en pedir al participante que exprese en voz alta su opinión
 - El personal debe facilitar esta comunicación y pedir pistas sobre qué está pensando el participante. No debe ayudarlo
 - <http://www.youtube.com/watch?v=QckIzHC99Xc>
 - Ventajas
 - Realimentación inmediata de las opiniones de los participantes sobre la interfaz y sobre cualquier problema o sorpresa
 - Puede ayudar a los usuarios a concentrarse durante la sesión de evaluación
 - Útil para recoger datos cualitativos
 - Desventajas
 - Algunos participantes lo encuentran forzado y les distrae
 - Pensar en voz alta puede frenar los procesos mentales del participante, y reducir artificialmente su desempeño y tasas de error
 - Puede ser extenuante para el usuario si dura demasiado

Durante la prueba

- ▶ Protocolo retrospectivo: preguntar a los usuarios sobre sus intenciones al realizar cada acción después de la prueba
 - Ventajas
 - No interfiere con las medidas de desempeño
 - Útil para recoger datos cuantitativos
 - Inconvenientes:
 - El participante puede haber olvidado la razón por la que hizo cierta acción
 - Algunos participantes pueden sentirse intimidados por las cámaras

Grabaciones

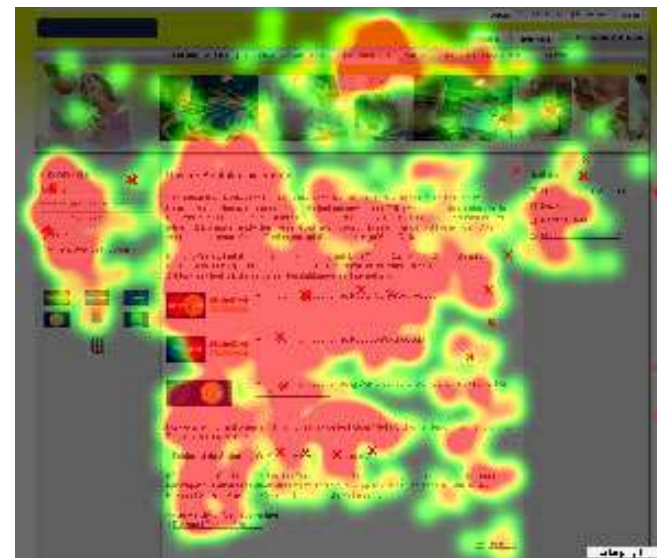
- ▶ Normalmente se graba a los participantes, para poder revisar posteriormente sus reacciones, errores, forma de trabajar, etc.
- ▶ Los métodos de tracking visual permiten calcular las zonas de la ventana que más atención atraen



cure.at



tobii.com



cure.at

Encuestas

- ▶ Son un método familiar y barato para capturar la opinión de un usuario o un experto
- ▶ Permite sondear la opinión de miles de usuarios de una forma estructurada
- ▶ Las claves para realizar encuestas útiles es tener claro sus objetivos y diseñar preguntas directas
- ▶ Pueden ser en papel u on-line

ENCUESTA DE OPINIÓN DEL ALUMNADO SOBRE LA ACTUACIÓN DOCENTE DEL PROFESORADO
ENQUESTA D'OPINIÓ DE L'ALUMNAT SOBRE L'ACTUACIÓ DOCENT DEL PROFESSORAT
STUDENT QUESTIONNAIRE FOR THE ASSESSMENT OF TEACHERS' ACADEMIC PERFORMANCE
CURSO/CURS/ ACADEMIC YEAR 2009-2010

DATOS IDENTIFICATIVOS / DADES IDENTIFICATIVES / IDENTIFYING INFORMATION

Professor/a _____
 Teacher _____
 Assigatura _____
 Subject _____

Grupa _____
 Grup _____
 Group _____

CONTESTA CON SINCERIDAD, SI NO TIENES SUFICIENTE INFORMACIÓN, NO OPINES
CONTESTA AMB SINCERITAT, SI NO TENS PROU D'INFORMACIÓ, NO OPINES
PLEASE ANSWER THE QUESTIONNAIRE HONESTLY. IF YOU DO NOT HAVE SUFFICIENT INFORMATION, PLEASE DO NOT ANSWER THE QUESTION

INSTRUCCIONES BÁSICAS PARA LA COMPLEMENTACIÓN DE LA ENCUESTA
INSTRUCCIONS BÀSIQUES PER A LA COMPLEMENTACIÓ DE L'ENQUESTA
QUESTIONNAIRE INSTRUCTIONS

Para responder a las preguntas del cuestionario deberá realizar una marca nítida en forma de x en el espacio (1). Recuerde que sólo puede hacer una marca por pregunta. Si se equivocara debe completamente el casillero (2) y marcar la respuesta correcta con un x en el (1).
 Per a respondre a les preguntes del qüestionari has de realitzar una marca nítida en forma d'x en l'espai (1). Recorde que nom cal fer una marca per pregunta. Si s'equivoca hauria de completarment el casiller (2) i marcar la resposta correcta amb un x en el (1).
 Please clearly cross the appropriate box (1). Please select only one option per question. If you make a mistake, please cross out the whole box (2) and then put a cross in the correct answer (1).

CÓDIGOS DEL ALUMNO / CODIS DE L'ALUMNE / STUDENT CODE

SEXO / SEXE / SEX _____
 HOMBRE / HOME / MALE _____
 MUJER / DONA / FEMALE _____

CONVOCATORIA EN LA ASIGNATURA _____
 CONVOCATORIA EN L'ASSIGNATURA _____
 NUMBER OF TIMES YOU HAVE TAKEN THIS SUBJECT _____

ESCALA DE VALORACIÓN / ESCALA DE VALORACIÓ / RESPONSE SCALE

1	2	3	4	5	6
TOTALMENTE EN DESACUERDO	MÁS BIEN EN DESACUERDO	TÉRMINO MEDIO	MÁS BIEN DE ACUERDO	TOTALMENTE DE ACUERDO	NO DISPONGO DE ELEMENTOS DE JUICIO PARA OPINAR
STRONGLY DISAGREE	DISAGREE	NEITHER AGREE OR DISAGREE	AGREE	STRONGLY AGREE	NO OPINION
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6
1	2	3	4	5	6

1 Parece dominar la materia que imparte
 Sembla dominar la matèria que imparteix
 The teacher has a good command of the subject being taught

2 Resuelve dudas y cuestiones sobre la materia con claridad y precisión
 Resol dubtes i qüestions sobre la matèria amb claredat i precisió
 The teacher answers questions about the subject clearly and precisely

3 Al inicio del curso proporciona información clara sobre la asignatura: objetivos, programa, metodología y criterios de evaluación
 A l'inici del curs proporciona informació clara sobre l'assignatura: objectius, programa, metodologia i criteris d'avaluació
 The teacher provides relevant information about the subject at the beginning of the course: objectives, contents, methodology, and assessment criteria

4 Tiene una buena programación del ritmo de las clases y del tiempo dedicado a cada tema
 Té una bona programació del ritme de les classes i del temps dedicat a cada tema
 The teaching pace and the time allocated to each topic are adequate

5 La metodología empleada y las actividades realizadas en la asignatura ayudan a aprender al alumnado
 La metodologia emprada i les activitats realitzades en l'assignatura ajuden l'alumnat a aprendre
 The methodology adopted and the activities carried out contribute to student learning

6 Los materiales utilizados y/o recomendados (bibliografía, documentos, recursos didácticos, etc.) son de gran ayuda para el aprendizaje de la asignatura
 Els materials utilitzats i/o recomanats (bibliografia, documents, recursos didàctics, etc.) són de gran ajuda per a l'aprenentatge de l'assignatura
 The materials used and/or recommended (bibliography, documentation, teaching resources, etc.) are helpful for learning this subject

7 Contribuye a crear un buen clima de trabajo y anima a los alumnos a participar en las clases
 Contribueix a crear un bon clima de treball i anima els alumnes a participar en les classes
 The teacher contributes to a positive working environment and encourages students to actively participate in the classroom

8 Su manera de explicar consigue motivar al alumnado y despertar el interés por la asignatura
 La seua manera d'explicar aconsegueix motivar l'alumnat i despertar l'interès per l'assignatura
 His/her teaching methodology motivates students and awakens their interest in the subject

9 Teniendo en cuenta las limitaciones, pienso que el profesor que imparte esta asignatura debe considerarse un buen profesor
 Tintant en compte les limitacions, penso que el professor que imparteix aquesta assignatura ha de considerar-se un bon professor
 All things considered, I think that the teacher of this subject should be regarded as a good teacher

Documento aprobado por el Consejo de Gobierno de la Universidad Politécnica de Valencia el 8 de noviembre de 2007
 Documente aprovat pel Consell de Govern de la Universitat Politècnica de València el 8 de novembre de 2007
 This questionnaire was approved by the Governing Council of the Universidad Politécnica de Valencia on 8th November 2007

Gracias por su colaboración
 Gràcies per la vostra col·laboració
 Thank you for your co-operation

Encuestas

▶ Ventajas

- Es más fácil no olvidar hacer alguna pregunta
- Es más fácil comparar las respuestas de distintos participantes, ya que todos ellos ven exactamente la misma pregunta
- Se pueden recoger datos cuantitativos como “3 de los 4 participantes dijeron que era fácil de navegar por la interfaz”
- Se puede demostrar el progreso durante el desarrollo cuando el resultado de las encuestas mejora

▶ Inconvenientes

- Es difícil diseñar una buena encuesta
- Las preguntas cerradas son más fáciles de analizar, pero no dan la razón por la que el usuario eligió la respuesta

Encuestas

Tipos de preguntas

- ▶ Escala Likert. El usuario debe mostrar su nivel de acuerdo con una afirmación: totalmente de acuerdo, de acuerdo, neutral, en desacuerdo, totalmente en desacuerdo
 - Preguntas:
 - Mejora mi rendimiento en la búsqueda y compra de libros
 - Me permite buscar y comprar libros más rápido
 - Hace más fácil la compra y búsqueda de libros
 - Aumenta mi productividad buscando y comprando libros
 - Normalmente usa 5 niveles, pero 7 o 9 también son posibles (la escala siempre es simétrica)

Encuestas

Tipos de preguntas

- ▶ Encuesta binaria sobre las reacciones del usuario al usar un procesador de texto:
 - Agradable o irritante
 - Simple o complejo
 - Conciso o redundante...
- ▶ Otra opción, que gradúa las reacciones:

◦ Hostil	1 2 3 4 5 6 7	Amigable
◦ Ambiguo	1 2 3 4 5 6 7	Específico
◦ Induce a error	1 2 3 4 5 6 7	Beneficioso
◦ Desalentador	1 2 3 4 5 6 7	Alentador

Encuestas

- ▶ Hay muchas encuestas disponibles ya validadas:
 - QUIS Questionnaire for User Interaction Satisfaction
 - SUS System Usability Scale
 - CSUQ Computer System Usability Questionnaire
 - WAMMI Website Analysis and MeasureMent Inventory
 - ...
 - Más información sobre encuestas en: <http://garyperlman.com/quest>

Encuestas

Cuestionarios tipo

- ▶ *QUIS: Questionnaire for User Interaction Satisfaction*
(<http://lap.umd.edu/quis>)
 - Fue diseñado para evaluar la satisfacción subjetiva del usuario con respecto a aspectos específicos de la interfaz
 - Estudia nueve factores específicos de la interfaz:
 - factores de pantalla, terminología y feedback del sistema, factores de aprendizaje, capacidades del sistema, manuales técnicos, tutoriales on-line, multimedia, teleconferencia e instalación del software
 - Usa una escala de nueve puntos
 - El cuestionario se debe ajustar a las particularidades de cada interfaz

Cuestionarios tipo

- PARTE 1: Experiencia con el sistema

- PARTE 6: Aprendizaje

difícil fácil

1 2 3 4 5 6 7 8 9 NA

difícil fácil

1 2 3 4 5 6 7 8 9 NA

difícil fácil

lento rápido

a 1 2 3 4 5 6 7 8 9 NA

Encuestas

Cuestionarios tipo

- ▶ *System Usability Scale (SUS)*
 - Más sencillo que el anterior
 - 10 afirmaciones, en las que el usuario debe valorar en una escala de 5 puntos
 - La mitad de las preguntas están formuladas afirmativamente, y la otra mitad negativamente

Brooke, John. SUS: A quick and dirty usability scale. Usability Evaluation in Industry, 1996

System Usability Scale

© Digital Equipment Corporation, 1986.

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Strongly disagree

Strongly agree

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

1	2	3	4	5

Encuestas

Cuestionarios tipo

► *Computer System Usability Questionnaire (CSUQ)*

- 19 afirmaciones, en las que el usuario debe valorar su coincidencia en una escala de 7

J. R. Lewis. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. Technical Report 54.786

1. Overall, I am satisfied with how easy it is to use this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE
---------------------------	----------	----------	----------	----------	----------	----------	----------	------------------------------

COMMENTS:

2. It is simple to use this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE
---------------------------	----------	----------	----------	----------	----------	----------	----------	------------------------------

COMMENTS:

3. I can effectively complete my work using this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE
---------------------------	----------	----------	----------	----------	----------	----------	----------	------------------------------

COMMENTS:

4. I am able to complete my work quickly using this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE
---------------------------	----------	----------	----------	----------	----------	----------	----------	------------------------------

COMMENTS:

5. I am able to efficiently complete my work using this system.

STRONGLY AGREE	1	2	3	4	5	6	7	STRONGLY DISAGREE
---------------------------	----------	----------	----------	----------	----------	----------	----------	------------------------------

COMMENTS:

Encuestas

Cuestionarios tipo

- ▶ *Website Analysis and MeasureMent Inventory (WAMMI)*
 - Servicio comercial de realización de cuestionarios on-line
 - 20 afirmaciones, en las que el usuario debe evaluar su coincidencia en una escala de 5 puntos
 - Dan el soporte web para la realización de los cuestionarios y luego generan un informe de los resultados

<http://www.wammi.com/samples/index.html>

Website Analysis and Measurement Inventory (WAMMI)

Demo Survey - Company Z



Thank you for helping us evaluate the Company Z web site. If you have not yet used this site, please go back to it now and fill out this questionnaire after you've used it.

The information you provide is kept completely confidential, and no information is stored on computer media that could identify you as a person. You are not in any way obliged to participate and you may freely withdraw at any time.

What is your age?

choose ...

What is your gender?

- ☐ Male
☐ Female

Which of these browsers do you have experience with?(select all that apply)

- ☐ Safari
☐ Google Chrome
☐ Opera Browser
☐ Mozilla Firefox
☐ Internet Explorer
☐ something else

What do you normally visit the web site for?

Statements 1 - 10 of 20

	Strongly Agree	Strongly Disagree
This web site has much that is of interest to me.	<input type="radio"/>	<input type="radio"/>
It is difficult to move around this web site.	<input type="radio"/>	<input type="radio"/>
I can quickly find what I want on this web site.	<input type="radio"/>	<input type="radio"/>
This web site seems logical to me.	<input type="radio"/>	<input type="radio"/>

Análisis de los resultados

- ▶ Después de la sesión, habrás recogido:
 - Datos personales sobre los participantes, notas de los observadores, grabaciones de audio o vídeo, formularios de captura de datos, datos cuantitativos sobre tiempos, errores y otras métricas de usabilidad, datos cuantitativos (y cualitativos) de las encuestas realizadas antes y después de la prueba, protocolos retrospectivos, una lista de problemas de usabilidad encontrados...
- ▶ Un defecto de usabilidad es un problema en una interfaz de usuario que puede llevar al usuario a confusión, error, retraso o incapacidad de completar una tarea

Resumiendo datos cuantitativos

- ▶ Varias opciones:
 - Tablas, gráficas, y rankings
 - Estadísticas descriptivas (media, mediana, moda)
 - Cuidado: supón que un usuario ha necesitado 5 segundos para completar una tarea, otro necesitó 6 segundos y un tercero 55 segundos
 - Media: 22 segundos
 - Estadísticas inferenciales (pruebas de significancia estadística)
 - Necesita muchos participantes

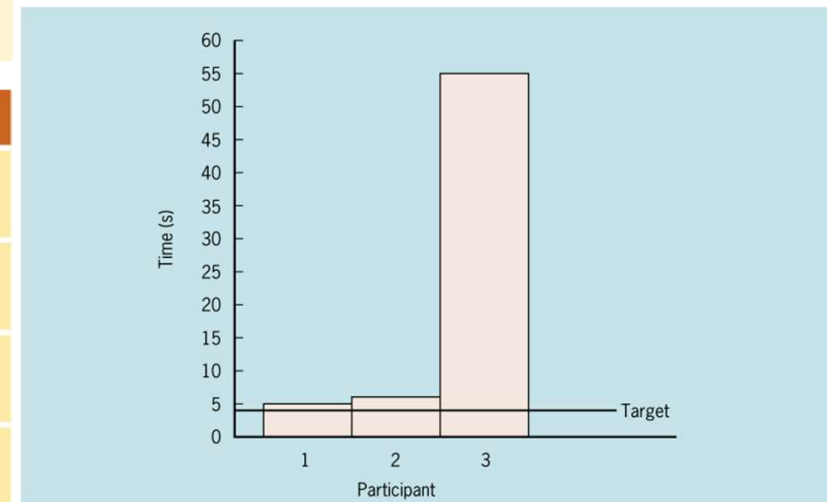
Resumiendo datos cuantitativos

► Tablas y representaciones gráficas

Example of Task Times

The task time for participant one (experienced) was five seconds. Participant two (experienced) took six seconds. The third participant was a novice and took 55 seconds. The target for task time is four seconds.

Participant	Experience level	Task time
Participant 1	Experienced	5 seconds
Participant 2	Experienced	6 seconds
Participant 3	Novice	55 seconds
Target		4 seconds



Recomendaciones

- ▶ Después de interpretar los resultados, se pueden realizar una serie de recomendaciones como:
 - Puntos fuertes de la interfaz
 - Defectos a solucionar
 - Posibles defectos o puntos fuertes, pero sin pruebas suficientes (hacen falta más pruebas)
 - Áreas de la interfaz que no se han probado
 - Cambios a los requisitos de usabilidad u otros requisitos

Evaluación experta

- ▶ Normalmente se realizan evaluaciones informales pidiendo opinión a colegas o clientes.
- ▶ Aunque una técnica más efectiva consiste en consultar a evaluadores expertos
 - internos o externos
- ▶ El resultado de una evaluación experta puede ser:
 - un informe formal con los problemas encontrados y recomendaciones
 - una presentación y discusión con los responsables o los diseñadores

Evaluación experta

▶ Ventajas:

- Los resultados están disponibles más rápidamente
- Más económicos que los estudios con usuarios
- Los expertos pueden sugerir soluciones a los problemas encontrados
- Ayudan a encontrar los errores obvios más fácilmente, para que no lleguen a las pruebas con usuarios

▶ Inconvenientes

- Los expertos no son usuarios reales, y sus predicciones sobre lo que harán los usuarios con la interfaz o sobre la importancia de un defecto pueden ser incorrectas
- Los expertos tienen sus propias preferencias sobre el diseño de interfaces de usuario, lo que puede sesgar la evaluación
- La evaluación experta depende mucho de la experiencia del evaluador

Evaluación experta

- ▶ El evaluador debería:
 - Replicar las condiciones del usuario final (recibir cursos de manejo de la aplicación, leer manuales, y usarla en un entorno parecido)
 - Tener tacto en sus recomendaciones,
 - Comprender que es difícil en tan poco tiempo comprender todas las decisiones de diseño y la historia del desarrollo
 - Tener experiencia en el tipo de aplicación evaluada
 - Dejar el desarrollo de las soluciones a los diseñadores
 - Ser exhaustivo en el informe
 - Revisar la consistencia en todas las ventanas de la aplicación

Evaluación experta

Tipos de evaluación

- ▶ Evaluación heurística
 - El experto evalúa si la interfaz sigue la lista de heurísticas de diseño (p.e., los principios de Nielsen)
- ▶ Revisión de las guías de diseño
 - El experto revisa que se hayan seguido las indicaciones de las guías de diseño
- ▶ Revisión de estándares
 - El experto comprueba que el diseño sigue un estándar (p.e., ISO 9241)
- ▶ Inspección de consistencia
 - Comprueba la terminología, fuentes, paletas, formatos de E/S, etc. tanto en las interfaces como en la documentación

Evaluación experta

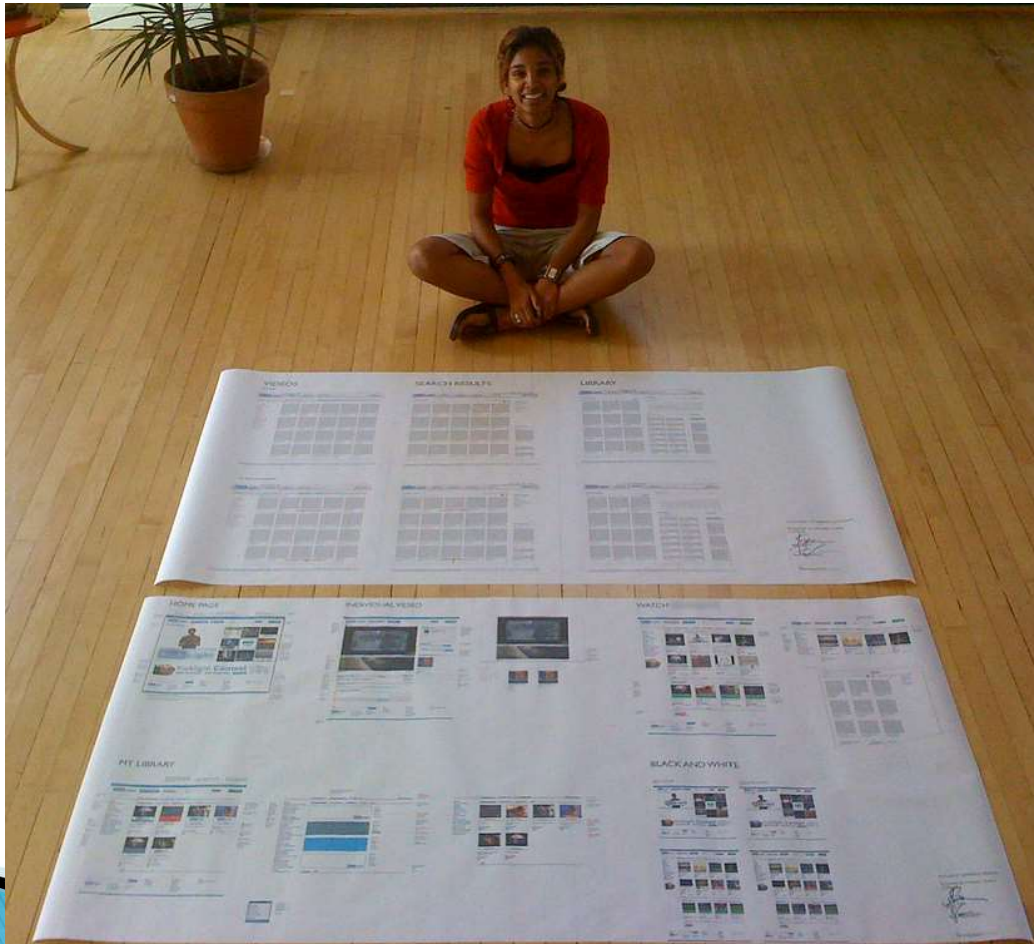
Informe

- ▶ Con respecto al informe, se debería:
 - Usar guías de diseño para estructurar el informe,
 - Separar los problemas en función del tipo de usuario al que afecta
 - Ordenar las recomendaciones por importancia
 - Describir cada recomendación a nivel conceptual
 - Ser consciente de las restricciones técnicas o de negocio
 - Resolver el problema completo, no sólo un caso especial
 - Dar recomendaciones específicas y claras, con ejemplos
 - Incluir pequeños detalles como faltas ortográficas, fallos de alineamiento de controles, inconsistencia en la posición de los controles, etc.

Evaluación experta

Otras técnicas

- ▶ Ojo de pájaro



[flickr.com/photos/juhansonin/](https://www.flickr.com/photos/juhansonin/)

Otros tipos de evaluación

- ▶ Entrevistas de grupos (*focus groups*)
 - Más o menos 8 personas dirigidas por un moderador discuten sobre algo
 - Se usa a menudo en marketing
 - Útil cuando aún no existe la interfaz
- ▶ Ordenación de tarjetas (*card sorting*)
 - Se pide a los usuarios que agrupen elementos
 - Útil para identificar categorías y estructuras subyacentes
 - Imprimir cada elemento en una tarjeta y pedir a los usuarios que las agrupen, y posiblemente que le den un nombre al grupo
 - Estudiar los resultados, tanto:
 - donde hay acuerdo general,
 - donde no hay acuerdo, y estudiar cómo resolverlo (quizá renombrar el elemento, o poner el elemento en varias categorías)

Otros tipos de evaluación

- ▶ Herramientas de validación automática
 - Validadores de accesibilidad y HTML
 - W3C Web Accessibility Initiative list: <http://www.w3.org/WAI/ER/tools/>
 - W3C CSS Validation Service: <http://jigsaw.w3.org/css-validator/>
 - SSB BART Accessibility Management Platform (AMP)
<http://www.ssbartgroup.com/amp>
 - ...

Pruebas de aceptación

- ▶ Son pruebas que realiza un cliente sobre el sistema para comprobar que cumple los requisitos
 - Por ejemplo, estableciendo unos casos de uso, y el tiempo medio de respuesta del conjunto hardware-software
- ▶ Para establecer requisitos de las interfaces de usuario, hay que usar criterios medibles
 - Tiempo necesario para aprender a usar ciertas funciones
 - Velocidad de completado de una tarea
 - Tasa de error de los usuarios
 - Retención de las órdenes a lo largo del tiempo
 - Satisfacción subjetiva del usuario

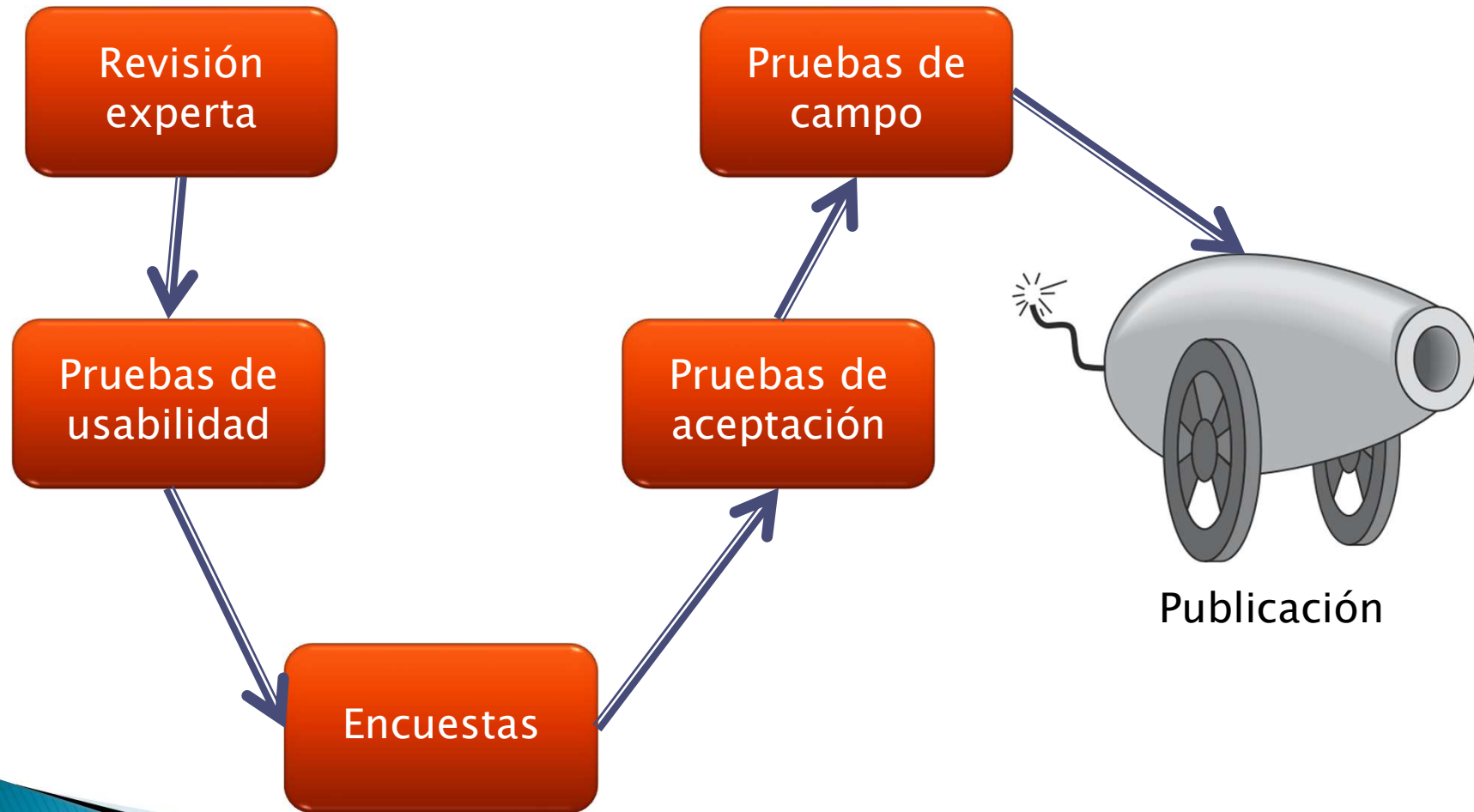
Pruebas de aceptación

- ▶ Ejemplo de caso de prueba: web de compra de alimentos
 - Los participantes serán 35 adultos (entre 25 y 45 años), nativos sin discapacidad, contratados a través de una agencia de empleo. Tendrán una experiencia moderada en el uso de la web: de 1 a 5 horas semanales por lo menos durante un año. Recibirán una demostración de 5 minutos sobre los características básicas. Al menos 30 de los 35 adultos deberían ser capaces de completar el banco de pruebas en menos de 30 minutos.

Pruebas de aceptación

- ▶ Otros elementos medibles:
 - Comprensión de la salida del sistema, tiempo de respuesta del sistema, procedimiento de instalación, documentación impresa, atractivo visual, etc.
- ▶ Unos criterios de aceptación precisos ahorra posteriores discusiones y ofrece una prueba objetiva de cumplimiento de contrato
- ▶ Se deberían realizar por una parte neutral
- ▶ Después de las pruebas de validación, aún se deberán realizar pruebas de campo antes de publicar el sistema

Pruebas antes de la publicación



Evaluación durante el uso

- ▶ Una vez que se ha publicado el sistema se debe hacer un seguimiento del mismo para mejorarlo
- ▶ Hay que minimizar los cambios en la interfaz para evitar molestar a los usuarios
- ▶ Herramientas
 - Entrevistas y discusiones en grupo
 - Captura continua de datos de rendimiento del usuario
 - Consultas online o telefónicas, buzones de sugerencias
 - Grupos de discusión, *wikis* y grupos de noticias
 - Herramientas de evaluación automática

Bibliografía

- ▶ D. Stone, C. Jarrett, M. Woodroffe. User Interface Design and Evaluation. Morgan Kaufmann, 2005.
 - Chapters 20-27
- ▶ Shneiderman, B. y Plaisant, C. Designing the User Interface. Pearson 5th ed., 2010
 - Capítulo 4

