

Silvia Tamayo Moreno
Diana Pérez Marín

**Evaluación heurística de Agente
Conversacional Pedagógico
Dr. Roland para el nivel de
Educación Secundaria**

Número 2017-18

Serie de Informes Técnicos DLSI1-URJC
ISSN 1988-8074
Grupo Docente de Lenguajes y Sistemas Informáticos I
Universidad Rey Juan Carlos

Índice

1	Introducción.....	5
1.1	Evaluación heurística, método de evaluación de la Interacción Persona Ordenador	5
2	Evaluación heurística del agente Dr. Roland: Educación Secundaria.	8
3	Conclusiones.....	13
	Referencias	13

Evaluación heurística de Agente Conversacional Pedagógico Dr. Roland para el nivel de Educación Secundaria

Silvia Tamayo Moreno y Diana Pérez Marín

Dto. de CC Comp., Arq. Comp., LSI y EIO, Universidad Rey Juan Carlos,
C/ Tulipán s/n, 28933, Móstoles, Madrid
silviatamayomoreno@gmail.com, diana.perez@urjc.es

Abstract. In this paper the heuristic evaluation of the pedagogical conversational agent Dr. Roland is carried out for the level of Secondary Education following the Nielsen heuristics. The heuristic evaluation is one of the inspection methods of the techniques of evaluation of the Human Computer Interaction.

Keywords: Evaluación heurística, Agente Dr. Roland, Educación Secundaria, Interacción persona Ordenador.

1 Introducción

En primer lugar se va a hacer alusión a la interacción persona ordenador (IPO). Seguidamente a técnicas de evaluación de la IPO [1], centrandó la atención en los métodos de inspección. Concretamente en la evaluación heurística.

A continuación se realizará la evaluación heurística de Agente Conversacional Pedagógico Dr. Roland [23, 24] para el nivel de Educación Secundaria.

1.1 Evaluación heurística, método de evaluación de la Interacción Persona Ordenador

Los sistemas interactivos son dispositivos hardware y software que actúan de interconexión entre las personas, permitiéndoles interactuar, y favoreciendo la realización de tareas y el alcance de los objetivos propuestos.

Con el desarrollo de los sistemas interactivos, y para tener una mejor comprensión de las relaciones persona-ordenador y mejorarla, surge la interacción persona ordenador (IPO). La IPO, conocida en la comunidad internacional como Human-Computer Interaction (HCI) o Computer-Human Interaction (CHI), según el grupo SIGCHI (Special Interest Group on Computer-Human Interaction, [2]) de ACM: “*Es la disciplina relacionada con el diseño, evaluación e implementación de sistemas informáticos interactivos para el uso de seres humanos, y con el estudio de los fenómenos más importantes con los que está relacionado*” [1].

Constantemente se presenta la situación en la que se diseñan productos que no se prueban con quienes van a ser los usuarios, sin prestar atención por tanto a la usabili-

dad de los productos y el software, como a las pruebas con usuarios. Los motivos suelen ser los costes y cambios que pueda suponer dicha evaluación. Se centran los esfuerzos y atención en la funcionalidad, y se tiene el pensamiento de que si quienes diseñan y programan, pueden usar el sistema y les gusta, es suficiente [3].

Evaluar algo implica hacer pruebas de funcionamiento y comprobar si cumple las expectativas, necesidades y requisitos. En el ámbito de los sistemas interactivos es clave y es importante que se realice durante todo el proceso de desarrollo. Los resultados de ello, obtenidos mediante la aplicación de una serie de técnicas, se pueden emplear para mejorar los sistemas. Lorés et al. [4] define la evaluación como “*La actividad que comprende un conjunto de metodologías y técnicas que analizan la usabilidad y/o la accesibilidad de un sistema interactivo en diferentes etapas del ciclo de vida del software*”.

Los prototipos se hacen para comprobar si el sistema reúne características que lo hagan usable para los usuarios finales. La evaluación del diseño conlleva comprobar si se adapta a lo esperado, permitiendo realizar las tareas, y comprobar el impacto en el usuario en aspectos como la facilidad de aprendizaje, manejabilidad, identificación de áreas que necesiten que sea recordada gran cantidad de información, etc. Con ello se identifican posibles problemas específicos del diseño, como todo aquello que cause resultados no esperados o confusos al usuario estando en su contexto de uso [5]. Según Dix [6] los objetivos de la evaluación son la comprobación de la funcionalidad del sistema, del efecto de la interfaz y la identificación de problemas específicos del sistema.

Los sistemas interactivos pueden ser evaluados en cualquier contexto que lo permita, donde se reúnan usuarios y evaluadores, desde espacios equipados como laboratorios, salas de reuniones o hasta en el entorno en el que los usuarios desarrollan sus tareas.

El abanico de métodos de para evaluar los diferentes aspectos es muy amplio, existiendo solapamiento entre algunos de ellos en cuanto a las actividades que desarrollan, no son independientes [7]. La elección de unos u otros depende de los costes y lo que se obtendrá con su uso.

En cuanto al tipo de técnica de comprobación utilizada, se distinguen:

- ***Métodos de inspección:*** Agrupa una serie de métodos que cuentan con expertos, llamados evaluadores, que explican el grado de usabilidad de un sistema. Se basan en la inspección de aspectos de la interfaz del sistema relacionados con la usabilidad ofrecida a los usuarios. Se considera opiniones, juicios o informes de los inspectores sobre elementos específicos de la interfaz como factor fundamental de la evaluación [8]. Permiten identificar, clasificar y contabilizar un gran número de errores potenciales de usabilidad a precio relativamente bajo, contribuyendo a ello el hecho de no usar usuarios [8, 9].
- ***Métodos de indagación.*** Consiste en alcanzar el conocimiento de una cosa mediante conjeturas y señales. Se llevan a cabo, hablando con los usuarios,

observándolos, usando el sistema en tiempo real y observando las respuestas a preguntas, ya sea orales o escritas. La información sobre gustos del usuario, necesidades o identificación de requisitos es importante, especialmente en fases tempranas del proceso de desarrollo.

- **Métodos de test:** En ellos usuarios, normalmente representativos, usan el sistema o prototipo para tareas concretas. Los evaluadores usan los resultados para saber la forma en la que la interfaz de usuario da soporte a los usuarios con sus tareas.

Como parte de los métodos de inspección se encuentra la evaluación heurística.

Evaluación heurística

Los métodos de evaluación por inspección como la evaluación heurística, que inicialmente fueron propuestas por Molich y Nielsen [10], son técnicas que no requieren la participación de usuarios, por lo que son más económicas [11].

En el diseño son los expertos quienes lo inspeccionan y analizan buscando problemas potenciales, comprobando los requisitos que tiene que cumplir el diseño para facilitar la comprensión y el uso por el usuario final, en definitiva, los principios de diseño usable. En relación a los expertos, es aconsejable que sean 3-5, para poder realizar un informe conjunto con las opiniones ya que la opinión de menos evaluadores sería sesgada, y contar con más evaluadores supondría un coste de la evaluación sin justificación [8]. Además, cuanto más experiencia se tenga será más fácil que detecten más problemas [12, 4]. Cada uno de ellos, analizará independientemente el diseño, informando de los problemas que se detecten. Finalmente, se pondrán en común y hará un informe conjunto [13].

Diversos autores han propuesto reglas de diseño que pueden emplearse como heurísticas (las 10 reglas heurísticas de usabilidad de Nielsen [14], las ocho reglas de oro de Schneiderman [15], o Tognazzini [16]). Otros ofrecen guías formadas por criterios heurísticos más específicos, más fáciles de aplicar por evaluadores que no son expertos. Además, existen diversos casos prácticos de aplicación de evaluación heurística que pueden ser consultados [17, 18, 19].

La evaluación heurística puede realizarse en cualquier momento del ciclo de desarrollo del proyecto (antes de las pruebas de usuarios es un buen momento), y dependiendo de cuando sea, los principios o criterios a comprobar podrían variar. En relación a las limitaciones y problemas [12]:

- Detecta más problemas menores de usabilidad, detecta menos problemas mayores que otras metodologías, por lo que no sustituye a la realización del test de usuarios.
- Puede reportar alarmas falsas, identificando como problema algo que no lo es.
- Para que ofrezca resultados importantes es necesario que participen varios evaluadores.

2 Evaluación heurística del agente Dr. Roland: Educación Secundaria.

Se ha llevado a cabo la Técnica Inspección Heurística, mediante la evaluación de las 10 heurísticas de usabilidad de Nielsen [14].

Heurística 1. Visibilidad del estado del sistema: Se aprecia en la Fig. como el sistema mantiene informado a los usuarios, mostrando su estado y ofreciendo respuesta y en un tiempo razonable. Puede comprobarse, por ejemplo, cómo el sistema informa del resultado de la acción previa del usuario, y cuando está en estado de espera pendiente de que realice una nueva acción el usuario, le ha informado previamente de la situación, y además le ofrece los medios para que el usuario responda, ya sea mediante selección de opciones o para que inserte texto. De esta manera, el usuario conoce que el siguiente en realizar alguna opción es él. Además, en todo momento se muestra el diálogo anterior, como se puede apreciar en la Fig..

La respuesta no es correcta ¿Quieres ver o introducir la o las palabras de nuevo?
ver ▾

La respuesta correcta es: diagonal de un rectángulo de lados, igual al lado de un cuadrado, diagonal
¿Entiendes qué es lo que te piden?
no ▾

¿Crees que puede haber algun pronombre importante? Introdúcelos, si crees que no hay, no introduzcas nada
.....

La respuesta no es correcta ¿Quieres ver o volver a introducir el o los pronombres?
ver ▾

La respuesta correcta es: cuánto
¿Entiendes qué es lo que te piden?
no ▾

En cualquier caso, para tener la información, escribe qué es lo que no entiendes
qué me están pidiendo que calcule


Habla con tu profesor sobre este ejercicio
¿Desea realizar otro ejercicio?
▾

Fig. 1. Interacción entre el usuario y agente Dr. Roland

H2. Relación entre el sistema y el mundo real: Se puede comprobar en la Fig. que el sistema usa un lenguaje entendible por los usuarios.

H3. Control y libertad del usuario: el usuario cuenta con una salida ante estados no deseados al que accedieron. Como por ejemplo, al tratarse de una aplicación web, siempre puede salir cuando quiera, simplemente cerrándola, o volver atrás. Además, dentro de las posibilidades cubiertas por el sistema, que es quien llevará el control principal de la interacción, permite al usuario libertad en determinadas acciones, como por ejemplo decidir si quiere realizar otro ejercicio, como puede apreciarse en la parte final de la Fig. .

How does information that you type in get passed from the keyboard to the hard disk?



viewing data, and for long-term storage of data.
Tutor: Now for something different.
Tutor: The figure you see shows that the CPU communicates with a number of devices. There are the input devices, random access memory, storage devices, and output devices. So, here's your question. How does information that you type in get passed from the keyboard to the hard disk?

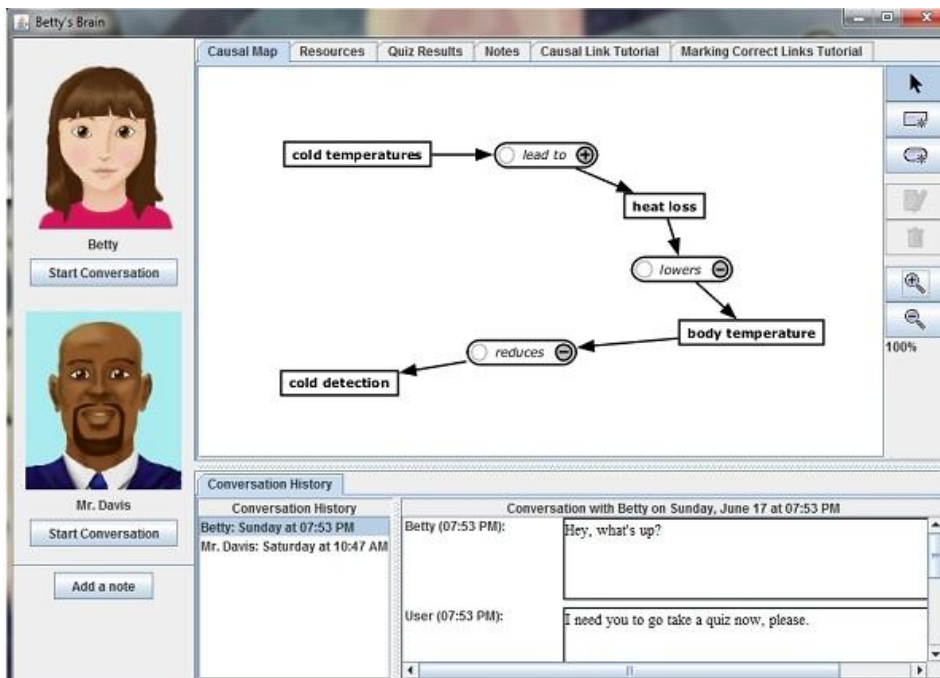
through the CPU

Submit Settings...

Fig. 2. Autotutor [20]

Betty's Brain

Causal Map Resources Quiz Results Notes Causal Link Tutorial Marking Correct Links Tutorial



Betty

Start Conversation

Mr. Davis

Start Conversation

Add a note

Conversation History

Conversation with Betty on Sunday, June 17 at 07:53 PM

Betty (07:53 PM): Hey, what's up?

User (07:53 PM): I need you to go take a quiz now, please.

Fig. 1. Interfaz del agente Betty [21]



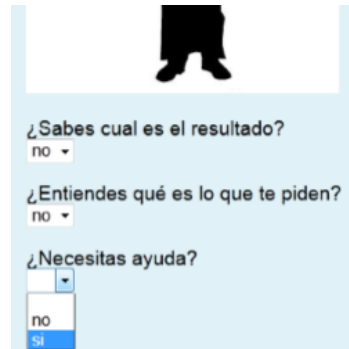
Fig. 4. Interfaz del agente Laura [22]

H4. Consistencia. En relación a la consistencia externa, que se refiere a la consistencia con otros agentes, puede comprobarse que la mantiene respecto algunos agentes como el Autotutor (Figura 2), Betty (Figura 3) o Laura (Fig.4). Mantienen una distribución de los elementos similar, ofreciendo la imagen del agente en la parte izquierda, realizan la interacción con el usuario por medio de cuadros de texto en los que el usuario introduce información, y el agente le va respondiendo, lo que también es compartido e al usuario en el agente Dr. Roland.

Respecto a la consistencia interna, que se refiere a la existente entre pantallas del propio agente, puede apreciarse, que tanto en la parte del profesor como en la del alumno, se mantienen los mismos patrones, la misma tonalidad y estilos parecidos. Esto puede apreciarse en la Figura 6.

H4. Estándares: En relación a los estándares, como en la metodología MEDIE y en el desarrollo del agente se ha seguido el DCU, se han abarcado lo relativo a la aplicación del DCU.

H5. Prevención de errores: el sistema trata de prevenir ciertos errores, especialmente a la hora de que el usuario inserte la información. Un ejemplo de ello es cuando las opciones de respuesta no son abiertas, sino que tienen un rango de valores, acotar las mismas ofreciendo las opciones disponibles en una lista desplegable, de tal manera, que el usuario solo tiene que seleccionar la opción que desee. Puede observarse en la Fig.5.



The image shows a light blue rectangular interface. At the top center, there is a black silhouette of a person's legs and feet. Below this, there are three questions, each followed by a dropdown menu:

- Question: ¿Sabes cual es el resultado? (Do you know what the result is?)
Dropdown menu: no (selected)
- Question: ¿Entiendes qué es lo que te piden? (Do you understand what they are asking you?)
Dropdown menu: no (selected)
- Question: ¿Necesitas ayuda? (Do you need help?)
Dropdown menu: si (selected)

Fig. 5. Ejemplo de respuestas mediante selección de opciones

H6. Reconocimiento antes que recuerdo: implica todo lo que te haga al usuario reconocer, y que mantenga patrones entre ello. En el agente se intenta dar visibilidad a objetos, acciones y opciones disponibles, de tal manera que el usuario no tenga que recordar la información. Por ejemplo, dispone de la opción de ayuda, que se encuentra a la vista del agente, en la que se explica al usuario (Figura 6). Otro ejemplo, es cuando se le ofrecen las opciones de respuesta disponibles en forma de selección (Figura 5). Ambos son ejemplo que evitan al usuario tener que recordar.

Soy Dr. Roland, bienvenido
Colegio San Diego y San Vicente de Paúl

Ayuda



Halla la longitud y el área de una circunferencia de diámetro 22 cm

¿Sabes cual es el resultado?
Si

¿Qué entiendes que te han pedido?
Hallar longitud y área de la circunferencia, considerando que el dato que me dan corresponde el diametro

Introduce el resultado
69'1 longitud, 380'11 área

La respuesta es correcta
Enhorabuena, has resuelto bien el ejercicio
¿Desea realizar otro ejercicio?

Fig. 6. Interfaz del estudiante con ayuda y selección de opciones en la respuesta

H7 Flexibilidad y eficiencia de uso: el sistema tiene cierta flexibilidad ya que permite que los usuarios adapten el sistema para usos frecuentes. Esto es algo que tiene sentido dada la particularidad de agente Dr. Roland que va dirigido a niños. La entrada se podría hacer tanto por teclado, como por ratón a la hora de responder a las preguntas y también la salida es flexible al poderse mostrar tanto en pantallas de ordenador, tabletas, móviles y pizarras digitales.

H8. Estética y diseño minimalista: los diálogos usados en la interacción contienen información concreta y relevante. Se ha tratado de no incluir información que no es útil y que puede llevar confusión.

H9. Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de errores. Lo que se ha tratado de hacer es prevenir el error. El agente responde a cada acción del usuario, habiéndose evitado la inclusión de errores, intentando redirigir la situación a alguna salida. No se encuentra todavía un plan de errores que detalle al usuario qué pasó o cómo recuperarse de ellos.

H10. Ayuda y documentación: el sistema ofrece ayuda, como puede comprobarse en la parte superior del mismo (Figura 6).

3 Conclusiones

Las heurísticas fueron revisadas por tres expertos, dos de ellos doctores en informática, para el agente Dr. Roland para el nivel de Educación Secundaria. Tras la evaluación de las mismas, puede concluirse que el sistema es usable porque, en mayor o menor grado cumple las heurísticas evaluadas.

Referencias

1. Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G. & Verplank, W. (1992). ACM SIGCHI Curricula for Human-Computer Interaction. Technical Report. ACM, New York, NY, USA.
2. SIGCHI Special Interest Group in Computer Human Interaction. Última fecha de consulta: 20-04-2017.
<http://www.sigchi.org/>
3. Granollers i Saltiveri, T., Lorés Vidal, J., & Cañas Delgado, J. J. (2012). Diseño de sistemas interactivos centrados en el usuario.
4. Lorés, J., Granollers, T. y Lana, S. (2002). Introducción a la interacción persona-ordenador. En J. Lorés (ed.) La interacción persona-ordenador. Lérida, 2001-2002, 20-40.
5. Saltiveri, G. (2007). MPIu+ a. Una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinares. Universitat de Lleida.
6. Dix, A. (1993). Human computer interaction. Prentice Hall, Englewood Cliffs, NJ.
7. Rodeiro, J. (2001). Representación y Análisis de la componente visual de la interfaz de usuario. Tesis doctoral. Universidad de Vigo.
8. Nielsen (1994), "Heuristic evaluation", Usability Inspection Methods, John Wiley & Sons, New York, NY.
9. Bias, R. & Mayhew D. (1991). Cost-justifying usability. IEEE Software.
10. Molich y Nielsen, "Improving a human-computer dialogue", Communications of the ACM, 3 (33), pp. 338-348, 1990
11. Hassan Montero, Y. & Ortega Santamaría, S. (2009). "Informe APEI sobre usabilidad".
12. M.P. González, A. Pascual, J. Lorés, "Evaluación Heurística", Lorés, J. (Ed.), 2006.
13. Manchón, E. Evaluación heurística (o por expertos) de la usabilidad. Última fecha de consulta: 20-04-2017
14. Nielsen, J. (1995). 10 usability heuristics for user interface design. Nielsen Norman Group, 1(1).
15. Schneiderman, B. (1986). Eight golden rules of interface design
16. Tognazzini, B. (2003). First principles of interaction design. AskTOG.
17. Marcos, M. C. (2006). Evaluación de la usabilidad en sistemas de información terminológicos online. Hipertext. net: Anuario Académico sobre Documentación Digital y Comunicación Interactiva, (4).
18. García Gómez, J. C. (2008). Análisis de usabilidad de los portales en español para personas mayores. No Solo Usabilidad, (7).

19. Candamil-Llano, M. and Guevara-Hurtado, A. F. "Análisis de sitios web universitarios colombianos mediante evaluaciones heurísticas para el proyecto del nuevo portal web Unicauca", 2009.
20. Person, N. K., & Graesser, A. C. (2000). Designing AutoTutor to be an effective conversational partner.
21. Leelawong, K. and Biswas, G. (2008) Designing Learning by Teaching Systems: The Betty's Brain System. In *International Journal of Artificial Intelligence in Education*.
22. Theodoridou, K. & Yerasimou, T. (2008). "Learning spanish with"laura": The role of an intelligent agent in a spanish language course," in *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, vol. 2008, no. 1, 2008, pp. 4907– 4912.
23. Tamayo-Moreno (2012). Propuesta de Desarrollo Centrado en el Usuario de un Agente Conversacional Pedagógico para la Comprensión Lectora de Ejercicios de Matemáticas a nivel escolar. Máster Universitario en Investigación en Sistemas Hardware y Software Avanzados de la Escuela Técnica Superior de Ingeniería Informática de la Universidad Rey Juan Carlos.
24. Tamayo-Moreno, S & Pérez-Marín, D. (2014). Diseño y gestión de diálogos educativos en agentes pedagógicos conversacionales para uso escolar. *Aplicaciones multidisciplinares de sistemas de diálogo*. López-Cózar Delgado, Ramón (Ed.), Coca.