

Silvia Tamayo Moreno
Diana Pérez Marín

**Evaluación del Agente
Conversacional Pedagógico
Dr. Roland para el nivel de
Educación Primaria mediante el
método de medida de las
prestaciones**

Número 2017-11

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	Medida de las prestaciones, método de evaluación de la Interacción Persona Ordenador.....	5
2	Evaluación del agente Dr. Roland para Educación Primaria mediante el método de medida de las prestaciones	7
3	Conclusiones.....	12
	Referencias	12

Evaluación del Agente Conversacional Pedagógico Dr. Roland para el nivel de Educación Primaria mediante el método de medida de las prestaciones

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 evaluation of the pedagogical conversational agent Dr. Roland is carried out for the level of Primary Education following the performance measurement method. The logging method is one of the test methods of the techniques of evaluation of the Human Computer Interaction.

Keywords: Medida de las prestaciones, Agente Dr. Roland, Educación Primaria, 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 test. Concretamente en el método de medida de las prestaciones.

A continuación se realizará la evaluación del Agente Conversacional Pedagógico Dr. Roland [11, 12] para el nivel de Educación Primaria usando el método de test de medida de las prestaciones.

1.1 Medida de las prestaciones, 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 usabilidad 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 test se encuentra el de medida de las prestaciones

Medida de las prestaciones

Este método se basa en tomar medidas sobre el rendimiento y otros tipos de aspectos subjetivos que afecten la usabilidad del sistema. Es importante la comprensión de lo que se pretende medir, se pueden seleccionar:

- Medidas de rendimiento, son medidas cualitativas, se puede contar las acciones, los comportamientos, número de personas, de errores, etc.
- Medidas subjetivas, es decir, percepciones de las personas, opiniones y juicios. Pueden ser cuantitativas o cualitativas.

Por ello, se necesita disponer de un sistema implementado o un prototipo que permita la evaluación de dichos aspectos. Los participantes del test son usuarios realizando tareas, se analiza la forma en la que usan el producto y se mide el tiempo que emplean. Es útil para ser realizado en laboratorios de usabilidad, aunque también puede hacerse en un entorno real. Es importante la selección de tareas que los usuarios tendrán que llevar a cabo, estas deberán demostrar problemas de usabilidad, siendo el criterio más importante para la selección que prueben problemas potenciales de usabilidad del producto. Deberán ser tareas sugeridas por la propia experiencia, ya que los desarrolladores tienen ideas sobre dónde encontrar problemas, qué partes del producto fueron más difíciles de diseñar y cuáles son los problemas que se tienen que probar [10].

2 Evaluación del agente Dr. Roland para Educación Primaria mediante el método de medida de las prestaciones

Se ha llevado a cabo la técnica de test medida de las prestaciones [10] para el agente Dr. Roland para Educación Primaria. El agente para dicho nivel puede verse, junto con un ejemplo de interacción entre el agente y el usuario en las figuras 1, 2, 3 y 4.



¿Cuántos jugadores forman un equipo de baloncesto?
OPCIONES: A) 3 B) 5 C) 10

¿Sabes cual es el resultado?
no ▾

¿Entiendes qué es lo que te piden?
no ▾

¿Necesitas ayuda?
si ▾

¿Qué frase crees que es la principal y que nos daría esta información?

Fig. 1. Agente Dr. Roland, ejemplo de interacción entre el estudiante y el agente (1)

¿Qué frase crees que es la principal y que nos daría esta información?
equipo de baloncesto

La respuesta es correcta

¿Entiendes qué es lo que te piden?
no ▾

¿Cuál o cuáles crees que son los verbos principales de la frase?
forman

La respuesta es correcta

¿Entiendes qué es lo que te piden?

Fig. 2. Agente Dr. Roland, ejemplo de interacción entre el estudiante y el agente (2)

¿Cuál o cuáles crees que son los verbos principales de la frase?

forman

La respuesta es correcta

¿Entiendes qué es lo que te piden?

si

¿Sabes resolver el ejercicio?

si

¿Qué entiendes que te han pedido?

Número de jugadores que forman un equipo

Fig. 3. Agente Dr. Roland, ejemplo de interacción entre el estudiante y el agente (3)

Introduce el resultado

¿Cuántos jugadores forman un equipo de baloncesto?

10
3
5

Enhorabuena, has resuelto bien el ejercicio

¿Desea realizar otro ejercicio?

Fig. 4. Agente Dr. Roland, ejemplo de interacción entre el estudiante y el agente (4)

En este apartado se va a hacer alusión a las medidas de las prestaciones del uso del agente Dr. Roland.

La Tabla 1 recoge una serie de medidas por experiencia, correspondiendo cada experiencia a cada uno de los colegios en los que ha sido puesto en práctica. Dos en este caso (Experiencias Pri-1 y Pri-2). Recoge números de ejercicios realizados (14 y 23 respectivamente), respondidos correctamente, tiempos de uso (35 y 55 minutos), tiempos medios de ejercicios (2,5 y 2,39 minutos) y una serie de porcentajes, como el

porcentaje de ejercicios respondidos correctamente sobre el total realizados (64,3% y 73,9%). En el primer colegio se usó con pizarra digital y en el segundo con ordenador.

Tabla 1. Evaluación - Medidas de las prestaciones en Primaria: medidas por sesión

Medidas	Experiencia Pri-1 (1 sesión)	Experiencia Pri-2 (2 sesiones)
Número Ejercicios totales hechos en cada experiencia	14	23
Número Ejercicios medios hechos en cada experiencia por sesión	14	11,5
Tiempo total de uso por experiencia (minutos)	35	55
Tiempo medio de uso por experiencia por sesión (minutos)	35	27,5
Dispositivo usado	Pizarra digital	Ordenadores
Número Ejercicios respondidos correctamente por experiencia	9	17
Número de Ejercicios respondidos correctamente por experiencia por sesión	9	8,5
Porcentaje de ejercicios respondidos correctamente sobre el total hechos por experiencia	64,286%	73,9%
Porcentaje de ejercicios respondidos correctamente sobre el total hechos por experiencia por sesión	64,286%	73,91%
Tiempo medio de uso por ejercicio por experiencia (minutos)	2,5	2,39
Tiempo medio de uso por ejercicio por experiencia por sesión (minutos)	2,5	2,39

Se recoge una serie de medidas por sesiones de cada una de las experiencias (Pri-1 y Pri-2) en la Tabla 2, entendiendo como experiencia aquello que se mencionó previamente, y como sesión cada uno de los días en los que se ha puesto en práctica el agente. Dos experiencias en este caso (Pri-1 y Pri-2), la primera (Pri-1) con una sesión y la segunda (Pri-2) con dos. Recoge los ejercicios realizados, tiempos de uso, dificultades en la interacción, porcentajes de ejercicios resueltos de voz o por teclado, uso de reintentos en los ejercicios, uso de la ayuda, ejercicios respondidos correctamente o entendimiento de los ejercicios.

Cabe destacar que de la ayuda se ha hecho uso en todos los ejercicios que no se han resuelto correctamente a la primera, coincidiendo de este modo, también con el porcentaje en los que se ha hecho uso del reintento en la respuesta. Añadir en el tema de los reintentos para Primaria, que dadas las características que tenían los ejercicios

en los que cuando respondías inmediatamente te mostraba el resultado de tu respuesta así como resaltada en verde la opción correcta (entre las existentes en el grupo test). Se ha considerado que no han sido respondidos correctamente aquellos ejercicios que en una misma realización de los mismos, no se han resuelto bien a la primera. Por ello, en este caso, matizar que el uso de reintentos en Primaria, pierde importancia por el motivo indicado.

El tiempo medio de uso por ejercicio por sesión (minutos), ha sido de 2,5 en Pri-1 y de 2,5 en la primera sesión de Pri-2 y de 2,31 en la segunda sesión de Pri-2. Se puede observar como en pri-2 en tiempo medio de uso de ejercicio se reduce, lo que puede suponer mayor destreza en el uso o resolución de ejercicios más rápida (por ejemplo sin reintentos). El porcentaje de ejercicios en los que se ha hecho uso de la ayuda del ejercicio es de 35,7%(Pri-1), 30% (Sesión 1 de Pri-2) y 23,1% (Sesión 1 de Pri-2) y para el porcentaje de ejercicios en los que se ha hecho uso de reintento en la respuesta es de 35,7% (Pri-1), 30% (Sesión 1 de Pri-2) y 23,1% (Sesión 1 de Pri-2). Se puede comprobar cómo en la segunda sesión de Pri-2 se redujo el porcentaje, lo que podría ir ligado a una mejora en los resultados y aprendizaje.

Tabla 2. Evaluación - Medidas de las prestaciones en Primaria: medidas por experiencia

Medidas	Experiencia Pri-1. Sesión 1	Experiencia Pri-2. Sesión 1	Experiencia Pri-2. Sesión 2
Número Ejercicios hechos en cada sesión	14	10	13
Tiempo total de uso por sesión	35	25	30
Número Ejercicios respondidos correctamente	9	7	10
Porcentaje de ejercicios respondidos correctamente sobre el total hechos	64,28%	70%	76,9%
Tiempo medio de uso por ejercicio por sesión (minutos)	2,5	2,5	2,31
Porcentaje de ejercicios en los que se ha hecho uso de la ayuda del ejercicio	35,7%	30%	23,1%
Porcentaje de ejercicios en los que se ha hecho uso de reintento en la respuesta	35,7%	30%	23,1%
Dificultad en la interacción (si/no)	Al inicio	Al inicio	No
Porcentaje de ejercicios en los que no entendían lo que se preguntaba	0%	0%	0%

Porcentaje de ejercicios resueltos de voz	No aplica	No aplica	No aplica
Porcentaje de ejercicios resueltos por teclado	100%	100%	100%

3 Conclusiones

El método de test de medidas de las prestaciones de la Interacción Persona Ordenador fue usado para la evaluación de Agente Dr. Roland para el nivel de Educación Primaria. Tras la evaluación puede concluirse, que el agente fue integrado en el aula y usado por los alumnos. De ello puede extraerse, que los estudiantes fueron capaces de interactuar con el agente, siendo el algoritmo RUANLP el que guía la interacción y el diálogo entre estudiantes y el agente.

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. Hassan Montero, Y. & Ortega Santamaría, S. (2009). "Informe APEI sobre usabilidad".
11. Tamayo-Moreno, S & Pérez-Marín, D. (2016). Analizando la interacción de estudiantes de educación Infantil y Primaria con un agente conversacional. Proceedings of IkaSABAR-GUIDE 2016, 9th International Conference. Gorka J. Palazio (ed.), servicio editorial de la Universidad del País Vasco.
12. Tamayo-Moreno, S & Pérez-Marín, D. (2016). Diseño e Integración en el Aula de Primaria de Agentes Pedagógicos Conversacionales. Actas de las Jornadas de Innovación y TIC Educativa (JITICE). In press.