

# DISEÑO Y EVALUACIÓN DEL INTERFAZ GRÁFICO DE UN CURSO WEB

Claudia Engel y Juan José Casares

Centro de Supercomputación de Galicia

*El presente trabajo describe las experiencias durante el proceso de diseño del interfaz de un curso Web para el autoestudio de Fortran 90, un lenguaje de programación. Se realizó un análisis sistemático y cuantitativo. El curso fue evaluado por un grupo de 12 usuarios, modificado en base a los resultados obtenidos y evaluado de nuevo. Se demuestra que las modificaciones mejoraron la puntuación de la evaluación significativamente y se destaca la importancia de la incorporación de usuarios "reales" durante todo el proceso de desarrollo.*

*This paper describes the experiences during the user interface design process of a Web course for the self- study of Fortran 90, a programming language. A systematic and quantitative analysis was performed. The course was evaluated by a group of 12 users. It was then modified based on the results obtained and re-evaluated. It is shown that the modifications improved the response values of the evaluation significantly, thus demonstrating the importance of the incorporation of "real" users during the whole development process.*

*DESCRIPTORES: Evaluación Normativa, World Wide Web, enseñanza distribuida, nuevas tecnologías de enseñanza, Internet, diseño gráfico, interfaz hombre-ordenador, multimedia, hipermedia, Fortran 90.*

## 1. Introducción.

Con el crecimiento de Internet y la informatización de universidades, colegios y otras instituciones de enseñanza se incrementa también el uso del World Wide Web (WWW) para acompañar y completar el proceso de aprendizaje y la enseñanza. Entre sus características más destacables, comparado con la enseñanza tradicional, cuentan la posibilidad de incluir multimedia (vídeo, audio, simulaciones, etc.), el formato hipertexto con la opción de enlazar sus componentes libremente y crear una estructura del contenido no lineal, y el acceso por ordenador, es decir la presentación en una pantalla de monitor y el manejo con ratón. Por lo tanto, el material educativo preparado para páginas Web requiere un tratamiento diferente, teniendo en cuenta dichas características junto con los objetivos del curso. Para facilitar el proceso existen una serie de guías de estilo y recomendaciones para el diseño de páginas Web en general (James, 1996; Eekma y Collis, 1996). Sin embargo, aunque la evaluación es una parte muy importante en el desarrollo de entornos de enseñanza, hay pocos estudios sistemáticos y cuantitativos de evaluación del interfaz de cursos Web (van Rennes y Collis, 1998) y menos aún sobre la incorporación de los resultados en el proceso de rediseño.

En los siguientes apartados se describe el proceso iterativo de evaluación (evaluación formativa) y rediseño del interfaz de un curso diseñado para el World Wide Web.

## 2. El curso.

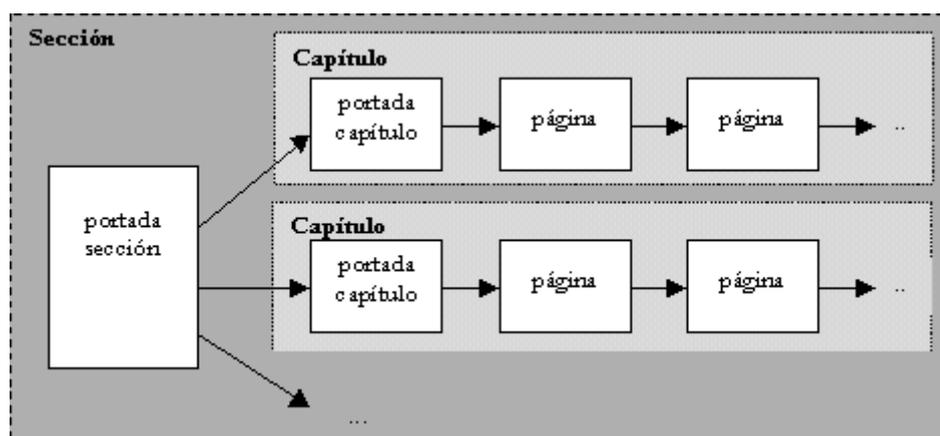
El objeto de la evaluación fue un curso Web sobre el lenguaje de programación Fortran 90 (<http://www.cesga.es:8800/F90/>). Este curso se desarrolló con el fin de facilitar a los usuarios del Centro de Supercomputación de Galicia (CESGA) el material necesario para el conocimiento del lenguaje con el fin de poder utilizar los recursos de ordenadores con procesadores escalar y vectorial del centro.

En la planificación del diseño se respetaron los siguientes aspectos:

1. **Acceso universal.** Los alumnos deberían acceder al material del curso desde cualquier ordenador independientemente del tiempo y lugar, y avanzar a ritmo individual.
2. **Enfoque.** El curso está pensando para el autoestudio. Por tanto, es necesario guiar al alumno por los temas relevantes con máxima brevedad y sencillez para limitar desvíos y evitar la desorientación.
3. **Interactividad.** El curso pretende enseñar el manejo práctico de Fortran 90, es decir, facilitar capacidades de programar en este lenguaje, por lo cual era importante incluir la posibilidad de practicar con el código.

El World Wide Web parece el medio idóneo para el acceso universal y la inclusión de material interactivo. Además, la red ATM de banda ancha de que dispone el Centro de Supercomputación de Galicia y el acceso directo a ella desde las aulas de Tele-enseñanza de las universidades gallegas (Engel y Casares, 1998) eliminaron el límite de velocidad de transmisión que impide normalmente el desarrollo de grandes aplicaciones para Internet.

Para guiar al usuario se eligió como metáfora en la que se basa la estructura del curso la de un libro. El contenido del curso se distribuye en "secciones", las secciones se componen de "capítulos". Cada capítulo consiste en una serie de "páginas" (cuadro 1).



Cuadro 1. Curso Web de Fortran 90 – estructura de contenidos

Según el concepto elegido se crearon páginas de los siguientes tipos:

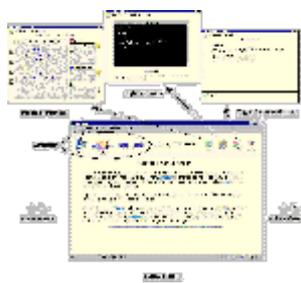
\* **Página de texto.** Se procuró dividir el texto en cantidades adecuadas y no llenar más que una pantalla.

\* **Página de ejemplo.** Para visualizar los efectos de comandos se presentan programas en Fortran 90 con explicaciones que acompañan cada línea del código. El alumno se puede mover libremente por el programa. Además es posible traer una copia del código al propio ordenador para uso individual.

\* Página de autoevaluación. Contiene preguntas sobre el respectivo contenido del capítulo en formato *multiple choice*. Al seleccionar una respuesta el alumno recibe contestación inmediata sobre el resultado.

\* Página de tarea. En este tipo de página se proponen ejercicios de programación relevantes para el tema del capítulo. A través de una conexión (*telnet*) a un ordenador remoto se da acceso a un compilador de Fortran 90, con lo cual se puede programar, compilar y ejecutar un código fuente.

El hilo conductor lo forman las páginas de texto. Desde ahí se accede a los otros tipos de páginas (tareas, ejemplos, autoevaluación; gráfico 1).



Como componentes gráficos la página de texto contiene iconos de navegación para moverse por el contenido e ir a los índices. También presenta iconos para enlazar a páginas de los otros tipos. No hay enlaces dentro del texto, ni tampoco existen enlaces a recursos externos, es decir, ajenos al propio curso.

Los futuros usuarios del curso son en su mayoría alumnos de cursos avanzados de químicas, físicas, matemáticas, etc. así como personal investigador de las universidades gallegas, por lo cual se presuponen conocimientos básicos de informática y de manejo de ordenadores.

### 3. La evaluación.

Una vez terminado el diseño y desarrollo del prototipo del curso y detectado y corregido la mayoría de los fallos técnicos del sistema, el objetivo de la evaluación era comprobar su "usabilidad" ("*usability*") (Shackel y Richardson, 1991) con un pequeño grupo de usuarios "reales". La usabilidad se refleja sobre todo en la facilidad de manejo de los componentes del sistema respecto al resultado deseado. Proyectos previos demuestran que la dificultad de entender y operar el sistema, la ineficacia en el uso, características inapropiadas para el aprendizaje, un comportamiento impredecible y poco fiable y una apariencia desagradable e incomoda en el manejo cuentan entre los problemas más frecuentemente mencionados por parte de usuarios. Para conseguir un producto de alta calidad es necesario incorporar pruebas de usabilidad, aplicar la técnica de diseño iterativo e integrar a los usuarios durante todo el proceso de desarrollo (Smith y Mayes, 1996).

En el caso del curso Web de Fortran 90 era importante comprobar la facilidad del uso intuitivo del sistema, ya que bajo condiciones reales el usuario tiene que ser capaz idealmente de estudiar con el material por su cuenta y sin ayuda externa.

Para realizar la evaluación se eligió un grupo de usuarios con un perfil similar al de los usuarios destino del curso, que trabajaron con la aplicación y respondieron a una serie de preguntas. Para

motivar la interacción con el material del curso y provocar experiencias comparables se dio a los usuarios una lista de tareas para realizar (cuadro 2).

- |  |
|--|
| <ol style="list-style-type: none"><li>1. elegir una sección del curso</li><li>2. visitar todas las páginas de la sección</li><li>3. examinar los ejemplos</li><li>4. contestar a las preguntas de la autoevaluación</li><li>5. editar y compilar un pequeño programa</li></ol> |
|--|

**Cuadro 2. Tareas a realizar por los usuarios evaluadores.**

Al terminar las tareas se recogieron las opiniones, impresiones y comentarios de los usuarios evaluadores mediante un cuestionario electrónico con 22 preguntas (<http://www.cesga.es:8800/DEMOS/EvalFortran.html>).

Las preguntas se centraron en los siguientes aspectos:

- i. la impresión general del curso,
- ii. el diseño de las páginas en general,
- iii. el manejo de las páginas específicas de cada tipo.

Se diseñaron las preguntas con el enfoque a las categorías "*instructional adequacy*" y "*cosmetic adequacy*" (adecuado respecto a la enseñanza y a la presentación) de análisis de software educativo (Hannafin y Peck, 1988). Se basan en las recomendaciones de evaluación de interfaces de QUIS (*Questionnaire for User Interface Satisfaction*) (Shneiderman, 1993). Con el fin de recibir respuestas enfocadas, adquirir unas medidas concretas y derivar sugerencias para mejorar la aplicación, se precisaron respuestas a través de una escala semántica de 7 puntos según Likert (1932) (cuadro 3). Para cada aspecto existía además la posibilidad opcional de comentarios libres.

Marque su impresión general del sistema:								
	1	2	3	4	5	6	7	
Terrible	<input type="radio"/>	Maravilloso						

**Cuadro 3. Ejemplo de pregunta del cuestionario.**

Todas las personas recibieron una introducción individual sobre el proyecto y el fin de la evaluación. También estaba una persona presente para responder en caso de preguntas, dudas o fallos del sistema. Adicionalmente, se tomó nota de problemas que se comentaron durante el proceso de la evaluación. No existiendo un límite de tiempo, las pruebas duraron entre 30 y 60 minutos por persona.

La evaluación se realizó en dos fases. En la fase 1 el prototipo del curso fue evaluado por un grupo de usuarios. Basado en los resultados se modificaron algunos aspectos de la aplicación. En la fase 2 se replicó la evaluación en un proceso idéntico de la primera fase, pero con el curso modificado y con otros usuarios sin conocimientos previos del curso.

Los evaluadores del grupo 1 eran 12 alumnos y profesores de la Facultad de Informática, Universidad de A Coruña (edades 21-32; 11 hombres, 1 mujer). Los evaluadores del grupo 2 eran

12 alumnos de la Facultad de Matemáticas, Universidad Santiago de Compostela (edades 19-26; 11 hombres, 1 mujer).

## 4. Resultados.

Un resumen de los resultados se presenta en la tabla nº 1. Se da el valor medio de todas las respuestas por grupo, basado en los valores obtenidos de la escala según Likert, de 1 (peor) a 7 (mejor). Por tanto, el valor "neutro" es el 4 y cuanto más alto el valor, más positiva la valoración.

A continuación, se describen las dos fases de la evaluación por separado.

### **Fase 1 (Grupo 1)**

En general, las opiniones sobre el curso fueron positivas, es decir, los valores medios superaron en más de una unidad al valor neutro (4). Negativamente solo destacaron dos parámetros ("Significado de iconos" y "Realización de tareas") con 4,2 y 4,0 respectivamente, indicando que el significado de los iconos no era suficientemente preciso y las tareas no siempre se podían realizar sin problemas.

En consecuencia, y juntamente con las observaciones del personal técnico durante las sesiones de la primera fase, se identificaron dos fuentes de problemas de usabilidad como más destacables:

**A. Navegación.** Como se indica anteriormente, las páginas de texto formaban el hilo conductor del curso, desde donde se accedía a los otros tipos de páginas (ejemplos, tareas y autoevaluación) como complemento a través de los iconos de la barra de navegación. Sin embargo, las páginas de ejemplos y autoevaluación no se cargaban en otra ventana adicional, sino que se abrían en el área designada para las páginas de texto, es decir, "encima" de la misma y por tanto, ocultándola. A parecer, esto resultó un concepto de diseño poco intuitivo y causante de confusión, a juzgar también por los comentarios y observaciones durante las pruebas.

Además, debido a este planteamiento, había que introducir la posibilidad de volver desde las páginas de autoevaluación y de ejemplos a la página de texto. El icono elegido para dicha función, una flecha señalando para abajo, creó mucha confusión al no aplicarse el concepto "linear" de "avance de página". Parece que una flecha hacia atrás hubiera sido más lógica como indicaron algunos de los usuarios. (Sin embargo, la flecha para atrás ya se utilizaba para la navegación por las páginas de texto, por lo cual esto no era una solución aplicable.)

Impresión general del curso		Grupo1	Grupo2	Diferencia
Terrible	Maravilloso	5,5	5,6	+0,1
Frustrante	Satisfactorio	5,9	6,3	+0,4
Aburrido	Estimulante	5,3	5,7	+0,4
Difícil	Fácil	6,3	5,3	-1,0 <sup>b)</sup>
Rígido	Flexible	5,3	5,0	-0,3
Feo	Atractivo	5,8	5,8	± 0

Páginas en general – organización			Grupo1	Grupo2	Diferencia
Estructura en pantalla	Confusa	Clara	5,8	6,2	+0,4

Secuencia de páginas	Confusa	Clara	5,7	6,0	+0,3
Organización de contenidos	Ilógico	Lógico	6,1	6,1	± 0
Cantidad de información por página	Inadecuada	Adecuada	5,8	5,3	-0,5
<b>Páginas en general – navegación</b>					
Siguiente página	Impredecible	Predecible	5,5	5,2	-0,3
Ir a la página anterior	Imposible	Fácil	5,9	6,8	+0,9
Significado de iconos	Ambiguo	Preciso	<b>4,2</b>	6,0	+1,8 <sup>a)</sup>

<b>Página de texto</b>			<b>Grupo1</b>	<b>Grupo2</b>	<b>Diferencia</b>
Tipografía	llegible	Perfectamente legible	6,2	6,2	± 0
Color y fuente letras	Confuso	Ayuda a entender	6,3	5,7	-0,6
Contenido	Incomprensible	Fácil de comprender	5,8	5,3	-0,5
<b>Página de ejemplos</b>					
Aprender a manejar	Difícil	Fácil	6,4	6,4	± 0
Interpretación	Difícil	Fácil	6,3	5,9	- 0,4
<b>Página de autoevaluación</b>					
Aprender a manejar	Difícil	Fácil	6,5	6,3	- 0,2
Interpretación de la respuesta	Difícil	Fácil	5,8	5,8	± 0
<b>Página de tareas</b>					
Aprender a manejar	Difícil	Fácil	5,2 (*)	5,8	+0,6
Tareas se pueden realizar (técnicamente) sin problemas	Nunca	Siempre	<b>4,0 (*)</b>	5,7	+1,7 <sup>a)</sup>

(\*) muestra solo de 11 personas

<sup>a</sup>  $p \leq 0.025$

<sup>b</sup>  $p \leq 0.05$

**Tabla 1. Valores de respuestas al cuestionario por grupo y diferencias entre ellos (Grupo2 – Grupo1)**  
**Basados en escala según Likert de 1 (peor) a 7 (mejor);  $n_{(\text{Grupo1})} = 12$ ,  $n_{(\text{Grupo2})} = 12$**

**B. Página de tareas.** No existía ninguna guía intuitiva del entorno para practicar con el código fuente. El interfaz y las herramientas para trabajar con los programas de Fortran 90 eran demasiado incomprensibles e incómodos de utilizar, como comentan estos usuarios:

*"Para realizar las tareas se hace necesario disponer de conocimientos acerca del uso de comandos UNIX, de un editor de textos ... y saber utilizar el propio compilador de Fortran 90, cuyo uso no se explica ... hasta llegar a la primera tarea."*

*"Considero que debería de haber en el curso ...[una información] donde se indicara como funciona el entorno donde se realizan las tareas, manejo de un editor, manejo del compilador..."*

En base de los problemas identificados se modificaron los siguientes componentes del curso:

#### Modificaciones en navegación:

Objetivo: mejorar el soporte del concepto no lineal a través del diseño: diferenciar entre hilo conductor y otros componentes (páginas) complementarios.

- \* añadir texto a todos los iconos
- \* eliminar icono de la flecha para abajo
- \* cargar páginas de ejemplo y autoevaluación en propia ventana aparte
- \* unificar el diseño de las ventanas en que se cargan las páginas tarea, ejemplo y autoevaluación, dándoles el mismo tamaño, borde, etc.

#### Modificaciones en página de tareas:

Objetivo: facilitar el trabajo autónomo con la aplicación.

- \* elaborar y añadir página de ayuda para las tareas
- \* añadir componentes adicionales (barra desplazamiento, botones)

### **Fase 2 (Grupo 2)**

Para la mayoría de los parámetros evaluados el segundo grupo respondía también con valores relativamente altos, es decir, mejores que 4 (tabla 1). Sin embargo, las puntuaciones de los parámetros "Significado de iconos" y "Realización tareas" suben a un valor de 6,0 y 5,7 respectivamente, resultando diferencias con el primer grupo de +1,8 y +1,7.

Se analizó, si dichas diferencias eran estadísticamente significativas. Para ello se aplicó el test  $\chi^2$  (sin corrección de continuidad; Bradley y otros, 1979), asumiendo que el pequeño tamaño de la muestra no permitía hablar de una distribución normal. Se compararon las distribuciones de los valores de la escala de Likert clasificados en tres categorías (1-3, 4 y 5-7) antes y después de las modificaciones.

Las dos diferencias consideradas resultaron significativas ( $p \leq 0,025$ , grados de libertad = 3), lo cual se puede interpretar que las manipulaciones resultaron en una notable mejora del curso respecto a las características evaluadas. Merece destacar que otro parámetro relacionado con el

tema de navegación ("Ir a página anterior") también subió a un valor muy alto de 6,8 (sin embargo, la diferencia no es significativa).

## Discusión.

La experiencia demuestra que los usuarios evaluadores tenían una actitud positiva hacia las nuevas tecnologías de enseñanza. Esto se puede observar en las "buenas notas" y comentarios favorables que recibió el curso por parte de los dos grupos de usuarios, y también se manifestó en una gran tolerancia hacia los fallos del sistema (en algunas ocasiones se quedó "colgado" y había que reiniciar el equipo). Las posibilidades de interactuar y aprender de una forma no tradicional a través de tecnologías educativas modernas crean una gran fascinación entre los alumnos:

"... me parece que hasta el momento no hay una manera de aprender mas estimulante que esta, por lo menos para un lenguaje de programación."

"Una muy interesante iniciativa que espero llegue a buen término..."

En concreto, y con ayuda de una cuantificación de resultados la prueba de usabilidad ayudó a identificar dos problemas del diseño del curso que impedían el trabajo con el material. Se mostró que los cambios introducidos eran adecuados ya que resultaron en una mejor evaluación en la segunda fase.

Una excepción fue la evaluación del parámetro "Impresión general", en donde la diferencia obtenida fue de -1,0 y estadísticamente significativa, indicando que el grupo 2 lo encontraba más difícil trabajar con el curso que el grupo 1. La explicación de este fenómeno podría ser una influencia indirecta causada por las modificaciones, pero debe buscarse probablemente en el diseño de la pregunta. No se especificó de forma precisa, a qué se hacía referencia con "difícil" y "fácil". Por las conversaciones durante las pruebas se sacó la impresión de que los usuarios del grupo 2 incluían en este concepto el contenido del curso, mientras los que los del grupo1 se centraron más en su diseño y manejo. El diseño del cuestionario debe ser muy bien pensado y las preguntas lo más precisas posibles. Varios autores recomiendan una pre-evaluación con el único fin de comprobar la calidad del cuestionario (Shneiderman, 1993).

El análisis intentó demostrar el valor de un método sistemático para la evaluación y el rediseño de material Web para la enseñanza, incluso con grupos de usuarios pequeños. La estrategia elegida reveló faltas de diseño anteriormente no detectadas y ayudó así a mejorar el producto respetando las necesidades de usuarios "reales". Con el incremento rápido de uso del Web para la enseñanza y aprendizaje habrá que enfrentarse a preguntas como: ¿Cuáles son los requerimientos particulares de los estudiantes como usuarios? ¿En qué se diferencian de otros usuarios del Web? En este sentido será necesaria la elaboración de guiones de diseño específicos para uso pedagógico, basados en análisis sistemáticos como el presentado en este trabajo.

## Referencias bibliográficas.

BRADLEY, D.R. y otros (1979). Type I error rate of the chi square test of independence in rc tables that have small expected frequencies. **Psychological Bulletin**, **86**, 1200-1297.

EEKMA, A.. y COLLIS, B. (1996). Design guidelines for WWW-based course environments. **Telektronikk (Norwegian Telecom Research Journal)**, **3/4**, 44-51.

ENGEL, C. y CASARES, J. (1998). Remote Learning over Broadband Networks. A pilot experience with ATM, en: **Proceedings of the**

**Fourth International Conference on Computer Aided Learning and Instruction in Science and Engineering CALISCE98**, Göteborg, Suecia, Junio 15-17, 1998.

HANNAFIN M. J. y PECK, K. L. (1988). **The Design, Development, and Evaluation of Instructional Software**. Macmillan, New York, 1988.

JAMES, D. (1996). Design Methodology for a Web-Based Learning Environment. <http://www.lmu.ac.uk/lss/staffsup/desmeth.html>

LIKERT, R. (1932). A Technique for the Measurement of Attitudes. **Archives of Psychology**, 140.

SHACKEL, B. y RICHARDSON, S. (eds) (1991). **Human Factors for Informatics Usability**. Cambridge University Press, Cambridge, UK.

SHNEIDERMAN, B. (1993). **Designing the User Interface. Strategies for Effective Human-Computer Interaction**. Segunda edición, Addison Wesley, Reading, MA..

SMITH, C. y MAYES, T. (1996). Usability Guide. Telematics Applications for Education and Training. Versión 2.1. **Cambridge Learning Systems**.

VAN RENNES, L. y COLLIS, B (1998). User Interface Design for WWW-Based Courses: Building Upon Student Evaluations. en: **Proceedings of ED-MEDIA/ED-TELECOM'98, Freiburg (Alemania)**.

Parte del trabajo fue financiado por la Comisión Europea (DEMOS ET-1007) y la CICYT (TEL97-0435).