

Gestión Calidad Software



GRADO II
ITINERARIO ING. SW
OPT. 4º CURSO 2º C.

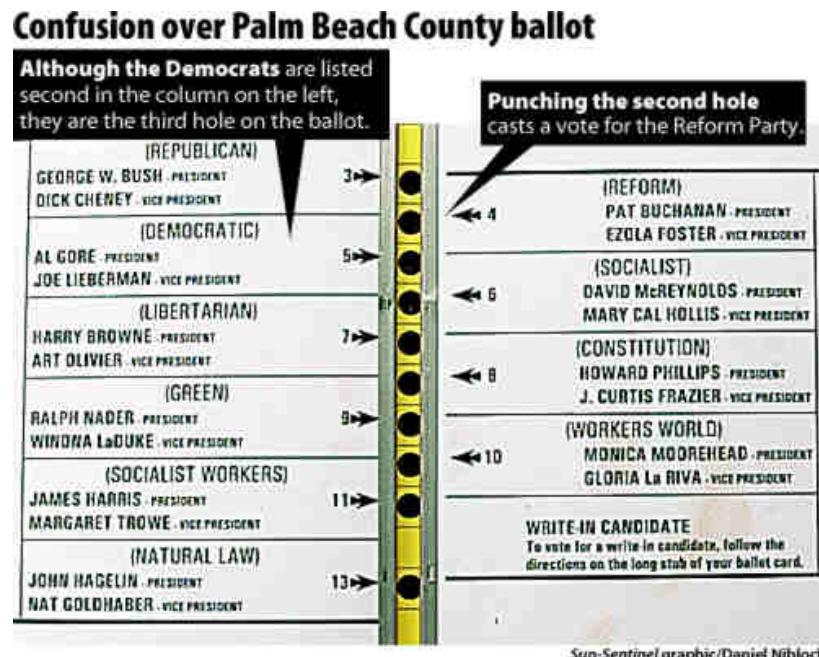
➤ ¿Por qué una fase de UT?

El propósito real del método científico es asegurarse de que la Naturaleza no te ha engañado al hacerte pensar que sabes algo que en realidad no sabes

Robert Pirsing, Zen and the Art of Motorcycle Maintenance

- Necesitamos asegurarnos de que hemos conseguido los objetivos de eficiencia, efectividad y satisfacción que nos pusimos (o nos pusieron) en un ppi.
- Para ello necesitamos disponer de técnicas efectivas

- ¿Por qué una fase de UT?
 - ¿Os acordáis de la famosa papeleta de las elecciones de estados unidos?
 - <http://www.asktog.com/columns/042ButterflyBallot.html>



- ¿Por qué UT?

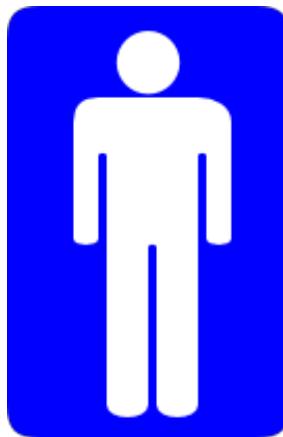
- Incluso siguiendo guías cuidadosas de diseño, existe el peligro de haber SOBREGENERALIZADO
 - Haber usado guías que se probaron exitosas (por mí o por otros) en un contexto y con un tipo de audiencia y haberlas aplicado aquí pensando que tb serían válidas, cuando en realidad no lo son
 - Ejemplo: Usborne (2005) describe un A/B/C test de tres versiones diferentes de una página que explica el contenido de un informe e incluye un formulario para llenar y comprar el informe.
 - Página A: página actual
 - Página B: mismo diseño pero cambios en el texto
 - Página C: rediseño que suponía cambiar de un diseño en una sola columna a un diseño en doble columna que permitía mostrar más contenido 'por encima del punto de scrolling'
 - Métrica: número de ventas
 - Página A: 244 ventas
 - Página B: 282 ventas
 - Página C: 114 ventas (-53,3%!!!)

¿Resultados?



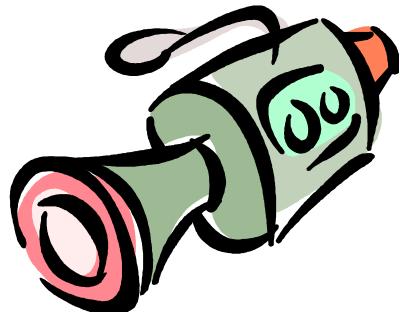
- ¿Por qué UT?
 - De hecho, muchas heurísticas/guías/patrones de diseño son dependientes de contexto, y evolucionan con el tiempo / uso. Algo cierto hace 10 años puede no serlo ahora, debido e.g. a cambios en el modelo mental de nuestros usuarios.
 - ¿Se os ocurre algún ejemplo?

- Cualquier tipo de asunción puede ser peligrosa



- <http://hystericalmarissa.blogspot.com/2010/08/go-where-sex-gender-and-toilets.html>
 - *The real purpose of the scientific method is to make sure that Nature has not deceived you by making you think that you know something that in fact you do not know.*
 - Robert Pirsing, Zen and the Art of Motorcycle Maintenance

- Nadie es immune a los desastres de usabilidad!
 - Vídeo: Microsoft Kin y Kin2



<http://www.elmundo.es/blogs/elmundo/el-gadgetoblog/2012/11/23/el-vergonzoso-error-de-kin.html>



**¿Qué cosas podemos
testear?**



- **Métodos Formativos/Sumativos**
- Métodos Cuantitativos/Cualitativos
- Métodos sobre Diseños Tempranos/Avanzados
- Métodos sin Usuarios/Métodos con Usuarios

Técnicas: Formativo vs Sumativo



¿Diferencias?



Técnicas: Formativo vs Sumativo

Testeo Formativo (PROBING)	Testeo Sumativo (IMPROVING)
Se realiza durante el proceso de desarrollo del sitio o producto	Se realiza al final del desarrollo
Se utiliza para determinar si un objetivo específico de diseño se está cumpliendo o encontrar errores	Se utiliza para validar que un producto ha alcanzado las metas propuestas (cobertura objetivos ROI)
No utiliza métodos estadísticos	Utiliza métodos estadísticos
<i>E.g. ¿Puede un usuario encontrar y/o usar el botón ‘calculadora de hipoteca’ para introducir de manera adecuada los datos del préstamo en la aplicación?</i>	<i>E.g. ¿Ha mejorado el nuevo formulario el ratio de conversión en la medida prevista?</i>

■ Tipos tests sumativos:

- **Tests de benchmark (comparación en distintos instantes de tiempo)**: permiten describir cómo de usable es nuestra aplicación con respecto a una serie de objetivos (e.g. medidas que informes internacionales dicen que son la media del sector). Permiten determinar tanto qué arreglar en la interfaz como proporcionar un conjunto de datos base para la comparación tras las sucesivas iteraciones.
- **Tests comparativos (comparación en el mismo instante de tiempo)**: Involucra más de una aplicación (e.g. La nuestra y la de nuestros competidores, o la versión actual y una versión previa, o varias versiones alternativas) que se testan a la vez.
 - **Diseño intra-sujeto**: los mismos sujetos realizan las mismas acciones (o comparables) en todos los productos
 - **Diseño inter-sujeto**: diferentes usuarios testan cada aplicación. Requieren más usuarios para detectar diferencias significativas.

- Métodos Formativos/Sumativos
- **Métodos Cuantitativos/Cualitativos**
- Métodos sobre Diseños Tempranos/Avanzados
- Métodos sin Usuarios/Métodos con Usuarios

Técnicas: Cuantitativos vs Cualitativos



¿Diferencias?



?

O

?



Técnicas: Cuantitativos vs Cualitativos

Cualitativo (pequeña escala)	Cuantitativo (gran escala)
Proporciona muchas oportunidades para observar y entender por qué ocurren las cosas	Su foco está en las medidas cuantitativas, no en observaciones individuales
Se basa en la experiencia y el juicio. <i>El especialista observa, interpreta, decide</i>	Se basa en el análisis de datos. <i>El especialista es neutro</i>
Los resultados no son generalizables	Los resultados se pueden generalizar a una población mayor
<i>Puede utilizar tamaños muestrales pequeños</i>	<i>Requiere grandes tamaños de muestra</i>



¿Cuántos usuarios hacen falta para un test cuantitativo? ¿Y para uno cualitativo?

- Testeo Cualitativo: ideal para iteraciones en la fase de diseño de la aplicación
 - Requiere pocos usuarios.
 - Normalmente usado para testeo formativo: entre 6 y 8 usuarios en cada iteración. Si hay varios grupos de usuario, mínimo 4 usuarios por grupo.
 - Testear con cinco usuarios tres veces es mejor que testear una vez con quince!! ☺
 - Algunos consejos:
 - Controla las expectativas de los managers: no se puede generalizar, es exploratorio. Por tanto, cuando presentes los resultados, no presentes estadísticas; sólo observaciones y recomendaciones
 - Pide al usuario que haga tareas realistas
 - **Utiliza buenas técnicas de muestreo (elige a los usuarios correctos para testar)**
 - No tomes ninguna medida cuantitativa: sólo busca errores

- **Testeo Cuantitativo: ideal para validar el resultado del proceso de diseño.**

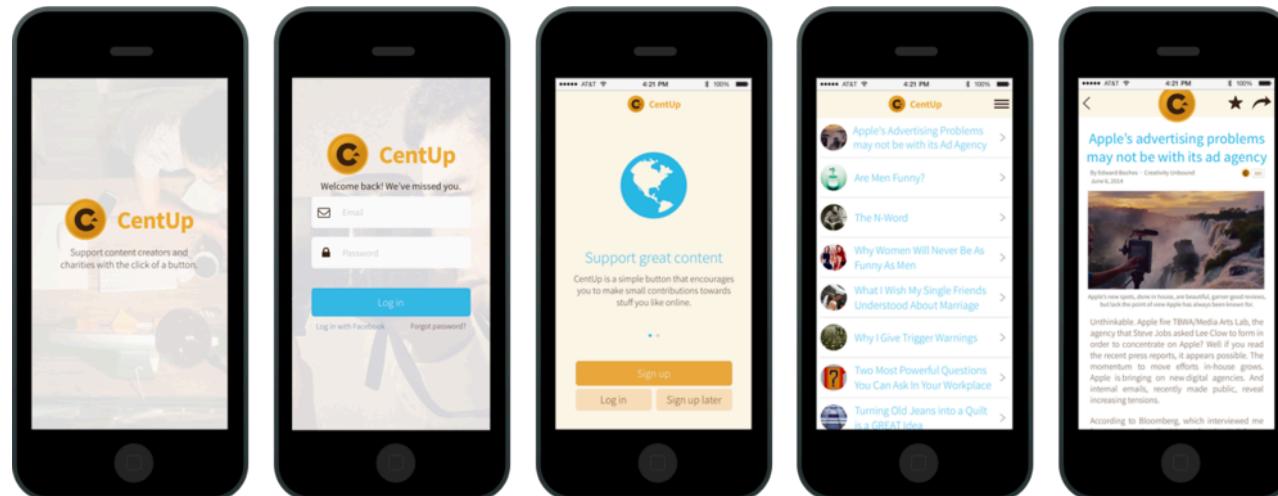
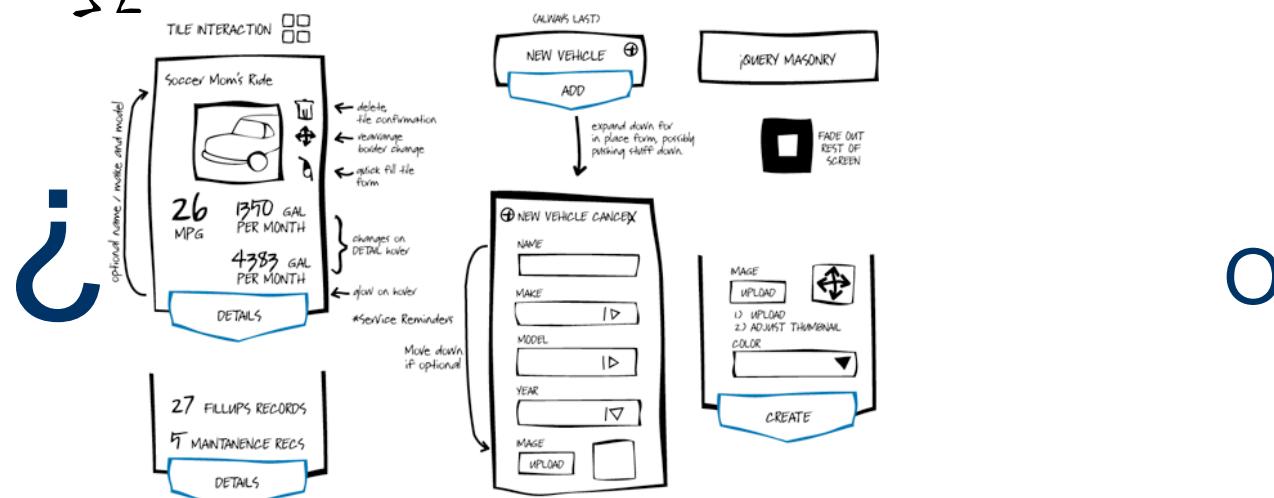
- Normalmente usado para Testeos Sumativos. Requiere un número importante de usuarios.
 - El número concreto dependerá del PODER que quieras conseguir con tu test, i.e. del tamaño del efecto que quieras ser capaz de detectar con un porcentaje dado de certeza (tipicamente un 95%).
 - A menos usuarios, menor poder estadístico, i.e. Menor capacidad de detectar que una determinada diferencia de medias es significativa. Además, a menos usuarios, mayor es el intervalo de confianza, i.e. Mayor será la variabilidad de tu resultado.
 - *Por tanto, si realizo cambios de diseño que suponen un cambio sutil en la experiencia de usuario... ¿necesitaré más o menos usuarios que si el cambio es menos sutil?*
 - Los tests cuantitativos requieren un diseño cuidadoso del instrumento de medición, y un control exhaustivo de las variables del experimento/cuestionario
 - ¡Usar siempre intervalos de confianza!

- Métodos Formativos/Sumativos
- Métodos Cuantitativos/Cualitativos
- **Métodos sobre Diseños Tempranos/Avanzados**
- Métodos sin Usuarios/Métodos con Usuarios

Técnicas: Prototipado temprano vs avanzado



¿Diferencias?



Técnicas: Prototipado temprano vs avanzado

➤ Recordad (Sesión de prototipado):



- Punto de vista



- Storyboard (task): visión de la app



- Paper prototypes/Sketches (task interface)

Tempranos



- Wireframes (grids, location, relative sizes).

baja fidelida



- Mockups (colors, fonts)

Avanzados
(alta fidelidad)



- Prototipos Nativos



- Código

- ¿Cuándo testamos sobre Low fidelity prototypes?
 - Tenemos una idea de app, y queremos saber si el usuario querría usar una app como la nuestra, y en qué modos podría serle útil.
¿Qué funcionalidades y rasgos querría el usuario encontrar en una app como esta?

Técnicas: Prototipado temprano vs avanzado

- Ventajas Prot Tempranos
 - Permiten involucrar diseñadores, usuarios, desarrolladores etc muy pronto en el proyecto
 - Tiempo de desarrollo
 - Facilidad de cambio
 - Foco sobre el diseño básico
 - No hay restricciones tecnológicas
 - Los usuarios critican más fácilmente
 - Los desarrolladores aceptan mejor las críticas
- Limitaciones:
 - No hay experiencia de uso real
 - Algunos comportamientos no son observables sobre prototipos
 - Entorno de ejecución no es realista
 - No se puede testear ni el tema gráfico, ni el emocional ni el rendimiento

Restaurant Finder - Anytown Chamber of Commerce

Overview Business Retail/Trade Housing Schools

Name	Type	Avg Price	Neighborhood	Closed	Brunch	Lunch	Dinner	Kids Menu	Hours
Applebee's	American	\$8.25	Central	Monday	Y	Y	Y	Y	Y
Demiglo's	American	\$7.50	River View	-	Y	Y	Y	Y	Y
Chili's	Mexican	\$7.00	Central	-	N	Y	Y	Y	Y
Chopsticks	Chinese	\$8.00	River View	Monday	N	Y	Y	N	Y
Dixie's	Continental	\$12.50	SE	Weekends	Y	Y	N	Y	Y
Doughnut Hut	Fast Food	\$1.25	Central	-	Y	Y	Y	N	Y
Frenchy's	American	\$6.50	SE	Weekends	Y	Y	N	Y	Y
Greek Café	Greek	\$8.25	University	Wednesday	N	Y	Y	Y	Y
International Cafe Asian	Asian	\$9.50	River View	-	N	Y	Y	Y	Y

Applebee's

Phone: 815-773-9270
Address: 234 - W. Third Ave.
Anytown, IL 12345

Open: Tuesday - Sunday
Hours: 6:00 AM to 10:00 PM
Closed: Monday

Description: Family-style American Eatery.
Eclectic menu with ethnic specialties:
- Mexican fajitas
- Italian pastas
Full bar with daily drink specials
Complete Children's Menu

Show Map



¿Qué NO tiene sentido comprobar con prototipos tempranos (contenido, navegación, presentación, interacción, PET, visión negocio)?

- ¿Cuándo testamos sobre High Fidelity prototypes?
 - Tenemos ya claros o queremos testar detalles específicos: cuántos elementos mostrar por cada línea de pantalla, colores, fuentes, el conjunto de tabs que hemos definido, ...
 - Queremos hacernos una idea de ciertas medidas de usabilidad.



■ Prototipo avanzado

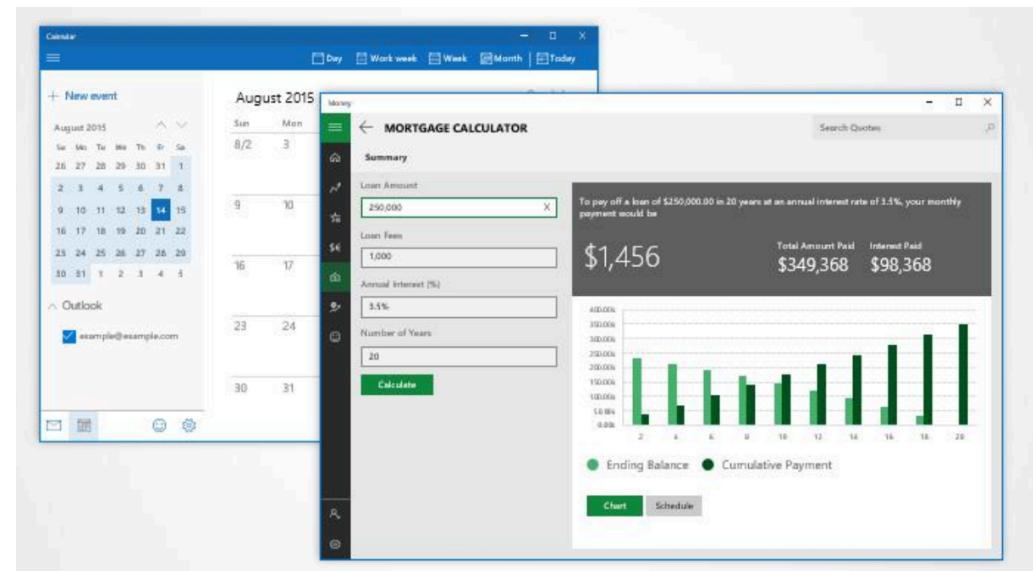
¿Ventajas e inconvenientes?

■ Ventajas

- Mayor profundidad del feedback de diseño
- Presentación cercana a la aplicación final
- Permite mayor nivel de interacción

■ Limitaciones:

- Más caro
- Menor habilidad para realizar cambios fácilmente
- Menor apertura del usuario a ser crítico
- Crítica más ‘pequeña’: normalmente centrada en layout (en lo concreto que estás mostrando) más que en funcionalidad
- El cliente puede pensar que ya está hecho !!!



Diseño

Prototipo en Papel

- Ejemplo real de prototipo temprano en papel preparado para testeо:



■ ¿Prototipos de baja o alta fidelidad? ¿Los testamos en papel o con ordenador?

HIGH-FIDELITY OR LOW-FIDELITY, PAPER OR COMPUTER? CHOOSING ATTRIBUTES WHEN TESTING WEB PROTOTYPES

Miriam Walker, Leila Takayama, and James A. Landay

- Experimento donde se compararon los resultados de realizar un test de usuario sobre prototipos de baja fidelidad (dibujados) vs alta fidelidad (HTML) tanto en papel como en ordenador.
- Resultados: Los prototipos de baja y alta fidelidad mostraron una capacidad similar para descubrir errores de usabilidad, independientemente del medio (papel u ordenador). Por tanto los diseñadores deberían elegir el modo que les resulte más apropiado.

Paper Prototyping - What is it good for? A Comparison of Paper- and Computer-based Low-fidelity Prototyping

Richard Sefelin, Manfred Tscheligi, Verena Giller

- Resultados: los prototipos en papel y sobre ordenador permitieron detectar prácticamente la misma cantidad y calidad de problemas críticos. Los sujetos prefieren prototipos en computadora. Aún así, testar sobre papel puede ser preferible cuando las herramientas de prototipado actual no soportan los componentes e ideas de tu diseño, cuando no se quiere excluir del test a personas sin habilidades de computación o cuando se desea que el test genere muchos nuevos sketches que pueden luego ser discutidos dentro del equipo de diseño.

- ¿Prototipos de baja o alta fidelidad? ¿Los testamos en papel o con ordenador?

The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. Juergen Sauer, Andreas Sonderegger (2009)

- Estudio empírico que introduce la fidelidad (paper prototype, computer prototype, fully operational appliance) y la estética del diseño (high vs. moderate) como variables independientes. 60 sujetos realizaron dos tareas típicas de móvil: enviar un mensaje de texto y suprimir un número de teléfono. Las variables dependientes fueron rendimiento y medidas subjetivas.
- Los resultados sugieren que el tiempo de compleción de una tarea puede ser sobreestimado si se usa un prototipo en papel. Además, una estética menos elaborada del prototipo parece que hace que los usuarios sobrepongan la estética de los prototipos de menor fidelidad. Las emociones de los usuarios se vieron afectadas de manera más positiva por el uso de los teléfonos más atractivos.
- ¡HAY MUCHOS MÁS PAPERS! Buscad en scholar.google.es

- Métodos Formativos/Sumativos
- Métodos Cuantitativos/Cualitativos
- Métodos sobre Diseños Tempranos/Avanzados
- **Métodos sin Usuarios/Métodos con Usuarios**

Técnicas: sin usuarios vs con usuarios



¿Diferencias?



¿

o



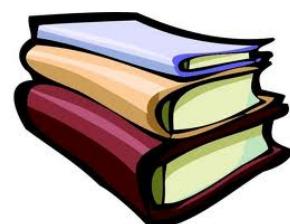
?

Técnicas: sin usuarios vs con usuarios

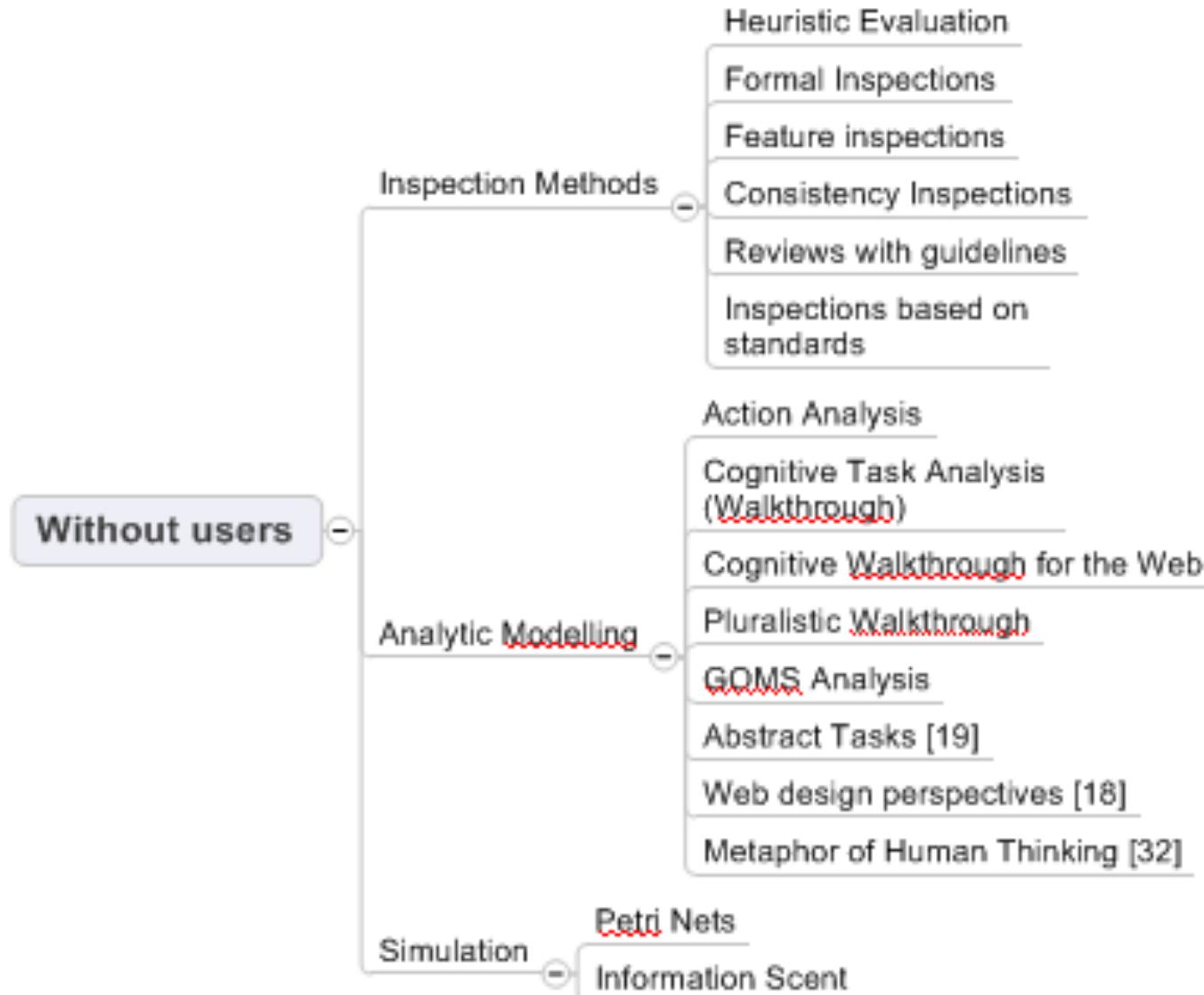
Sin usuarios	Con usuarios
<ul style="list-style-type: none">• No requiere usuarios, sino expertos (entre tres y cinco)• Ventajas: precio, tiempo. Permite descubrir más problemas que el testeo con usuarios• Desventajas: los resultados se ven afectados por la experiencia del evaluador tanto por lo que respecta a la usabilidad como por lo que respecta al dominio. Pueden escaparse errores de usabilidad importantes, y crearse falsas alarmas.	<ul style="list-style-type: none">• Requiere usuarios reales que sean representativos de la audiencia real. Se basa en las acciones/opiniones de esos usuarios.• Ventajas: permite detectar los problemas de usabilidad más importantes (más recurrentes, más globales o con mayor impacto en la eficiencia/eficacia/satisfacción del usuario)• Desventajas: más caros. Requiere artefactos que parezcan completos. Normalmente implica sacar al usuario de su entorno, lo que puede influir en los resultados. Estos resultados también pueden depender de la experiencia del facilitador.



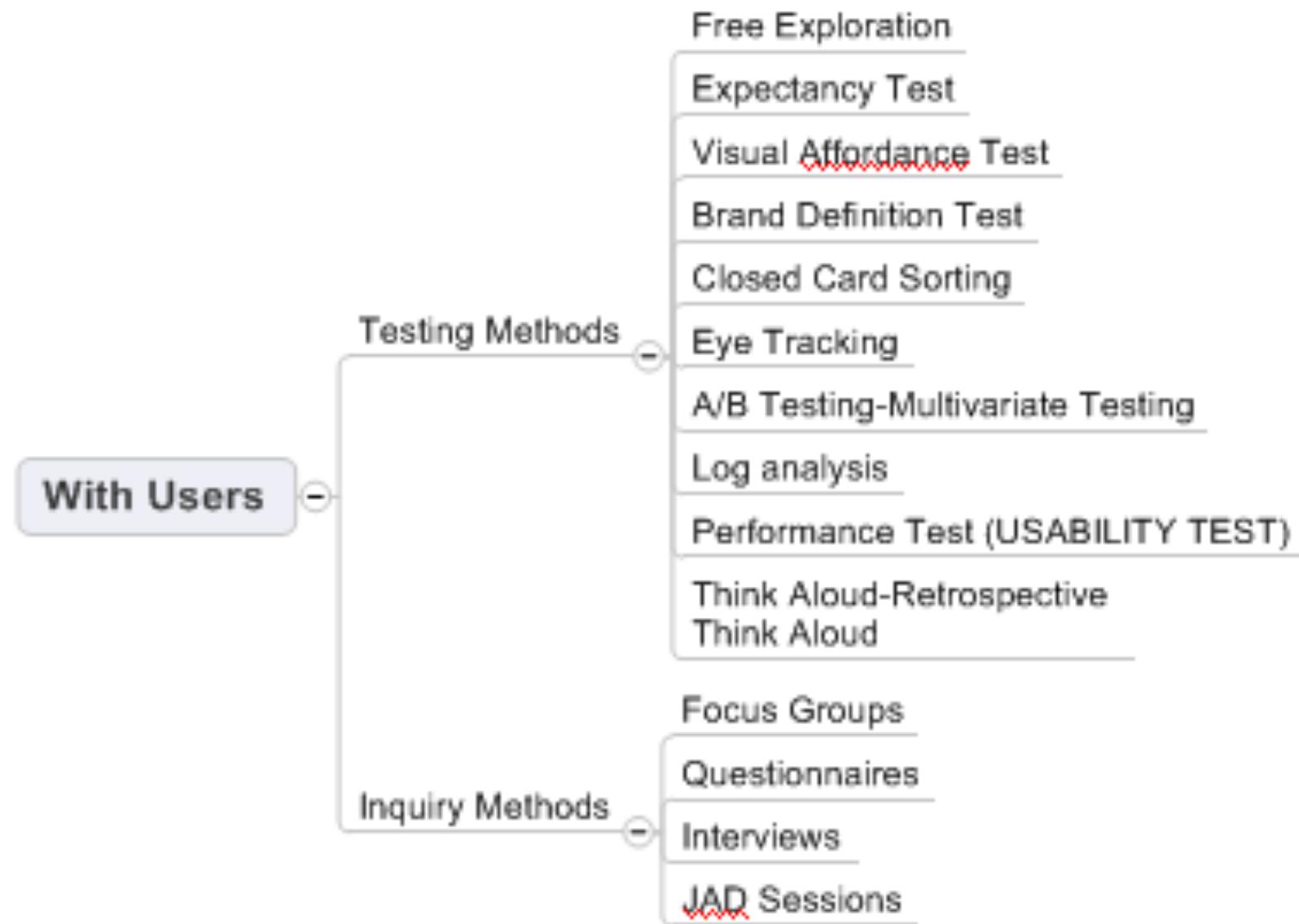
Reading:
[UsabilityInspectionMethods.pdf](#)



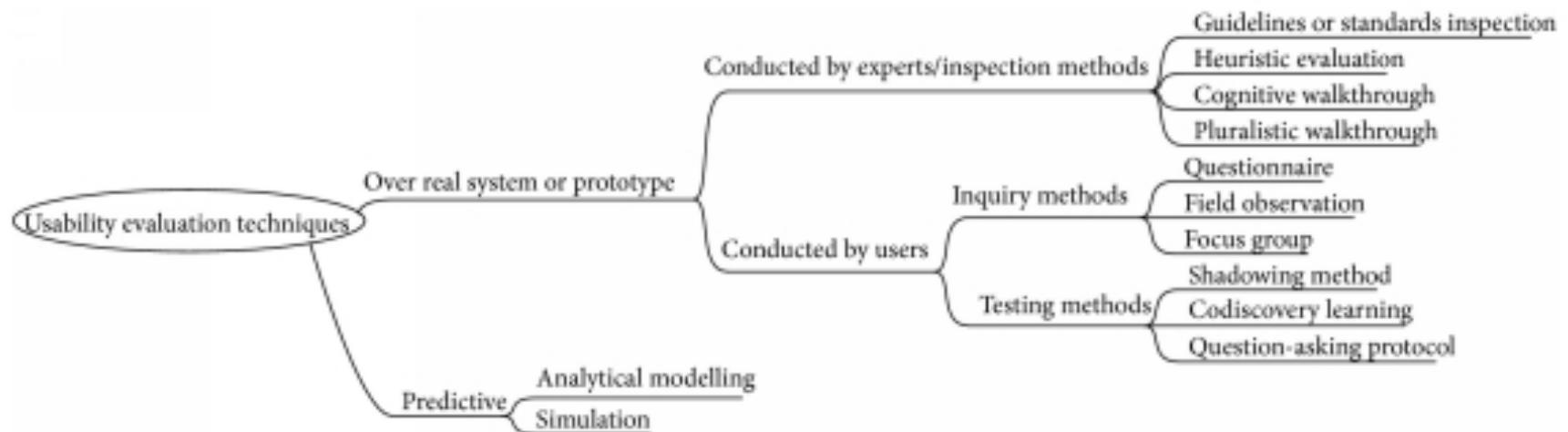
Técnicas: sin usuarios vs con usuarios



Técnicas: sin usuarios vs con usuarios



Técnicas: sin usuarios vs con usuarios



- Es importante elegir la técnica más apropiada para cada tarea:
 - **Confiabilidad/Fiabilidad/Precisión**
 - Si repitiera el test otra vez, ¿daría resultados similares?
 - **Generalizabilidad**
 - ¿Los resultados que obtengo con el test que han realizado mis alumnos a cambio de nota o de un cheque regalo son generalizables a una población más general?
 - **Realismo**
 - ¿Verías este comportamiento también en el mundo real?
 - **Comparación:** ¿Cómo variará la experiencia de usuario con distintos diseños? Va más allá de un mero 'me gusta/no me gusta'
 - **Coste:** ¿cuánto cuesta aplicar la técnica, en relación con el resultado que espero obtener?