

Dimensiones para Evaluación de Sistemas Tutores Inteligentes

Zulma Cataldi y Fernando J. Lage
zcataldi@posgrado.frba.utn.edu.ar, fernandojlage@yahoo.com.ar

Facultad Regional Buenos Aires Universidad Tecnológica Nacional. CABA.

RESUMEN

Un sistema tutor inteligente actúa como un tutor particular del estudiante, por lo que debe poseer libertad para actuar de acuerdo a las necesidades del estudiante. Los trabajos se centran actualmente en proveer una alternativa al tutor humano, cuando no puede dedicar más tiempo a sus estudiantes y es una opción para los estudiantes que buscan aprender en forma más autónoma. Por otra parte, a la hora de elegir, es necesario crear herramientas que permitan su evaluación.

Palabras claves: evaluación, criterios, indicadores, Sistemas Tutores Inteligentes

CONTEXTO

En esta comunicación se da cuenta de los avances del PID: *Modelado del tutor basado en redes neuronales para un Sistema Tutor Inteligente* Código 25/C099, correspondiente al Programa de Incentivos Secretaría de Políticas Universitarias del Ministerio de Ciencia y Tecnología de la Nación y se inscribe en el: *Programa “Electrónica, Informática y Telecomunicaciones”*.

1. INTRODUCCION

En la presente comunicación se analizan las categorías para evaluación de software educativo propuestas por diferentes autores. Para el caso de los STI se siguen lineamientos teóricos centrados en: aprender haciendo, la construcción del conocimiento, el conflicto cognitivo, la zona de desarrollo próximo, aprendizaje por descubrimiento, cambio conceptual, el aprendizaje para la comprensión y las normativas para diseño de interfaces. Se presenta una propuesta de evaluación basada en la actualización tecnológica de criterios de varios autores, se agregan criterios nuevos a las categorías existentes y se generan nuevas categorías de análisis. El dominio de aplicación hace que se deban tener en cuenta algunas características especiales como el diseño de los ejercicios y problemas y el sistema de autoevaluación y evaluación.

1.1. Evaluación de STI

Los programas para uso didáctico se evalúan a través de plantillas considerando los aspectos pedagógicos y didácticos, comunicacionales, organizativos, técnicos y económicos (Cataldi, 2001). También se realiza la evaluación contextualizada que surge como un complemento a las evaluaciones interna (realizada por el grupo de desarrollo) y externa (a través de especialistas en el tema que no participaron del desarrollo). En los programas didácticos, la evaluación generalmente se efectúa a partir de la consideración de criterios de calidad que se expresan a través de indicadores que se identifican en los materiales que se evalúan (Casanovas, 2005).

Un STI (Sistema Tutor Inteligente) actúa como un tutor particular del estudiante ya que como un entrenador humano, el cual puede actuar adecuándose a las necesidades de cada estudiante. Es decir, son: *“sistemas que modelan la enseñanza, el aprendizaje, la comunicación y el dominio del conocimiento del especialista y el entendimiento del estudiante sobre ese dominio”* (VanLehn, 1988).

Los resultados de la evaluación se suelen plasmar en plantillas que consideran diversos aspectos: identificación del producto, valoración del mismo de acuerdo con ciertos indicadores, evaluación global del producto con el agregado de comentarios (Casanovas, 2005). Para los materiales didácticos se buscan indicadores que sean fáciles de identificar y que representen las características que los materiales deben tener para cumplir con los criterios de calidad predeterminados, donde cada criterio podrá dar lugar a varios indicadores. En general se utilizan las listas de control que son instrumentos que facilitan los procesos de evaluación de los materiales didácticos, en las cuales aparecen los indicadores que se deben identificar de acuerdo con los criterios de calidad establecidos para cada tipo de material. Algunas de ellas, suelen proveer además una escala a través de la cual se puede valorar el nivel en que aparece la característica que representa cada indicador. Su nivel detallado dependerá de cada material y de la situación en particular. Los STI permiten la emulación de un tutor humano para determinar qué enseñar, cómo enseñar y a quién enseñar a través de un módulo del dominio: que define el dominio del conocimiento (ver Figura 1), un módulo del estudiante: que es capaz de definir el conocimiento del estudiante en cada punto durante la sesión de trabajo, un módulo del tutor: que genera las interacciones de aprendizaje basadas en las discrepancias entre el especialista y el estudiante y finalmente la interface con el usuario: que permite la

interacción del estudiante con un STI de una manera eficiente (conocimiento sobre cómo presentar los contenidos).

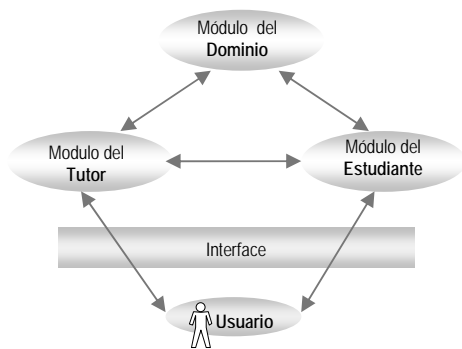


Figura 1: Interacción de los Módulos de un STI.

1.2. Categorías de análisis para la evaluación

Cabero y Sancho (1996) plantean una evaluación a través de criterios o categorías con sus respectivos indicadores y afirma que un problema tradicional en los medios de enseñanza consiste en determinar de qué manera pueden diseñarse para cumplir eficazmente la misión didáctica para la que fueron pensados a fin de facilitar los aprendizajes. Cabero (2001) propone una plantilla con categorías para evaluación de software educativo que se observan en la Tabla 1, pero, dado los avances tecnológicos desde la propuesta de las mismas, se hace necesario una actualización. Por otra parte, el dominio de cada área de aplicación requiere que se deban tener en cuenta algunas características especiales, tales como el diseño de los ejercicios, las situaciones problemáticas, el sistema de autoevaluación y el sistema de evaluación que poseen.

Para sustentar la elección de las nuevas categorías e indicadores, se tomarán como eje de análisis de las categorías adicionales los parámetros que se expresan en la Tabla 3, ellos sintetizan las teorías tomadas en cuenta y los representantes principales de cada una de ellas.

Calidad de contenidos
1. Nivel de actualización
2. Calidad científica
3. Conocimientos previos requeridos
4. Secuenciación y estructuración
5. Originalidad de presentación
Aspectos técnicos
6. Calidad del sonido
7. Calidad y tamaño de gráficos
8. Sincronización imagen-sonido
9. Calidad de uso de herramientas
10. Variedad de presentaciones
11. Calidad de animaciones
Motivación
12. Interés que despierta
13. Duración
Organización interna de la información

14. Inclusión de ejemplos y tutoriales
15. Síntesis de aspectos significativos
16. Velocidad de presentación
17. No redundancia
Costo económico
18. Relación costo-calidad
19. Relación costo-durabilidad científica
20. Relación costo-durabilidad física
Valor didáctico
21. Favorecimiento del proceso de aprendizaje
Claridad del programa
22. Claridad de la información
23. Claridad de explicaciones
24. Facilidad de manejo
25. Intuitividad
26. Menús de ayuda
Adecuación a los receptores
27. Adaptación al currículum
28. Adecuación a características de desempeño de los alumnos
29. Adecuación del vocabulario al ámbito profesional
30. Adecuación al nivel cognitivo alcanzado por el alumno
31. Adecuación de contenidos a los requerimientos académicos
Objetivos
32. Explicitación de los objetivos o logros esperados
33. Relación entre contenidos y evaluación
34. Relación entre contenidos y objetivos

Tabla 1. Plantilla de evaluación Cabero (2001).

Las teorías seleccionadas buscan explicar la metodología seguida a través de sus causas, efectos y leyes generales del aprendizaje.

Característica a considerar	Representantes	Descripción
Pedagogía activa, aprender haciendo	Dewey (1949,1950)	Una escuela cuyo centro y origen sean actividades constructivas a través de una tarea que se desarrolle en dos direcciones: una dimensión social de esta actividad constructiva y en contacto con la naturaleza.
La vida real en el aula	Métodos Dalton (Parkhurst, 1920) y Winnetka (Washburne, 1945)	Se basa en la actividad, individualidad y libertad siendo su objetivo principal desarrollar la vida intelectual. Se deja al alumno la posibilidad de escoger los trabajos y los momentos de realizarlos con actividades colectivas.
Leyes Gestálticas	Teoría de Gestalt (1912)	Esta corriente propone que el intelecto organiza los datos según leyes y formas llamadas Gestálticas. La experiencia sensible es un material desorganizado que proporciona contenidos; donde el orden, la relación y estructura son impuestos por las personas.
La construcción del conocimiento, conflicto cognitivo.	Piaget (1978)	Piaget delimitó el aprendizaje asociativo a un caso particular de aprendizaje por reestructuración, a través de una equilibración entre dos procesos: la asimilación y la acomodación. El aprendizaje se produce cuando tiene lugar un desequilibrio o un conflicto cognitivo. Asume una posición constructivista estática mediante la categorización de conceptos y dinámica a través de la formación de categorías.

Zona de desarrollo próximo (ZDP)	Vygotsky (1978 a y b)	Vygotsky (1978) considera la importancia de la interacción social en el aprendizaje y las funciones superiores se originan a través de las relaciones entre los seres humanos. Define a la zona de desarrollo potencial o de desarrollo próximo (ZDP) como la diferencia entre lo que el individuo es capaz de hacer sin ayuda de otras personas y lo que sería capaz de hacer con otras personas o instrumentos.
Aprendizaje significativo, organizadores previos.	Ausubel (1983)	La resolución de problemas es una actividad en la que la representación cognitiva de las experiencias previas y los componentes de una situación problemática actual, se reorganizan para lograr el objetivo: el aprendizaje por descubrimiento implica la formulación de hipótesis, la obtención de datos, su organización y su verificación.
Aprendizaje por descubrimiento, aprendizaje por problemas	Ausubel (1983)	Otra forma válida de aprendizaje es el descubrimiento guiado o dirigido mostrando al sujeto la solución que deberá ser recordarla y repetida, indicando verbalmente el principio de solución del problema y mostrando los pasos sin enunciar el principio que lleva a la solución.
Aprendizaje para la comprensión	Perkins (1995), Stone Wiske (1999)	La comprensión es una habilidad para desempeñarse con el conocimiento que se tiene" (Stone Wiske, 2007). Implica traducir o asimilar una información nueva a los conocimientos previos, así el aprendizaje requiere que se activen estructuras de conocimiento previas a las cuales poder asimilar la nueva información.
Cambio conceptual	Perkins (1995) Pozo (1998)	Proporcionar a los aprendices modelos o teorías alternativas desde los que reinterpretar la realidad y sus propios conocimientos como una condición necesaria para el cambio conceptual (Pozo, 1998).
Conflicto cognitivo, conflicto conceptual	Piaget (1978), Pozo (1998)	La instrucción basada en el conflicto cognitivo produce la reestructuración de los conocimientos como consecuencia de someter al estudiante a un conflicto, empírico (con la realidad) o teórico, buscando una teoría más explicativa.
Conflicto conceptual	Nussbaum y Novick (1982)	Modificar las creencias ingenuas, a través de una estrategia triple: descubrir las ideas preconcebidas, crear un conflicto conceptual y fomentar la acomodación cognitiva.
Efecto motivacional de los medios, efecto novedad	Salinas (1995), Cabero (2001, 2007)	Existe un efecto inicial de alta motivación por parte de los estudiantes al usar tecnologías, superando esta instancia, efecto novedad desaparece y el medio comienza a ser usado en forma constante e invisible.
Diseño de interfaces	ISO 9241 (1998)	La interacción computador humano concierne al diseño, evaluación e implementación de interfaces entre sistemas de computación y humanos. Comprende el estudio de la relación entre humanos y máquinas, con soporte de varias áreas de la computación: a) computación grafica, b) sistemas operativos, c) lenguajes de programación, d) metodologías de desarrollo; sobre los aspectos humanos: a) teoría de la comunicación, b) diseño grafico, c) lingüística, d) ciencias sociales, d) psicología cognitiva.

Requisitos ergonómicos para diseño de interfaces	ISO 11064 (2000), Bastien y Scapin (1993)	El diseño de interfaces debe responder a cuestiones tales como: su diferenciación del modelo para diseño de software, los involucrados en su realización, el nivel de creatividad, el diseño centrado en el usuario y sus implicaciones, la integración del conocimiento, la experticia en las disciplinas que contribuyen a su diseño, y la alta interacción requerida para alcanzar los requerimientos.
Diseño de interfaces	Smith y Mosier (1996)	Deben hallarse formas de atraer la atención hacia la información mostrada en la pantalla, a través de reglas como: no demasiada (ni poca) información en la pantalla, agrupación y ordenamiento con algún criterio tal como la usabilidad que significa que cualquier sistema diseñado para ser usado por personas debe ser fácil de aprender.

Tabla 2. Métodos y teorías

2. LINEAS DE INVESTIGACION y DESARROLLO

- Clasificación de población estudiantil de acuerdo a un perfil en relación con la información, ya que le estilo evoluciona y cambia de acuerdo a las variables de entorno y ambientales que afectan al estudiante.
- Selección de modelos pedagógicos adaptables al entorno virtual.
- Ajuste de estrategias reelaboración del cambio conceptual ante conceptos erróneos (misconceptions).
- Análisis de STI disponibles en la web.

3. RESULTADOS OBTENIDOS/ESPERADOS

3.1. Propuesta de Evaluación. Nuevos ítems y criterios de evaluación

Se propone una plantilla con actualización de los criterios para evaluación aplicables a los STI siguiendo la línea de investigación de Casanovas (2005) para los indicadores de evaluación de los simuladores didácticos que subsume las propuestas de evaluación de Cabero y Sancho (1996) y Cataldi (2001). De acuerdo al tipo de aplicación que son los STI se parte de la plantilla inicial de Cabero (2001) (Tabla 1) que se extendió a 80 ítems, agregando *criterios nuevos a las categorías existentes y generando nuevas categorías de análisis*, obteniéndose la Tabla 3.

Esta planilla ha sido proporcionada a un grupo de 15 docentes quienes hicieron los primeros ajustes. Actualmente, se ha aplicado la evaluación a tres tutores gratuitos para la enseñanza de programación, buscando validar la herramienta a fin de disponer de una opción válida para su selección de acuerdo a las necesidades educativas.

Calidad de contenidos
1. Nivel de actualización
2. Calidad científica
3. Conocimientos previos requeridos
4. Secuenciación y estructuración
5. Originalidad de presentación
Aspectos técnicos
6. Calidad del sonido
7. Calidad y tamaño de gráficos
8. Sincronización imagen-sonido
9. Calidad de uso de herramientas
10. Variedad de presentaciones
11. Calidad de animaciones
Motivación
12. Interés que despierta
13. Duración
14. Posibilidad de alcance de logros intermedios.
Organización interna de la información
15. Respuesta rápida a las acciones.
16. Respuesta rápida a las solicitudes de información.
17. Vínculos con la web
18. Registro sitios web visitados
19. Abundancia de información temática (para exploración de conceptos)
20. Uso de índices de contenidos disponibles.
21. Uso de hipertextos.
22. Uso de multimedia.
Costo económico
23. Relación costo-calidad
24. Relación costo-durabilidad científica
25. Relación costo-durabilidad física
Valor didáctico
26. Favorecimiento del proceso de aprendizaje
Claridad del programa
27. Claridad de la información
28. Claridad de explicaciones
Facilidad de manejo
29. Intuitividad
30. Menús de ayuda
31. Utilización de señales de disponibilidad de ayuda.
32. Opciones de uso para discapacitados
33. Facilidad de aprendizaje de manejo.
34. Comodidad de manejo.
Adecuación a los receptores
35. Adaptación al currículum
36. Adecuación a características de desempeño de los alumnos
37. Adecuación del vocabulario al ámbito profesional
38. Adecuación al nivel cognitivo alcanzado por el alumno
39. Adecuación de contenidos a los requerimientos académicos
40. Nivel de complejidad presentado acorde a los conocimientos previos requeridos
41. Estrategias didácticas tenidas en cuenta
42. Estrategia de formación relacionada al ámbito laboral
43. Registro de respuestas para elaborar perfil de comportamiento
44. Navegación libre siguiendo esquema de etapas (no inducida)
45. Selección reflexiva de opciones o caminos alternativos
46. Retroalimentación para reflexión y estudio de alternativas de solución.
Objetivos
47. Explicitación de los objetivos o logros esperados
48. Relación entre contenidos y evaluación
49. Relación entre contenidos y objetivos
50. Vinculación entre conceptos nuevos y existentes
51. Validación de conceptos o procedimientos a través de comprobaciones
Calidad en el diseño
52. Coherencia de estilo gráfico
53. Estabilidad de zonas en la pantalla
54. Estabilidad y equilibrio de cambio de fondos
55. Movimiento según ángulos de visión
56. Composición de pantalla simple
57. Recarga escénica (cognitiva)
58. Contraste de fondos y objetos

59. Densidad de objetos en pantalla
60. Diferenciación de zonas en pantalla
61. Jerarquización de pantallas
62. Originalidad en aplicación de recursos didácticos
Diseño de los ejercicios y problemas
63. Claridad en los enunciados
64. Planeo de problemas realistas
65. Relevancia del problema en la realidad profesional
66. Permite aprendizaje autodirigido
67. Permite razonamiento divergente y perspectivas múltiples ante un conflicto
68. Inclusión de herramientas colaborativas
69. Registro de evolución de los aprendizajes
70. Ponderación de opciones si las hubiere
71. Nivel de dificultad de los problemas
72. Dificultad acorde con conocimientos previos
Evaluación
73. Diferenciación entre actuación individual y grupal
74. Autoevaluación por unidades o ejes temáticos.
75. Claridad de aspectos evaluados
76. Posibilidad de experimentación y exploración
77. Selección de métodos válidos para solución
78. Relación del problema con el ámbito laboral
79. Retroalimentación de errores
80. Elaboración del cambio conceptual

Tabla 3: Propuesta para evaluación

4. FORMACION DE RECURSOS HUMANOS

La producción del grupo de investigación se plasma en comunicaciones a eventos y artículos (2007a,b y 2009), formación de los investigadores y tecnólogos participantes en el proyecto y dos tesis de grado finalizadas (Adriana Alvarez y Mónica Mollo) y una tesis en curso (Adriana Barbella) y dos tesis de maestría en curso (Ing. Octavio Rodríguez Angulo e Ing. Sebastián Carnotta) y dos tesis doctorales en curso (MD. Oscar Bruno y MD Silvia Sileo).

5. BIBLIOGRAFIA

- Ausubel, D.P.; Novak, J.D.; Hanessian, H. (1983) *Psicología Educativa: Un punto de vista cognitivo*, Editorial Trillas: México.
- Baecker, R. y Buxton, W. (1987) *Readings in Human Computer: a Multidisciplinary Approach*, Los Altos, CA, Kaufmann
- Bastien, J. M. C., y Scapin, D. L. (1993). *Ergonomic criteria for the evaluation of human-computer interfaces* (Report No. 156). Rocquen-court, France: Institut National de Recherche en Informatique et en Auto-matique.
- Cabero, J. (2001) *Tecnología educativa: diseño y evaluación de materiales para la enseñanza*. Barcelona, Paidós.
- Cabero, J. (2007) *Lecciones de las pasadas investigaciones* Ponencia en Edutec 07. UTN. 23-26 octubre. Buenos Aires.
- Casanovas, I. (2005) *La didáctica en el diseño de simuladores digitales para la formación universitaria en la toma de decisiones*, Tesis de Magister en Docencia Universitaria: UTN-FRBA.

- Cataldi, Z. (2001) *Diseño y Evaluación de Programas Didácticos Hipermediales*. Tesis para Magíster en Docencia Universitaria, UTN-FRBA.
- Cataldi, Z. Lage, F. 2009. *Categorías de análisis para la evaluación de Sistema Tutores Inteligentes*. Edutec 2009. 15, 16 y 17 de setiembre. Manaos.
- Cataldi, Z. y Lage, F. (2007a). *El problema del modelado del estudiante en Sistemas Tutores Inteligentes*. II Congreso de Tecnología en Educación y Educación en Tecnología. TE&ET'07. 12-15de junio. Facultad de Informática, Universidad Nacional de La Plata.
- Cataldi, Z. y Lage, Fernando J. (2009). Sistemas tutores inteligentes orientados a la enseñanza para la comprensión *Edutec-E Revista Electrónica de Tecnología Educativa*. Núm. 28, marzo. ISSN 1135-9250
- Cataldi, Z; Salgueiro, F. Y Lage, F. (2007b). *Fundamentos para el Submódulo Evaluador en Sistemas Tutores Inteligentes: Diagnóstico, predicción y autoevaluación*. CACIC 2007. 1-5 de octubre. Universidad Nacional del Nordeste Facultad de Ciencias Exactas y Naturales y Agrimensura. Corrientes y Universidad Tecnológica Nacional, Facultad Regional Resistencia.
- Dewey, J. (1949) *Experiencia y Educación*, Losada.
- Dewey, J. (1950) *Las escuelas de mañana*. Losada.
- ISO 11064-1 (2000) *Diseño ergonómico de los centros de control*.
- ISO 9241 (1996) *Requisitos ergonómicos para trabajos de oficina con pantallas de visualización de datos (PVD)*.
- Nussbaum, J. y Novick, S. (1982). *Alternative frameworks, conceptual conflict and accomodation: toward a principled teaching strategy*. Instructional Science, 11, pp. 183-200.
- Parkhurts, H. y otros (ca1920) *La escuela laboratorio Dalton*. Madrid: Ediciones de la Lectura
- Perkins, D. (1995) *La escuela inteligente*, Gedisa.
- Perkins, D. (1999) *¿Qué es la comprensión*, en Stone Wiske, M. (comp.) *La enseñanza para la comprensión*, Paidós.
- Perkins, D. y Blythe, T. (1994) *Putting Understanding up-front*. Educational Leadership 51 (5), 4-7.
- Piaget, J. (1978) *La Equilibración de las Estructuras Cognitivas*, Madrid, Ed. Siglo XXI.
- Pozo, J. I. (1998). *Aprendices y maestros*. Alianza
- Salinas, J. (1995) *Diseño, Producción y evaluación de Videos Didácticos*, Palma : Universitat de les Illes Balears, Servei de Publicacions i Intercanvi Científic, D.L.
- Sancho, J. (1996) *Para una Tecnología Educativa*, Barcelona, Horsori
- Shneiderman, B., Plaisant, C. (2006) *Diseño de interfaces de usuario*, 4ª Ed, Pearson,
- Smith, S. y Mosier, J. (1996) *Guidelines for Designing User Interface Software*, Ma, MITRE Corp
- Stone Wiske, M. (2007a) *Conferencia Enseñar para la comprensión con nuevas tecnologías*. Universidad de San Andrés. 8 de mayo.
- Stone Wiske, M. (2007b) Entrevista Clarín 27 mayo
- Vanlehn, K (1988). *Student Modelling*. M. Polson. Foundations of Intelligent Tutoring systems. Hillsdale. N.J. Lawrence Erlbaum Associates, 55-78
- Vygotsky, L. (1978a) *Mind in Society*, MA, Harvard University Press.
- Vygotsky, L. (1978b) *El desarrollo de los procesos psicológicos superiores*, Harvard University Press.
- Washburne, C. W. (1945) *La escuela individualizada: el sistema de Winnetka*. Losada.