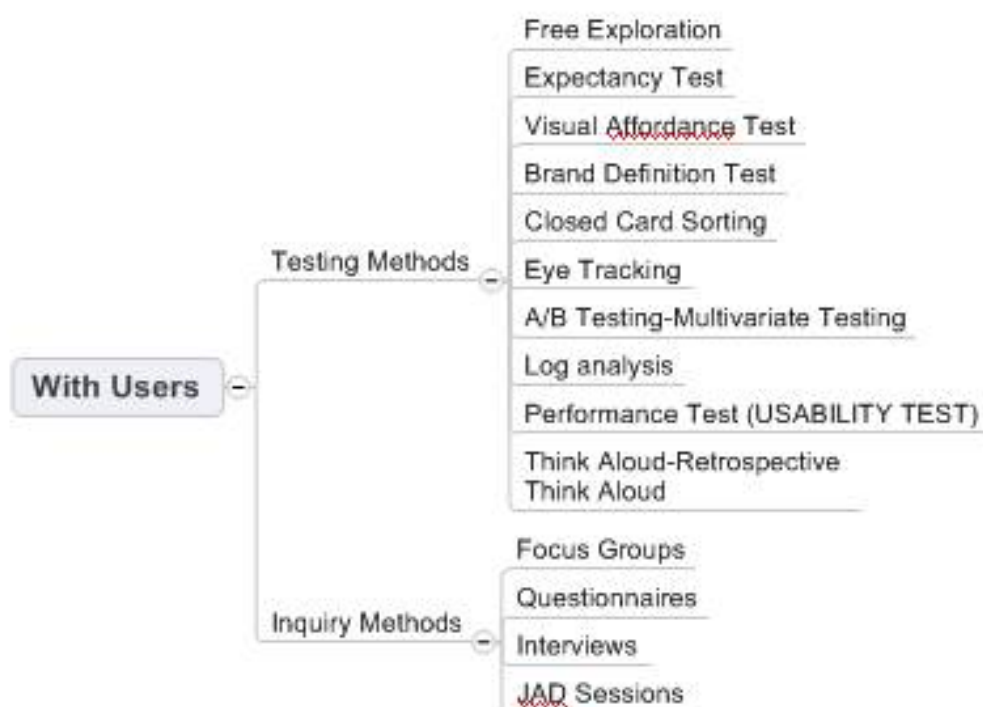


Gestión Calidad Software



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

UT con usuarios Técnicas



UT con usuarios: Cuestionarios

Introducción

- Método de indagación: preguntas a los usuarios sobre lo que hacen, pero no lo observas directamente.
- Existen varios cuestionarios estandarizados que han demostrado su validez interna y externa, y que se aplican a distintos niveles del test
 - A nivel global
 - A nivel de tarea
- Además, los cuestionarios pueden incluir preguntas abiertas que también pueden ser analizadas y priorizadas

UT con usuarios: Cuestionarios

Introducción

- Las actitudes son un indicador de la conducta, pero no la conducta en sí.
- Actitudes tienen dos propiedades:
 - Dirección: positiva o negativa
 - Intensidad: alta o baja
- Tipos de escalas para medir opinión
 - Método de escalamiento Likert
 - Diferencial semántico
 - Escala Guttman

UT con usuarios: Cuestionarios

Introducción

- Método de escalamiento Likert (Rensis Likert, 1932)
 - Conjunto de ítems presentados en forma de afirmaciones o juicios, ante los cuales se pide la reacción de los participantes.
 - Escala de 5, 7, 9 puntos... A cada punto se le asigna un valor numérico
 - Números más altos indican actitudes más POSITIVAS.
 - La puntuación total de la escala Likert se obtiene sumando los valores alcanzados respecto de cada frase (ESCALA ADITIVA)
- Construcción de la escala:
 - La escala se construye a partir de un número de afirmaciones acerca de la actitud que se pretende medir
 - A continuación se mide la correlación de cada ítem con la puntuación global de la escala. Todos los ítems deberían tener una correlación mayor de 0.3
 - Alpha de Crombach > 0.7 (para toda la escala)

5

UT con usuarios: Cuestionarios

Introducción

- Diferencial semántico (Osgood, Suci y Tannenbaum (1957)
- Califica al objeto de actitud a partir de un conjunto de adjetivos bipolares; entre cada par de éstos se presentan varias opciones. La persona selecciona aquella que en mayor medida refleje su actitud.
- La elección de objetivos depende del objeto de actitud a evaluar.



6

UT con usuarios: Cuestionarios

Introducción

- Escalograma de Guttman
 - Al igual que Likert se fundamenta en juicios ante los que el encuestado debe expresar su opinión (totalmente en desacuerdo.. Totalmente de acuerdo)
 - Diferencia con respecto a Likert: los ítems miden lo mismo, pero con un grado de intensidad distinto.
 - Es unidimensional y acumulativa.
 - Ejemplo:

Ítems	1	0
Usted sabe sumar y restar números de tres cifras	(si)	(no)
Usted sabe multiplicar números de tres cifras	(si)	(no)
Usted sabe dividir números de tres cifras	(si)	(no)

7

- <https://www.youtube.com/watch?v=aA6HB2QwRbc>

UT con usuarios: Cuestionarios

Cuestionarios Globales

- Por qué usar cuestionarios estandarizados en lugar de inventar uno propio
 - FIABILIDAD: los usuarios contestan consistentemente (siempre parecido)
 - VALIDEZ: el test mide lo que en realidad pretendía medir
 - SENSITIVIDAD: el test es capaz de detectar diferencias significativas
 - UTILITY: Corto, fácil de contestar, fácil de administrar y fácil de puntuar

8

UT con usuarios: Questionarios

Questionarios Globales

■ A nivel global (Post-test): Listado de los más conocidos

- CSUQ: Computer System Usability Questionnaire [Lewis 1995]: 19 frases, todas en positivo. Miden cuatro dimensiones: utilidad del sistema, calidad de la información, calidad de la interfaz y satisfacción general. Pensado para ser administrado off-line
- PSSUQ: Post-Study System Usability Questionnaire [Lewis 1995]: variación del CSUQ para ser administrado en persona
- QUIS: Questionnaire for User Interface Satisfaction [Chin 88]. 27 sentencias divididas en cinco categorías: reacción general, pantalla, Terminología/Info sistema, Aprendizaje, Capacidades del Sistema.
- **EUCS** : End-user computing satisfaction [Abdinnour-Helm 2005]. 12 items que representan cinco dimensiones: content, accuracy, format, ease of use and timeliness. Todos se combinan en una medida EUCS global. Cinco puntos por ítem (Almost never... almost always)
- **USE**: Usefulness, Satisfaction and Ease of Use [Lund 2001]. 30 ítems divididos en cuatro categorías: utilidad, satisfacción, facilidad de uso y facilidad de aprendizaje. Para cada uno, escala de siete puntos.
- **SUS**: Systems Usability Scale [Brooke 1996]: 10 preguntas, 5 formuladas en positivo y 5 en negativo. Da una medida global de la usabilidad del sitio o app.
- **Net promoter Score (NPS)** ©: Una sola medida sobre la lealtad del usuario
- SUMI [Kirakowski 1996]
- **Questionarios que incluyen ranking (basado en BD)**
 - **SUPR-Q**: Cuestionario de UX . Incluye usability (4 items), credibility (trust, value & comfort, 5 items), loyalty (2 items, uno de los cuales es la Net Promoter Score) y Appearance (2 items). Ítems de 5 puntos (strongly disagree to strongly agree excepto el Item 10 (net promoter question))
 - **WAMMI**, ACSI, Opinion Lab
 - UEQ (Industria)
 - ACSI
 - OpinionLab

UT con usuarios: Questionarios

Questionarios Globales: EUCS

■ Ejemplo: EUCS [2005]

Item	Question	Dimension
EUCS1	Does the site provide the precise information you need?	Content
EUCS2	Does the site information content meet your needs?	Content
EUCS3	Did the site provide help that seemed to be just about exactly what you need?	Content
EUCS4	Did the site provide sufficient information?	Content
EUCS5	Was the site accurate?	Accuracy
EUCS6	Were you satisfied with the accuracy of the site?	Accuracy
EUCS7	Did you think the site information is presented in a useful format?	Format
EUCS8	Was the site information clear?	Format
EUCS9	Was the site user friendly?	Ease of Use
EUCS10	Was the site easy to use?	Ease of Use
EUCS11	Did you get the site information you need quickly?	Timeliness
EUCS12	Did the site provide up-to-date information?	Timeliness

UT con usuarios: Cuestionarios

Cuestionarios Globales: USE

Ejemplo: USE [Lund 2001]

Usefulness

- It helps me be more effective.
- It helps me be more productive.
- It is useful.
- It gives me more control over the activities in my life.
- It makes the things I want to accomplish easier to get done.
- It saves me time when I use it.
- *It meets my needs.*
- It does everything I would expect it to do.

Ease of Use

- It is easy to use.
- It is simple to use.
- It is user friendly.
- It requires the fewest steps possible to accomplish what I want to do with it.
- *It is flexible.*
- *Using it is effortless.*
- *I can use it without written instructions.*
- *I don't notice any inconsistencies as I use it.*
- *Both occasional and regular users would like it.*
- *I can recover from mistakes quickly and easily.*
- *I can use it successfully every time.*

Ease of Learning

- I learned to use it quickly.
- I easily remember how to use it.
- It is easy to learn to use it.
- *I quickly became skillful with it.*

Satisfaction

- I am satisfied with it.
- I would recommend it to a friend.
- It is fun to use.
- It works the way I want it to work.
- It is wonderful.
- I feel I need to have it.
- It is pleasant to use.

Users rate agreement with these statements on a 7-point Likert scale, ranging from strongly disagree to strongly agree. Statements in *italics* were found to weight less heavily than the others.

11

UT con usuarios: Cuestionarios

Cuestionarios Globales: SUS

Ejemplo: SUS

- SUS (Systems Usability Scale) [Brooke 1996]: 10 preguntas, 5 formuladas en positivo y 5 en negativo. Da una medida global de la usabilidad del sitio (no se divide en componentes).
 - Es el que ha mostrado mayor fiabilidad [Tullis and Stetson, 2004]
- La media en SUS es 66%, el percentil 25 está en 57% y el percentil 75 está en 77%. Esto sugiere que deberíamos intentar conseguir un SUS mayor o igual a 80% (teniendo en cuenta intervalos de confianza).

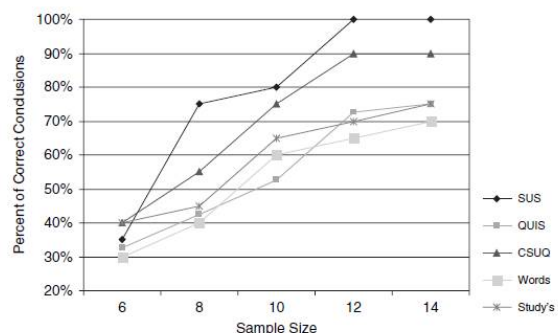


FIGURE 6.13

Data illustrating the accuracy of the results from random subsamples ranging from size 6 to size 14. This graph shows what percentage of the random samples yielded the same answer as the full dataset at the different sample sizes. Source: Adapted from Tullis and Stetson (2004).



Estructura y evaluación del SUS

12

UT con usuarios: Questionarios

Cuestionarios Globales: SUS

SUS: System Usability Scale

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	

UT con usuarios: Questionarios

Cuestionarios Globales: SUS

PUNTUACIÓN TOTAL: 22

PUNTUACIÓN SUS: $22 \times 2.5 = 55$

Cómo calcular el valor total de la escala (0..100)

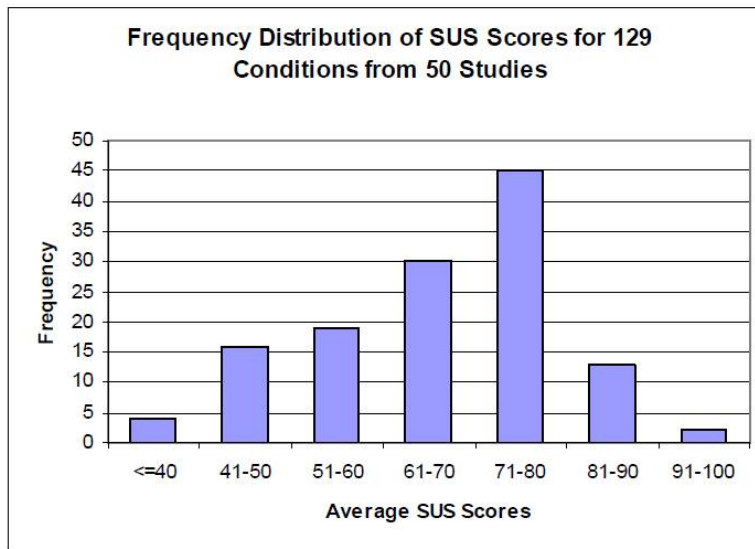
- Da la vuelta a los ítems 2,4,6,8,10
- Suma posiciones para cada ítem (0 to 4)
- Multiplica por 2.5

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	4
1	2	3	4	5	1
1	2	3	4	5	1
1	2	3	4	5	4
1	2	3	4	5	1
1	2	3	4	5	2
1	2	3	4	5	1
1	2	3	4	5	1
1	2	3	4	5	4
1	2	3	4	5	3

UT con usuarios: Questionarios

Questionarios Globales: SUS



Percentiles:	
10th	47.4
25th	56.7
50th	68.9
75th	76.7
90th	81.2
Mean	66.4



¿CUÁL DEBERÍA SER
NUESTRO OBJETIVO SUS?

15

UT con usuarios: Questionarios

Questionarios Globales: SUS

- El SUS puede adaptarse a la idiosincrasia del sistema particular que estemos evaluando
 - Example: [OHIM](#)

	Strongly disagree							Strongly agree
1. I think that I would like to use this web site frequently	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5			
2. I found the web site unnecessarily complex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5			
3. I thought the web site was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5			
4. I think that I would need the support of a technical person to be able to use this web site	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5			
5. I found the various functions in this web site were well integrated	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5			
6. I thought there was too much	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

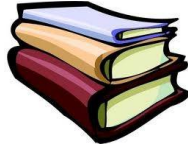
- Calculadora de SUS:
www.measuringux.com/SUS-scores.xls

16

UT con usuarios: Cuestionarios

Cuestionarios Globales: SUS

- Investigaciones más recientes sugieren que el SUS puede tener en realidad dos factores: Usable (8 ítems) y 'Aprendible' (Learnable) (2 ítems: 4 y 10).



SUS Factor Analysis

- También algunas investigaciones muestran que el SUS se correlaciona fuertemente con la medida de NPS

17

UT con usuarios: Cuestionarios

Cuestionarios Globales: NPS

- Net Promoter Score: medida obtenida a través de una sola pregunta sobre la lealtad del usuario:
 - *How likely is it that you'll recommend this product to a friend or colleague? (0..10)*
- Mide el impacto de negocio de la UX. Tres segmentos:
 - Promotores: Respuestas de 9 a 10
 - Pasivos: Respuestas de 7 a 8
 - Detractores: Respuestas de 0 a 6
- NPS (-100..+100): %Promotores-%Detractores

Not at
all
Likely

0
○

1
○

2
○

3
○

4
○

5
○

6
○

7
○

8
○

9
○

10
○

Neutral

Extremely
Likely

18

UT con usuarios: Cuestionarios

Cuestionarios Globales: con ranking

- Cuestionarios que proporcionan rankings que comparan tus resultados con los de otras webs (de tu sector o genéricas) (benchmark dataset)
 - **SUPR-Q** (Standardized Universal Percentile Rank, Jeff Sauro)
 - Validación del cuestionario:
SUPR-Q: A comprehensive Measure of the Quality of the Website User Experience. (Journal of Usability Studies)
 - **WAMMI** (www.wammi.com) (sucesor del SUMI)
 - **UEQ: User Experience Questionnaire**
 - **ACSI** (www.theacsi.org): especialmente interesante para government websites. American Customer Satisfaction Index.
 - OpinionLab (www.OpinionLab.com)

19

UT con usuarios: Cuestionarios

Cuestionarios Globales con ranking: SUPR-Q

- SUPR-Q: 8 ítems de rango 5+NPS

Usability:

1. This website is easy to use.
2. It is easy to navigate within the website.

Credibility: (Trust, Value & Comfort)

3. The information on the website is credible.
4. The information on the website is trustworthy.

I feel comfortable purchasing from this website. (Alternate for eCommerce websites)

I feel confident conducting business with this website. (Alternate eCommerce websites)

Loyalty:

5. How likely are you to recommend this website to a friend or colleague?
6. I will likely visit this website in the future.

Appearance:

7. I found the website to be attractive.
8. The website has a clean and simple presentation.

- Cálculo global: Suma de la puntuación de todos los ítems + $\frac{1}{2}$ puntuación NPS (5)

- La puntuación SUPR-Q se expresa como un percentil (VENTAJA)

21

UT con usuarios: Cuestionarios

Cuestionarios Globales con ranking: SUPR-Q

■ Ejemplo: adaptación a Amazon

Please rate your level of agreement to the following statements about the Amazon website.

	Strongly Disagree 1	2	3	4	Strongly Agree 5
I will likely return to the Amazon website in the future.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is easy to navigate within the Amazon website.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The information on the Amazon website is credible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I find the Amazon website to be attractive.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The information on the Amazon website is trustworthy.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Amazon website has a clean and simple presentation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Amazon website is easy to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Pregunta final: ¿Cómo de probable es que recomiendes esta web a un amigo o compañero? (NPS)

Not at all Likely	Neutral						Extremely Likely			
0	1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

22

UT con usuarios: Cuestionarios

Cuestionarios Globales con ranking: SUPR-Q

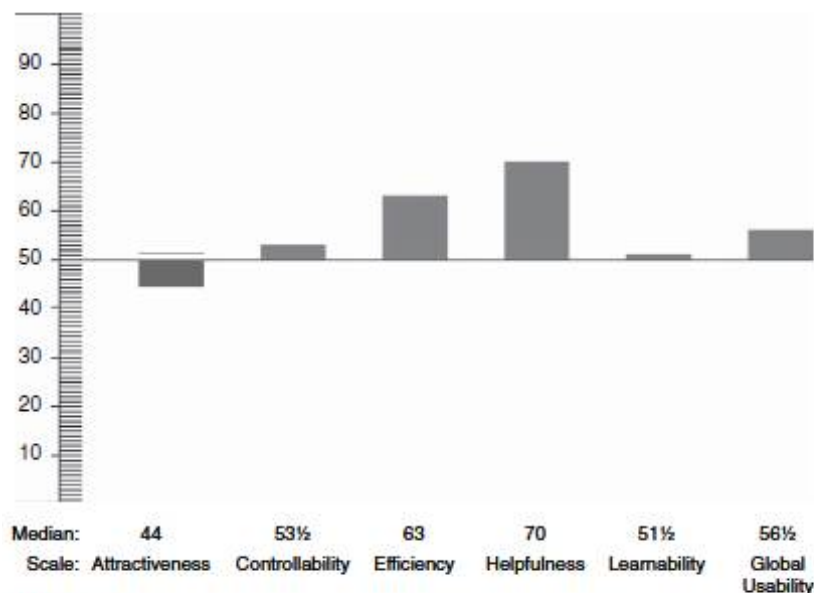
- Además de los valores para cada factor, SUPR-Q está respaldado por una BD que permite generar un ranking de percentiles
 - E.g. tu web está en el percentil 75 en usability, lo que significa que está entre el 25% mejor.
- Con la versión antigua (13 ítems) el factor de usabilidad del SUPR-Q muestra una correlación alta con la puntuación SUS, $r = .96$, $p < .001$, lo que indica que sólo cuatro preguntas explican el 93% (0.96^2) de la variación en el SUS.
 - Por tanto estas preguntas pueden ser un buen sustituto del SUS para sitios Web.
- Validez y fiabilidad del SUPR-Q: <http://www.suprq.com/>

23

- Cuestionario WAMMI:
<http://www.wammi.com/samples/index.html>
 - Los resultados se dividen en cinco áreas: atractivo, controlabilidad, eficiencia, ayuda (helpfulness) y aprendibilidad, más una medida global de usabilidad.
 - Las preguntas están estandarizadas (a partir de la comparación con la BD de referencia), por lo que un resultado de 50 es medio, y un resultado de 100 indica una usabilidad perfecta según este cuestionario.

24

- WAMMI
 - Gráfico de ejemplo que muestra cómo un sitio web se posiciona en relación a puntuaciones medias en cada eje.



25

UT con usuarios: Questionarios

Questionarios Globales con ranking: UEQ

- UEQ: User Experience Questionnaire
 - Industria
 - <http://www.ueq-online.org/>
 - Seis dimensiones: atractivo, familiaridad, eficiencia, dependabilidad, estimulación, novedad.

	1	2	3	4	5	6	7		
desagradable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	agradable	1
no entendible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	entendible	2
creativo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sin imaginación	3
fácil de aprender	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	difícil de aprender	4
valioso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	de poco valor	5
aburrido	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	emocionante	6
no interesante	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	interesante	7
impredecible	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	predecible	8
rápido	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	lento	9
original	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	convencional	10
obstrutivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	impulsor de apoyo	11
bueno	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	malo	12
complicado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	fácil	13
repeler	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	atraer	14
convencional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	novedoso	15
incómodo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	cómodo	16
seguro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	inseguro	17
activante	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	adormecedor	18
cubre expectativas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	no cubre expectativas	19
ineficiente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	eficiente	20
claro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	confuso	21
no pragmático	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	pragmático	22
ordenado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	sobrecargado	23
atractivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	feo	24
simpático	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	antipático	25
conservador	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	innovador	26

26

UT con usuarios: Questionarios

Questionarios Globales con ranking: ACSI

- ACSI: American Customer Satisfaction Index
- <https://community.verint.com/b/customer-engagement/posts/acsi-american-customer-satisfaction-index-score-its-calculation>
 - Indicador económico que mide la satisfacción global de los consumidores de un país. Tb se utiliza para medir la satisfacción de los clientes de una determinada empresa

What is your overall satisfaction with [our product or service]?

Very dissatisfied ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very satisfied

To what extent has [our product or service] met your expectations?

Falls short of your expectations ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Exceeds your expectations

How well did [our product or service] compare with the ideal [type of offering]?

Not very close to the ideal ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very close to the ideal

27

UT con usuarios: Questionarios

Questionarios Globales con ranking: ACSI

- ACSI: American Customer Satisfaction Index
- Benchmark: resultados por industria. En cada sector el peso de cada pregunta varía. E.g. La media de Computer Software está en 71

ACSI Score

85	Personal Care & Cleaning			
84	Credit Unions	Pet Food		
83	Breweries	Electronics (TV/VCR/DVD)	Food Manufacturing	Soft Drinks
82	Automobiles	Express Delivery	Internet Retail	
81	Ambulatory Care	Property & Casualty Insur.		
80	Apparel	Full Service Restaurants	Search Engines	Major Appliances
79	Athletic Shoes			
78	Cigarettes	Health Stores	Life Insurance	Limited Service Restaurants
76	Specialty Retail Stores	Supermarkets		
75	Banks	Hospitals	Hotels	Internet News
74	Dept. & Discount Stores	Energy Utilities	Gasoline Stations	Personal Computers
73	Fixed Line Telephone	Health Insurance		
71	Cellular Telephones	Computer Software		
70	Motion Pictures			
69	Network/Cable TV News			
68	Wireless Telephone			
67	Broadcasting TV News			
64	Cable & Satellite TV	Newspapers		
62	Airlines			Source: TheACSI.org

28

UT con usuarios: Questionarios

Questionarios Globales con ranking: ACSI

- Una aproximación (fórmula del Estado de Ohio) es

$$\begin{aligned}
 & ((Satisfaction-1)*.3885 + \\
 & (Expectancy-1)*.3190 + \\
 & (Performance-1)*.2925) / 9 * 100
 \end{aligned}$$

Nota: si puntúas de 0 a 9 no hace falta restar 1

29

UT con usuarios: Cuestionarios

VoC studies

- Estos cuestionarios pueden ser utilizados en VoC (Voice of client) studies
 - Se aplican en live sites, y utilizan servicios especializados para interceptar usuarios (e.g. userzoom).

30

UT con usuarios: Cuestionarios

Cuestionarios Globales

- ¡CUIDADO!
 - Las opiniones expresadas en un post-test están sujetas al peak-end effect: las partes de la experiencia últimas y las más intensas, sean éstas positivas o negativas, son las que más impactan en la evaluación final del usuario.
 - Un bajo nivel de usabilidad global no nos da pistas sobre qué parte de la aplicación causó esa baja puntuación.

31

UT con usuarios: Questionarios

Questionarios Tarea

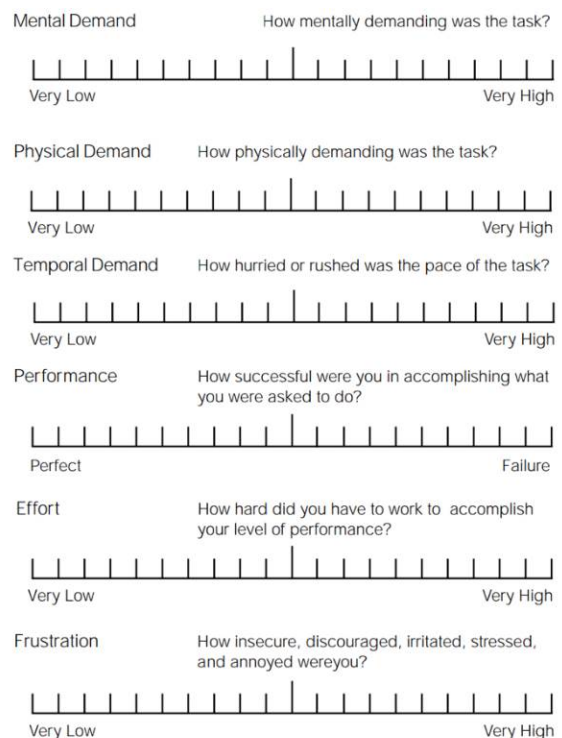
- A nivel de tarea (pre-post escenario): tienen la ventaja de que permiten detectar qué partes de tu interfaz (workflow) son las más problemáticas. Además, al estar la pregunta más cerca de la tarea, el usuario puede proporcionar una información clara de su actitud frente a la experiencia.
- Listado de los más conocidos
 - Single Ease Question (SEQ) (1 pregunta)
 - NASA-TLX: Post-task Task Load Index (6+1 preg)
 - SMEQ (1 preg)
 - UME (1 preg)
 - Expectation Rating
 - ASQ (3 preg)
 - Awareness-Usefulness and

32

UT con usuarios: Questionarios

Questionarios Tarea: Expectation Rating

- A nivel de tarea
 - SEQ
 - Considerada globalmente, esta tarea fue...
 - 1 (muy difícil)...7 (muy fácil)
 - NASA-TLX: 6 cuestiones, cada una con una escala de 21 puntos sin etiquetar, de muy bajo a muy alto. Cada pregunta aborda una dimensión de la carga mental percibida: demanda mental, demanda física, presión temporal, éxito percibido, esfuerzo percibido, frustración.
 - Posteriormente, los usuarios marcan qué categorías fueron más importantes para ellos para la tarea que estaban haciendo
 - Hay una app IOS gratuita (NASA TLX) para hacer el test



UT con usuarios: Questionarios

Questionarios Tarea: Expectation Rating

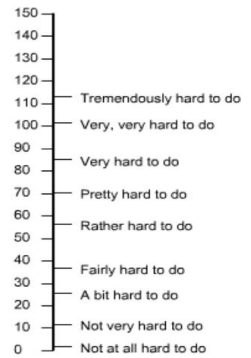
■ A nivel de tarea

■ SMEQ: Subjective Mental Effort Questionnaire

- No es una escala Likert: : evita efecto cielo y suelo, y mejora las propiedades de escala de intervalo
 - Versión online : <http://www.usablesurveys.com/>
 - Estudio Sauro&Dumas 2009: SMEQ mejor que UME

■ UME: Usability Magnitude Scale

- El usuario crea su propia escala basada en 'juicios de ratio' (Tampoco es una escala Likert)
- Imaginemos que estamos midiendo dificultad de la tarea: el usuario asigna un número (el que quiera) a la primera tarea, luego a la segunda, etc.
 - Se suele dar una tarea base muy sencilla con una medida de dificultad 10 e.g. El resto de tareas se puntúan tomando como referencia ese 10.
- Al final no nos importan los números sino el ratio de las puntuaciones (se asume que si asigno un 10 a la primera tarea y un 20 a la segunda, la segunda fue el doble de difícil que la primera)
- Difícil de explicar y más complicado convertir las puntuaciones individuales en una sola medida.



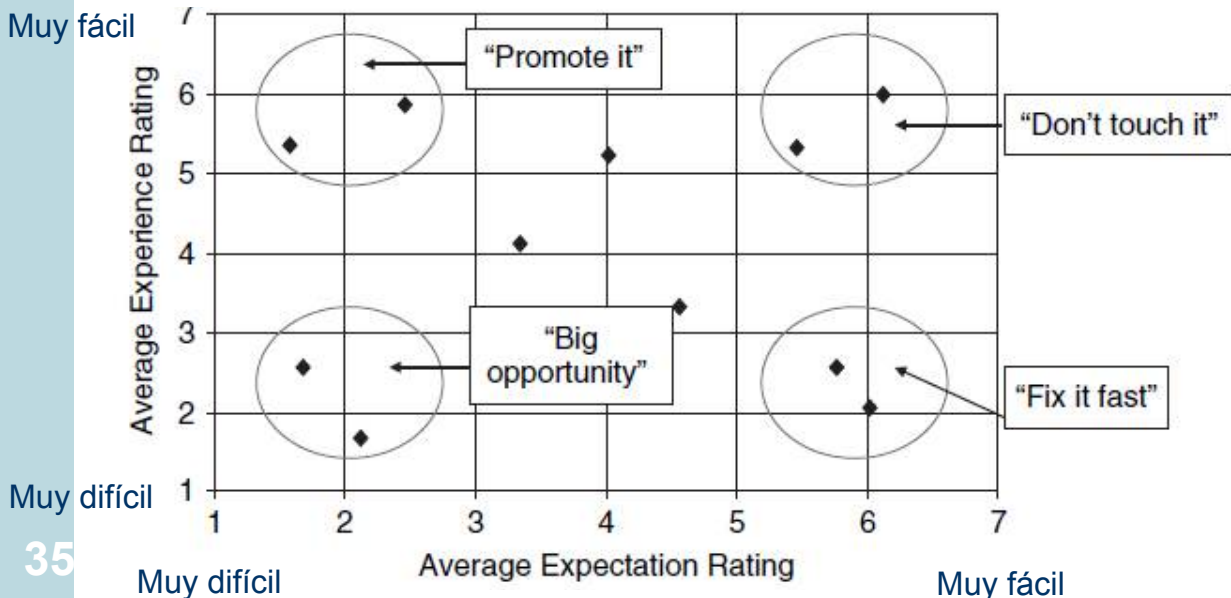
UT con usuarios: Questionarios

Questionarios Tarea: Expectation Rating

■ A nivel de tarea

■ Pre/Post-Escenario: Expectation Rating (7 puntos):

- Pre-task: ¿Cómo piensas que va a ser de fácil/difícil la siguiente tarea?
- Post-task: ¿Cómo ha sido de fácil/difícil para ti realizar la tarea?



UT con usuarios: Cuestionarios

Cuestionarios Tarea: ASQ

- A nivel de tarea
 - Post-Escenario: ASQ (After-Scenario-Questionnaire: 7points) [Lewis 1991]
 - Estoy satisfecho con la facilidad con la que he completado esta tarea
 - Estoy satisfecho con la cantidad de tiempo que me ha llevado completar esta tarea
 - Estoy satisfecho con la información de soporte (ayuda en línea, mensajes, documentación) durante la compleción de esta tarea.
 - Sugerencia: si vuestro sistema no ofrece ayuda en línea, cambiad la pregunta por: *Estoy satisfecho con la ayuda proporcionada por la interfaz durante la compleción de esta tarea*

36

UT con usuarios: Cuestionarios

Cuestionarios Tarea: Awareness-Usefulness gap

- A nivel de tarea
 - Awareness-usefulness gaps
 - “¿Eras consciente de la existencia de esta funcionalidad antes de realizar este estudio? (sí o no)
 - En una escala de 1 a 5, ¿cómo de útil te resulta esta funcionalidad?
 - Convertir el ítem likert en binario (e.g. 4 ó 5 útil, resto no útil) y graficar
 - Las diferencias entre awareness y usefulness indican qué tareas deberías promocionar (a través de un rediseño de la jerarquía visual e.g.) en tu web
 - ¡Tareas que se consideran muy útiles pero de las que había poca consciencia deben ser promocionadas!

37

➤ Graficación awareness-usefulness

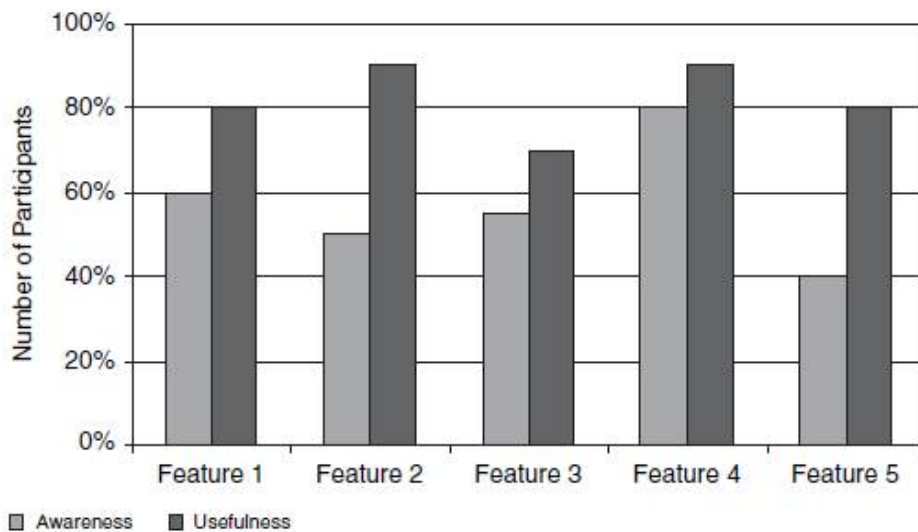


FIGURE 6.29

Data from a study looking at awareness–usefulness gaps. Items with the greatest difference between the awareness and usefulness ratings, such as Features 2 and 5, are those you should consider making more obvious in the interface.

38

UT con usuarios: Cuestionarios

Cuestionarios Globales

- Limitaciones de los cuestionarios post-test y post-task
 - No dejan de recoger datos auto-reportados, y por tanto pueden ser poco fiables
 - Miden percepción subjetiva, no rendimiento objetivo, y por tanto son más útiles cuando se acompañan de medidas de rendimiento
 - **No indican ni el porqué de la puntuación ni qué hacer para mejorar**

40

UT con usuarios: Cuestionarios

Preguntas abiertas

- Preguntas abiertas (post-test o post-task)
 - Razones por las que los usuarios son promotores o detractores de un producto
 - Insights de usuarios de estudios de campo
 - Quejas sobre un producto enviadas al servicio de atención al cliente
 - Por qué una tarea fue difícil de completar
 - ...
- El proceso suele consistir en convertir comentarios abiertos en categorías, cuantificarlos y someterlos a análisis estadístico.
- <http://www.measuringusability.com/blog/quantify-comments.php>

41

UT con usuarios: Cuestionarios

Comparación de cuestionarios de usabilidad

- ACTIVIDAD OPTATIVA: Leer el artículo “Comparison of Usability Questionnaires” y elaborar una tabla comparativa de las bondades/ problemas de los distintos cuestionarios según el estudio.

Criterio	SUS	QUIS	CSUQ	Microsoft	Tullis SUMI, WAMMI



[Lectura: ComparisonUsabilityQuestionnaire.pdf](#)



42

UT con usuarios: Questionarios

Validación de cuestionarios

- Todos los cuestionarios deben ser validados con respecto a su fiabilidad y validez. Una vez validados, deben ser revalidados cuando cambia el contexto en que se aplica (el tipo de aplicación).

	Critical Issues	Software/Hardware Environment			
		Web ^a	General ^b	Mobile ^c	Games ^d
Competitive environment	Speed of change	Rapid	Moderate	Rapid	Moderate
	Ease of entry and exit	User: high Provider: high	User: low Provider: very low	User: high Provider: low	User: medium Provider: very low
	Competitive advantage timeframe	Very short	Medium	Short	Medium
Marketing environment	Communication model	Many to many	One to many	Many to many	Either
	Functionality vs. flow Locus of motivation	Both required Either intrinsic, extrinsic, or both	Functionality required Extrinsic	Both required Either	Both required Intrinsic
Usage	Scope Location specific	General Yes	Specific No	Specific Yes	Specific No
Usability	Design flexibility Customer experience	High Use, then buy	Low Buy, then use	Moderate Buy, then use	Low Buy, then use

^aWeb refers to Web sites used on the PC (e.g., www.landsend.com).

^bGeneral refers to general application software used on the PC (e.g., productivity software).

^cMobile refers to software used on mobile devices such as PDAs or cell phones (e.g., Documents to Go).

^dGames refers to game software on the PC (e.g., Doom).

43

UT con usuarios: Questionarios

Validación de cuestionarios

- E.g. en el caso de EUCS, la valoración de sus distintas dimensiones se ve afectada por las limitaciones intrínsecas de la aplicación.

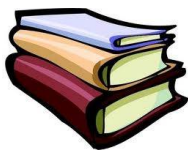
EUCS Dimension	Critical Issue	Software/Hardware Environment			
		Web	General	Mobile	Games
Content	Scope	Virtually unlimited because of thin client model	Limited to specific release (fat client)	Virtually unlimited because of thin client model	Can be either
Accuracy	Functionality	Potential tradeoff of rapid updating with accuracy loss	Potential tradeoff of slow updating with accuracy gain	Potential tradeoff of rapid updating with accuracy loss	Potential tradeoff of rapid updating with accuracy loss
Format	Flow	Information presentation focus; frequent changes to retain user interest; high site-specific variance, but some consistent design principles	Operation presentation focus (e.g., menus, dialog boxes); changes often avoided due to user learning curve requirements	Both information and operation focus; some design principles shared	Operation focus predominant, information focus often present
Ease of use	Design flexibility	Standard interface/simple usage	Release specific interface; more complex than Web usage	Interface depends on specific hardware/software combination	Release specific interface; more complex than Web usage
Timeliness	Speed of change	Frequent/rapid updates possible	Infrequent updates, must be installed by user	Frequent updates to software only	Real time updating possible

44

UT con usuarios: Cuestionarios

Puntos en ítems de las escalas

- Parece que escalas con pequeños rangos de respuesta (5 o menos) proporcionan puntuaciones que son en general menos válidas y menos discriminantes que aquellas con seis o más categorías de respuesta (Preston and Colman, 2000)



- Lectura: User Experience Rating Scales with 7, 11 or 101 points:
- does it matter?

UT con usuarios: Cuestionarios

Referencias

[Lewis95] Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7, 57–78.

[Lund 2001]

[Chin 88] Chin, J. P., Diehl, V. A., and Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In: *CHI '88. Conference Proceedings on Human Factors in Computing Systems*. New York: ACM, pp. 213–218.

[Brooke 1996] Brooke, J. (1996). SUS: A “quick and dirty” usability scale, In Jordan, P. W., Thomas, B. T., and Weerdmeester, B. A. (Eds.), *Usability Evaluation in Industry*. UK: Taylor and Francis, pp. 189–194.

[Kirakowski 1996] Kirakowski, J. (1996). The software usability measurement inventory: Background and usage. In Jordan, P., Thomas, B., and Weerdmeester, B. (Eds.), *Usability Evaluation in Industry*. UK: Taylor and Francis, pp. 169–177.

[Lewis 1991] Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *SIGCHI Bulletin*, 23(1), 78–81.

[Abdinnour-Helm 2006]. *Using the End-User Computing Satisfaction (EUCS) Instrument to Measure Satisfaction with a Web Site*.