



**CONGRESO
IBEROAMERICANO**
DE CIENCIA, TECNOLOGÍA,
INNOVACIÓN Y EDUCACIÓN

BUENOS AIRES, ARGENTINA
12, 13 Y 14 DE NOVIEMBRE 2014

**CONGRESSO
IBERO-AMERICANO**
DE CIÊNCIA, TECNOLOGIA,
INOVAÇÃO E EDUCAÇÃO

BUENOS AIRES, ARGENTINA
12, 13 Y 14 DE NOVEMBRO 2014

ODISEO: ALGORITMO DE EVALUACION INTELIGENTE EN E-ELARNING

SIMANCA, F; ABUCHAR, A; RIVERA, P

ODISEO: ALGORITMO DE EVALUACION INTELIGENTE EN E-ELARNING

Fredys Simanca Herrera

Fundación Universitaria los Libertadores
fredyssimanca@hotmail.com

Alexandra Abuchar Porras.

Universidad Distrital Francisco José de Caldas
alexandraabuchar@yahoo.com.mx

Pedro José Rivera Osorio

Fundación Universitaria los Libertadores
invprivera@gmail.com

Resumen.

Este artículo es el resultado de un proceso de investigación llevado a cabo sobre el impacto de evaluación de desempeño de estudiantes de educación superior y está enfocado en las pruebas de desempeño académico realizadas por el estado colombiano a estudiantes próximos a obtener su título universitario, siendo este requisito indispensable para su titulación y los resultados obtenidos es uno de los factores para medir la calidad de la educación de las Instituciones.

La investigación se concentra específicamente en ingeniería de sistemas ya que los estudiantes son evaluados con indicadores específicos en lo disciplinar con los componentes y contenidos referenciales propios y la cognitiva que comprende las competencias de interpretación, argumentación y proposición. Lo que genero una propuesta que fortalece y potencializa el proceso evaluativo y así obtener mejores resultados, es allí donde los gestores de la investigación proponen una estrategia que rompe las barreras de espacio y tiempo, ODISEO el algoritmo de evaluación inteligente en e-learning, plataforma en donde podrán realizar: entrenamientos, simulacros, ver su nivel de conocimiento y progreso. Para el desarrollo de ODISEO fue necesario conceptualizar y realizar un compendio de temas como: los procesos de evaluación académica y e-evaluación, teoría clásica de los test (TCT), Teoría de Respuesta al Ítem (TRI), Test Adaptativos Informatizados (TAI), Sistemas de Tutor Inteligente (STI), teorías y diseños metodológicos instruccionales para el desarrollo de plataformas e-learning y e-evaluación, diseño de algoritmo y desarrollo del software web entre otros.

Palabras clave.

Algoritmo, evaluación inteligente, test adaptativos, Algoritmo de Evaluación Inteligente, TAI, TRI, TCT.

Abstract.

This article is the result of an investigation process conducted on the impact assessment performance of students in higher education and its focused in the academic performance of students in the proofs done by the Colombian state to students coming to get their degrees, this being a prerequisite for their titration

requirement and the results are one of the factors for measuring the quality of education in the institutions.

The research focuses specifically on systems engineering since the students are evaluated with specific disciplinary indicators, that includes components and own referential contents, and the cognitive that cover the interpretation competitions, argumentation and proposition. What entail to generate a proposal to strengthen and empower the process and obtain better results in the institutions and is there where research managers propose a strategy that break down the space and time bounders, ODISEO the algorithm of intelligent evaluation on e-learning, platform where they can conduct training drills, see their level of knowledge and progress.

For ODISEO's development it was necessary to conceptualize and identify a compendium of topics as: processes of academic assessment and e-assessment, classical test theory (CTT), Item Response Theory (IRT), computerized Adaptive Test (IAT), Intelligent Tutoring Systems (ITS), instructional theories and designs for the development of the e-learning and e-evaluation platform, methodological models of instructional develop, algorithm design and software web setting-up, among others.

Key word.

Adaptative algorithm, Intelligent Evaluation, test, Algorithm of Intelligent Evaluation.

Introducción.

La evaluación es esencial en todas las actividades del ser humano, es ese proceso interno que lleva a una mejora continua permitiendo el poder desarrollarse, progresar y mejorar la calidad de vida. Por ende la evaluación está presente en todos los procesos implícitos o explícitos del ser humano y ésta tiene una relación directa con los procesos de cómo y para que se aprende, es en ese aprender, donde este artículo les llevará, la evaluación, de desempeño académico, ésta presente en dos niveles: La evaluación interna: la que se lleva a cabo dentro de la institución, al interior de cada asignatura, tiene su propia dinámica y los resultados ayudan a la toma de decisiones para mejorar los procesos internos. La evaluación externa, es la que realiza un agente externo a la institución y determina el grado de calidad que se da siendo esta última el centro de este artículo.

Cuando se realiza la evaluación externa es importante que la institución evaluada conozca los diferentes criterios en que será evaluada ya que con los resultados obtenidos se realizará la retroalimentación para robustecer a aquellos aspectos en que les fue bien y reforzar aquellos en que no les fue también, estos resultados les ayudarán a tomar decisiones, y potenciar a un auto-conciencia de todos los actores involucrados en el proceso evaluativo según Arakaki y Florez [1].

En Colombia se realizan pruebas censales: Estas pruebas son aplicadas en diferentes momentos de la vida académica de los estudiantes y dan indicadores de calidad educativa entre las instituciones y regiones. El estudiante en Colombia es evaluado cinco veces: Saber grados 3°, 5° y 9°; Saber grado 11, Saber Pro (antes de salir de la Universidad). Las Pruebas Saber Pro, se aplican a los estudiantes que cumplan con el 75% de los créditos académicos. Es importante destacar la necesidad de contar con estrategias que ayuden a potencializar los resultados obtenidos por los estudiantes en estas pruebas, el resultado de la investigación es una estrategia, denominada ODISEO, una plataforma de e-evaluación y como primera aplicación directa un pilotaje de las pruebas Saber Pro para los estudiantes de Ingeniería de sistemas ya que estos son evaluados con indicadores específicos en lo disciplinar, el cual comprende componentes y contenidos referenciales propios como: Ciencias básicas (CB), como la

Física, las matemáticas; Ciencias básicas de ingeniería, (BI), como: análisis numérico, probabilidad y estadística e investigación de operaciones; Ingeniería aplicada, (IA), Matemáticas discretas, Programación y algorítmica, Informática básica, Arquitectura del computador, Redes y comunicaciones, Administración de la Información, Ingeniería de software; Formación complementaria (C), Ciencias económico administrativas y la cognitiva que comprende la competencias de interpretación, argumentación y proposición según ICFES [2].

1. Procesos de evaluación académica y e-evaluación

La virtualidad brinda entornos y herramientas potencializando todos los procesos a través de la web en que el ser humano está inmerso y siendo la educación un factor importante en el desarrollo del ser, la virtualidad ofrece un sin número de posibilidades, entre ellas las plataformas e-learning para el aprendizaje y dentro de ellas la e-evaluación, la cual tiene características propias; se realizan por un medio *computarizado*, debe tener acceso a Internet, la comunicación puede ser asíncrona, cuando es en tiempo real o síncrona, cuando es simultánea; además presenta una variedad de apoyos académicos disponibles en red según Churchill [3].

1.1 Principios de la evaluación con TIC

Estos principios expresados a continuación están presentes en los procesos de evaluación presencial y trascienden a la evaluación computarizada, teniendo en cuenta que en la evaluación virtual el proceso es más complejo por la no presencialidad, por lo tanto los evaluados deben tener madurez personal, formación ética y ser responsables en proceso de aprendizaje, ya que se parte que el evaluado es quien realiza la evaluación.

- a) Confiabilidad
- b) Validez
- c) Objetividad
- d) Autenticidad

1.2 Generaciones de la evaluación computarizada.

La evaluación en las plataformas virtuales ha evolucionado y según Bunderson, Inouye y Olsen [4], plantean cuatro generaciones de evaluación por computador. Cada una de ellas representa un avance sobre la otra.



Imagen 1: Generaciones de evaluación Computarizada

Fuente: realizada por los autores

1.3 Clasificación evaluación con TIC

El tema de la clasificación de la evaluación con la TIC ha sido trabajada por varios autores siendo la taxonomía de Elena Barberá [5] la que se trabajó en esta investigación ya que el autor da una clasificación y conceptualización sobre evaluación con la utilización de las TIC:

- a) Evaluación automática
- b) Evaluación enciclopédica
- c) Evaluación colaborativa

La evaluación automática: es aquella se hace son preguntas que permiten visualizar inmediatamente la respuesta correcta validando la respuesta.

La evaluación enciclopédica: Es aquella que se realiza mediante trabajos escritos como: artículos, ensayo, ponencias entre otros sobre un tema específico y los estudiantes deben recurrir a la búsqueda de bases de datos en internet.

La evaluación colaborativa: Es aquella en donde se utilizan recursos virtuales ejerciendo una praxis colaborativa permitiendo evaluar el producto y el proceso ejemplo de esta evaluación está la participación en foros, debates, grupos de discusión, grupos de trabajo, entre otros.

Por lo tanto, Rodríguez Gómez e Ibarra Sáiz [6] en donde definen la e-evaluación como un “proceso de aprendizaje, mediado por medios tecnológicos, a través del cual se promueve y potencia el desarrollo de competencias útiles y valiosas para el presente académico y el futuro laboral de los estudiantes como profesionales estratégicos “ y teniendo en cuenta el estudio hecho por Rocío Yuste y Florentino Blázquez [7] apoya esta afirmación, señala que a través de la evaluación online es posible valorar y evaluar el aprendizaje a lo largo de un proceso formativo, utilizando tareas basadas en metodologías activas. La e-evaluación tiene un principio fundamental, aprender y valorar los logros conseguidos con la uso de recursos de: telecomunicación, plataformas virtuales y la interacción entre los actores del proceso. Un aspecto a tener en cuenta sobre la evaluación son los ítems o preguntas que se realizan en los diferentes instrumentos utilizados y es allí donde nace la teoría de los test.

2. Teoría clásica de los test (TCT)

Es importante definir la palabra test según Lord [8] “un test psicológico o educativo es un instrumento para obtener una muestra de conducta”, por lo tanto es una forma de determinar el grado de conocimiento, la aptitud de una persona, frente a un tema específico, según [5] Yela, Sperman fueron los primeros en proponer un modelo para medir los test, en donde la puntuación obtenida es la puntuación empírica y conocida como variable (X) y está compuesta por dos componentes; la puntuación verdadera identificada con (V) y la falsa como (e), de lo que se puede formular $X = V + e$, es la teoría de los test más conocida en psicometría, su característica se basa en la exactitud de la medida y la determinación exacta hacia el error de medición.

Personajes como: Sir Francis Galton, Alfred Binet, James Mckeen Cattell, Terman, Navas, Spearman, Thorndike, Thurstone, Allen, Gulliksen, Guilford, Yen, Meliá,

Cronbach, Rajaratnam y Gleser, entre otros, identifican las principales teorías en la medición de test en el campo de la psicometría: siendo esta un aspecto metodológico dentro de la psicología y su objetivo principal medir o cuantificar las variables psicológicas de las personas incluyendo sus implicaciones teóricas y prácticas, estas teorías son:

- Teoría clásica de los test
- Teoría de la generalizabilidad
- Teoría de respuesta a los ítems

2.1 La Teoría de respuesta al ítem (TRI)

Teoría respuesta al ítem, propuesta por George Rasch en el año 1960. La cual examina el comportamiento de los test, esta utiliza un modelo matemático logístico para describir la relación entre el nivel de habilidad del examinado y la probabilidad que éste dé una respuesta correcta a un ítem del test. Los test de medida consisten en la selección informatizada de los ítems que puedan medir mejor la habilidad de un individuo. Las respuestas a cada ítem se pueden catalogar como respuesta verdadera o falsa, de tal manera que si se asocia una variable X_{ij} para representar la respuesta del individuo j al ítem, esta variable tomará el valor 1 si la respuesta es verdadera y 0 si la respuesta es falsa. La probabilidad de responder correctamente al ítem i dependerá del parámetro Y_i , que es el valor que toma la capacidad de cada individuo. Por consiguiente, la probabilidad $P_i () Y_i$, es una función creciente del parámetro Y_i . Esta función es conocida con el nombre de **función de respuesta al ítem** o curva característica de ítem.

El **parámetro de dificultad** la probabilidad de responder correctamente al ítem es del 50 %. aquí indica la posición de la curva característica del ítem o CCI en donde los ítems fáciles estarán a un extremo y los difíciles al otro, teniendo en cuenta el parámetro de discriminación X_i está asociado con la pendiente de la CCI en un punto Z_i , cuanto mayor sea la pendiente de la curva, mayor será la diferencia en las probabilidades $P_i () Y_i$ de valores próximos, lo que indica que el ítem en cuestión tiene una gran capacidad de discriminación entre los individuos con escasa capacidad y los que tienen elevada capacidad, cuando un ítem tiene un parámetro de discriminación negativo debe ser eliminado del test ya que su definición es errónea (un parámetro a_i negativo significa que la probabilidad $P_i () Y_i$ es menor para los individuos de capacidad elevada que para los de baja capacidad).

Por otro lado, no es habitual que el valor de a_i exceda de 2, implica que el rango del parámetro de discriminación es 0; 2. Valores de a_i próximos a 2, y son ítems con gran discriminación, igualmente, los valores de a_i próximos a 0 se relacionan a ítems discriminatorios según Rizopoulos [9], por lo tanto la TRI genera unos resultados así:

- a) **Parámetros.** Como se ha comentado anteriormente, tanto la dificultad, como la discriminación son dos elementos esenciales en la valoración de un ítem. Lo mismo puede decirse sobre el pseudo-azar. Los parámetros pueden ser:
- a. Índice de discriminación
 - b. Índice de dificultad
 - c. Probabilidad de acierto al azar
 - d. Error por descuido
- El parámetro Θ es la habilidad o el rasgo latente, estos elementos tienen implicaciones directas en la construcción de test y pruebas académicas.

- b) **Curva característica del ítem.** Es un elemento fundamental en la TRI porque, por un lado, ofrece una representación gráfica de las características del ítem, y por otra parte muestra la relación entre el nivel de habilidad y la respuesta al ítem, en términos probabilísticos.
- c) **Puntuación verdadera en el test.** Se identifica con la suma de las probabilidades estimadas con relación al nivel de habilidad, para el total de ítems del test.
- d) **Curva características del test.** Permite ver de manera gráfica, la relación entre puntuación verdadera y habilidades. La Curva Característica del Test (CCT) muestra la relación existente entre la puntuación total en un test (no solamente en un ítem como sucede con la CCI) y el nivel de habilidad de una persona.
- e) **Nivel de información.** Es un indicador de la precisión de la estimación. La TRI afirma que si una persona tiene un nivel n de conocimiento, tendrá asociado a dicho valor una probabilidad de que conteste correctamente a una pregunta. Expresado a la inversa, si una persona que responde correctamente a un ítem, es posible suponer que esta persona posee un determinado nivel de conocimiento.

Para cada nivel de conocimiento o rasgo, existe una probabilidad de contestar correctamente al ítem. Esta probabilidad es pequeña para sujetos con bajo nivel de rasgo, y grande para sujetos con altos niveles de rasgo. La relación entre el nivel de aptitud y la probabilidad de respuesta correcta al ítem se representa con la curva características del ítem (CCI).

2.2 Modelos de la TRI

La teoría de respuesta al ítem presenta tres modelos según el número de parámetros:

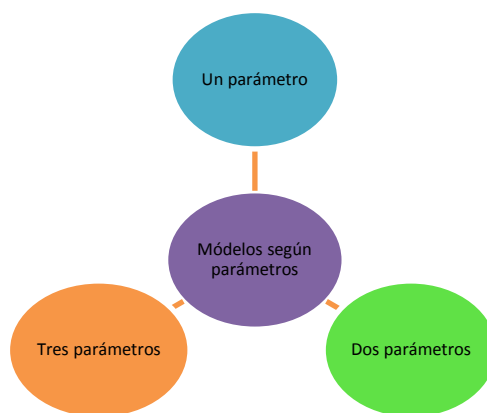


Imagen 2, Modelos según parámetros

Fuente: elaboración propia

1. Modelo de un parámetro, también llamado modelo de Rasch. Se estima la dificultad del ítem (b).
2. Modelo de dos parámetros. Se estima la dificultad (b) y la discriminación del ítem (a).
3. Modelo de tres parámetros, o modelo de Birnbaum, este modelo estima la dificultad (b), la discriminación (a) y el pseudo-azar(c).

También los modelos en función del número de respuesta de los ítems así:

- Modelo dicotómico: donde las respuestas a los ítems tienen dos opciones, valoradas como 1 y 0.

- Modelos politómicos: En este caso los ítems tiene más de dos opciones.

2.3 Función de información

Se puede calcular la función de la información de un ítem y la función de información del test completo:

- a) Función de información del ítem (FII). Aunque se puede calcular para todos los niveles de habilidad, en los modelos de uno y dos parámetros, los ítems miden con mayor precisión alrededor del valor de dificultad.
- b) Función de información del test (FIT). Calcula la cantidad de información proporcionada en el test total para los distintos niveles de habilidad. Se calcula sumando las funciones de información de todos los ítems para cada nivel de habilidad.

2.4 Procedimiento para el análisis de un test por medio de la TRI

La estructura a seguir para analizar las características de un test según la TRI, son los siguientes:

- a) Preparar los datos para su análisis.
- b) Analizar si los datos cumplen los supuestos básicos de la TRI.
- c) Estimar los parámetros del modelo elegido.
- d) Elaborar los resúmenes y gráficas correspondientes.
- e) Analizar el ajuste del modelo a los datos.
- f) Estimar los niveles de habilidad de los participantes.

2.5 Test Adaptativo Informatizado

El Test Adaptativo Informatizado (TAI) surge en el año 1996 y es una de las principales aplicaciones derivadas de la Teoría de la Respuesta al Ítem (TRI), el cual presenta ítems y agrupa las respuestas de los evaluados mediante los procesos llevados en sistemas informatizados y tiene la capacidad de adaptarse al desempeño y rendimiento, comenzando con una estimación inicial del nivel de conocimiento del evaluado.

Componentes de un TAI

- a) Banco de Ítems que conllevan las características psicométricas establecidas previamente desde un modelo de la Teoría de la Respuesta al Ítem (TRI).
- b) Procedimiento para implementar el inicio y el final de la prueba, además, la estrategia para establecer gradualmente los ítems más sobresalientes.
- c) Estadísticas de estimación de los rasgos de las personas del mismo nivel.

El uso de los TAI optimiza los tiempos y dan seguridad en su aplicación y se debe considerar los siguientes aspectos según Barrada [10]:

- a) **Las propiedades psicométricas**, donde se mide con exactitud para el cual es diseñado el test y su validez de contenido y de respuestas.
- b) **El banco de ítems**, almacena los rasgos y habilidades más sobresalientes en los diferentes test aplicados, de allí se pueden extraer los más característicos durante un test determinado. Se recomienda almacenar hasta 100 ítems característicos.

- c) **La estimación de los parámetros de los ítems**, proporciona los modelos de elaboración para las escalas de medición de los diferentes parámetros que compone el banco de ítems.
- d) **La comparación de las estimaciones de los parámetros**, convierte los diferentes sistemas de unidades entre test. Y al confrontarse diferentes test pueden ser equivalentes los puntajes deseados.
- e) **Un programa** a utilizar que convierte los test aplicados a un TAI, el objetivo es ofrecer una visión actualizada en la aplicación de los test. Para ello, se presenta la estructura básica de un TAI según Cattell [11], como el que se muestra en la imagen 3.

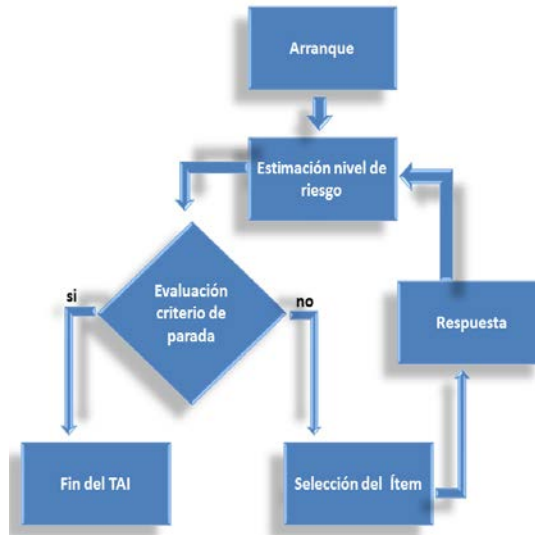


Imagen 3. Diagrama de flujo de un TAI
Fuente: elaboración propia

En los últimos años, se han desarrollado diversos TAI y es probable que, dadas las ventajas que ofrece esta técnica, sean bastantes más lo que se hagan disponibles próximamente. Se hace especial énfasis en la selección de ítems, la psicometría parte fundamental para la adaptabilidad del test según Arbero, M^a [12].

2.5.1 Componentes de un TAI



Imagen 4 Componentes de TAI
Fuente, elaboración propia

Los componentes centrales de un TAI lo integran proceso iterativo con tres pasos:

- a) La evaluación del criterio
- b) La selección de ítems
- c) La respuesta

a) La evaluación del criterio

Los criterios para la aplicación de un TAI en los test de evaluación pueden estar basados en determinadas reglas características como:

- Alcanzar un cierto número de ítems para administrar
- Reducir la incertidumbre en la estimación del nivel de rasgo
- Considerar la exactitud de ítems en los factores a evaluar deseados. Combinaciones de estos criterios resultan posibles.

b) La selección de ítems

En cuanto a la selección de los ítems se puede distinguir dos condiciones según Arbero [9], la primera condición cuando los ítems son seleccionados de uno en uno el test tiene tantos puntos de aplicación a la ejecución como el número de ítems que van a ser administrados, la segunda condición cuando los puntos de aplicación se son tres o cuatro, se conoce como Test Multietápico (Luecht y Nungester, 1998).

El trabajo realizado por Díaz y Ponsoda [13], el cual cita la investigaciones y trabajos realizados por Parshall [14] en donde identifican cuatro objetivos básicos a cumplir mediante un TAI.

- Permitir la estimación precisa del nivel de rasgo de los evaluados.
- Limitar la probabilidad e implicaciones de una filtración de ítems.
- Garantizar el ajuste a las especificaciones de contenido de la prueba.
- Facilitar el mantenimiento del banco de ítems.

c) La Respuesta

Éste es el único elemento del TAI en el que interviene el evaluado ya que los ítems suministrados han sido ajustados bajo alguno de los modelos de TRI.

La explicación de las puntuaciones de un TAI puede estar orientada a criterio o a norma buscando clasificación por categorías; ya sea apto o no apto, nivel bajo, medio o alto.

La evolución de un test a un TAI admite cambios: primero, automatizar el test, luego, hacerlo adaptativo. Esto quiere decir que en la medida que los test computarizados se van difundiendo, también lo van haciendo los TAI.

En el sistema educacional, la aplicación de TAI comienza con una estimación inicial del nivel de conocimiento del estudiante.

- a)** Todas las preguntas son examinadas para determinar cuál será la mejor para ser propuesta a continuación, según el nivel de conocimiento estimado del estudiante.
- b)** La pregunta es planteada y el estudiante responde.
- c)** De acuerdo con la respuesta del estudiante, se realiza una nueva estimación de su nivel de conocimiento.

Los pasos del 1 al 3 se repiten hasta que se cumpla alguno de los criterios de terminación definidos.

2.5.2 Fases para la construcción de un test adaptativo informatizado (TAI).



Figura 5 Fases TAI; Fuente, elaboración propia

Fase I: Planificación y prospección del TAI, en esta fase el objetivo es organizar lo que se va a exponer a lo largo de la construcción del TAI, por ello se debe tener claro; el objetivo final que se pretende con la evaluación, la elección de los ítems y las puntuaciones que se va a utilizar en la evaluación. Elementos a tener en cuenta en la fase:

- a) Definición operativa de la variable.
- b) Definición de la población.
- c) Definición del objetivo de evaluación. Es importante tener claro a qué población va dirigido el test, ¿qué es lo que se va evaluar?, ¿Cuáles son las características de la población?
- d) Idoneidad del rasgo.
- e) Formato y estructura de los ítems. hay que determinar qué tipo de ítems se van a manejar en el test, los más comunes son: los de tipo banco de Ítems Calibrados, dicotómicos, unidimensionales con alternativas múltiples.

Fase II: Producción del banco de ítems, El principal objetivo es crear una en base de ítems que contengan una recopilación de características de forma estandarizada. Elementos a tener en cuenta en la fase:

- a) Estudio piloto de la aplicación del TAI.
- b) Aplicación del test a una muestra piloto.
- c) Estimación de los parámetros clásicos psicométricos del test con los datos obtenidos.
- d) Reglas generales de los ítems y selección de descriptores.
- e) Tamaño del banco de ítems y diseño de las pruebas.
- f) Muestra de los evaluados.

Fase III: Calibración y ensamblado del banco de ítems, en esta se hace es efectuar una auditoria de las respuestas de los evaluados y verificar las irregularidades, en los ítems y en los evaluados para así depurar filas y columnas de la matriz de datos, posteriormente se realiza el control para filtrar la captura y obtención de datos para

luego realizar un análisis de cada prueba que permita localizar ítems incompatibles con los modelos TRI y finalmente se verifican las pautas de respuesta de los evaluados. Elementos a tener en cuenta en la fase:

- a) Análisis previos de las respuestas.
- b) Calibración de los ítems por bloques.
- c) Estimación de los parámetros de la TRI.
- d) Valoración: Ajuste, funcionamiento diferencial y dimensional.

Fase IV: Implementación y ejecución del TAI, una vez que se conocen las posibilidades del banco de ítems, se implementa y ejecuta el TAI con el objetivo de diligenciar las características y parámetros dentro de un software de administración determinado. Elementos a tener en cuenta en la fase:

- a) Selección de ítems en función de los resultados obtenidos en la calibración.
- b) Selección de ítems de anclaje para la aplicación del banco de ítems en la siguiente fase.
- c) Estudio normativo.
- d) Elaboración de los test para la su aplicación.
- e) Elaboración de un sistema informático para la aplicación de los test.
- f) Administración de ítems; Inicio, continuación y final de la prueba.

Fase V: Explotación y gestión del TAI, una vez elegido el procedimiento de administración de la prueba es necesario generar el ambiente propicio que permita ejecutar la prueba; es importante que el evaluado conozca las instrucciones y este familiarizado con la mecánica de la prueba, Elementos a tener en cuenta en la fase:

- a) Aplicación de la prueba a la muestra.
- b) Selección del algoritmo adaptativo.
- c) Elaboración del banco de ítems definitivo.
- d) Fiabilidad y validez del TAI.

Fase VI. Mantenimiento y renovación del TAI, se debe realizar un mantenimiento del proceso en si del TAI, del banco de ítems, esto se hace con una permanente actualización de parámetros de los ítems de renovar y ajustar el banco de ítems existente, el proceso de mantenimiento es el mismo de la construcción del TAI; desarrollo del banco inicial de ítems, estudio piloto, estudio normativo y estudio de simulación. Elementos a tener en cuenta en la fase.

2.6 Los sistemas tutores inteligentes (STI)

Son sistemas informatizados enfocados a la docencia cuyo objetivo es simular la capacidad del desarrollo del aprendizaje donde el proceso es guiado por el tutor y este debe analizar el comportamiento del estudiante para conocer el estado del conocimiento y así satisfacer los requerimientos para ser aplicados en cada momento. Por lo tanto un tutor inteligente, *“es un sistema de software que utiliza técnicas de inteligencia artificial (IA) para representar el conocimiento e interactúa con los estudiantes”*. También se puede definir un (STI) como *“un sistema que modela la enseñanza, el aprendizaje, la comunicación y el dominio del conocimiento del especialista y el entendimiento del estudiante sobre ese dominio”*. Es un sistema que agrega técnicas de IA (Inteligencia Artificial) a fin de crear un ambiente que considere los diversos estilos cognitivos de los estudiantes que utilizan el programa” según Giraffa [15]. En el Tabla 2 se pueden ver algunas de las características propias de los Sistemas Tutores Inteligentes.

Categoría	Característica	Metas principales
Interfaces inteligentes	Representación del usuario individual y del dominio	Facilitar el uso del software para una amplia variedad de usuarios
Sistemas Adaptativos de Lenguaje Natural (NL)	Inferencia de necesidades y metas del usuario	Facilitar la comprensión y generación del lenguaje natural, al ajustar el dominio al usuario particular
Tutores Inteligentes (ITS)	Capaces de reconocer errores y conceptos erróneos, de observar el comportamiento del usuario e intervenir de ser necesario	Enseñar conceptos al usuario estudiantes (se asume la meta del usuario)
Soporte inteligente	Incluyen las características de ITS; uso de lenguaje natural, orientados a la solución de problemas	Ayuda activa, apoya al usuario en la ejecución de la tarea, no en su aprendizaje
Sistemas de explicación	Agregan a lo anterior, la capacidad de explicar el comportamiento del sistema	Mostrar al usuario el porqué del comportamiento
Agentes cooperativos inteligentes	Sistemas independientes, que pudieran entrar en las categorías anteriores; agregan elementos de trabajo en equipo y negociación	Mayor poder de resolución de problemas, en grupos. Conseguir la adaptación tanto respecto al exterior (humano) como interior (otro agente).

Tabla 2 Características de los STI.

Autores como Ataldi [16]. El cual brinda una estructura general de un STI con la división de los submódulos en funciones específicas soporta una configuración distribuida, que se puede ver en la en la imagen 6.

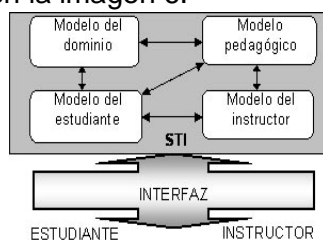


Imagen 6 Esquema de un STI con sus módulos correspondientes.

A través de la interacción entre los módulos básicos, como se muestra la figura 6, los STI son capaces de determinar lo que sabe el estudiante y sus progresos, por lo que la enseñanza, se puede ajustar según las necesidades del estudiante, sin la presencia de un tutor humano.

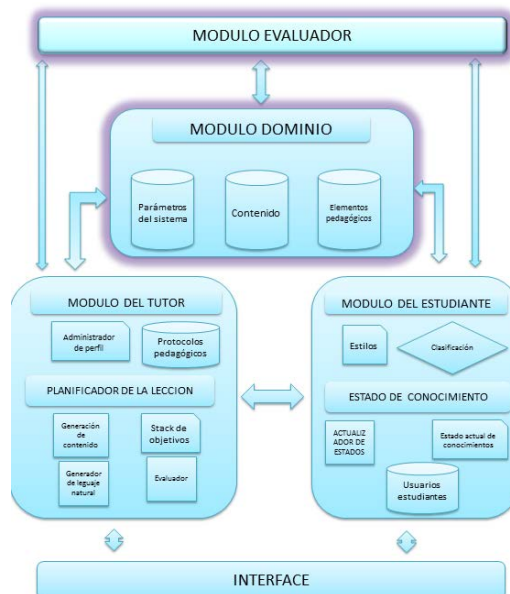


Imagen 7: Estructura de un STI.

Modulo Tutor. Modulo que define y aplica estrategia pedagógica de enseñanza, contiene los objetivos a ser alcanzados y los planes utilizados para alcanzarlos. Selecciona los problemas, monitorea el desempeño, provee asistencia y selecciona el material de aprendizaje para el estudiante.

Módulo Estudiante

Tiene por objetivo realizar el diagnóstico cognitivo del estudiante, y el modelado del mismo para una adecuada retroalimentación del sistema. Los datos se almacenan en una base de datos del estudiante a través del uso de un gestor. Para el módulo estudiante se han planteado los siguientes submódulos.

- a) **Estilos de aprendizaje:** Está compuesto por una base de datos con los estilos de aprendizajes disponibles en el sistema es la forma de clasificar el comportamiento de un estudiante de acuerdo a la manera en que toma la información, es la forma de agrupar o clasificar el estudiante de acuerdo a un perfil,
- b) **Perfil psico-sociológico del estudiante,** Para determinar el perfil psico-sociológico se usa la Teoría de las Inteligencias Múltiples de Gardner [17], quien señala, “no existe una inteligencia única en el ser humano, sino una diversidad de inteligencias que evidencian las potencialidades y aspectos más significativos de cada individuo, en función de sus fortalezas y debilidades para la expansión de la inteligencia.
- c) **Estado de conocimientos,** Contiene el mapa de conocimientos obtenido inicialmente a partir del módulo del dominio y que el actualizador de conocimientos irá modificando progresivamente a través de los resultados obtenidos en las evaluaciones efectuadas por el módulo del tutor quien le enviará dichos resultados procesados.
- d) **Perfil psico-sociológico del estudiante,** Para determinar el perfil psico-sociológico se usa la Teoría de las Inteligencias Múltiples de Gardner [18], quien señala, “no existe una inteligencia única en el ser humano, sino una diversidad de inteligencias que evidencian las potencialidades y aspectos más significativos de cada individuo, en función de sus fortalezas y debilidades para la expansión de la inteligencia”.

Módulo Dominio. Tiene el objetivo global de almacenar todos los conocimientos dependientes e independientes del campo de aplicación del STI. Los submódulos son los siguientes:

- a) **Parámetros Básicos del Sistema:** los cuales se almacenan en una base de datos.
- b) **Conocimientos:** son los contenidos que deben cargarse en el sistema, a través de los conceptos, las preguntas, los ejercicios, los problemas y las relaciones.
- c) **Elementos Didácticos:** Son las imágenes, videos, sonidos, es decir material multimedia que se requiere para facilitarle al estudiante apropiarse de conocimiento en la sesión pedagógica. Los temas relacionados con el almacenamiento de conocimiento han sido tratados ampliamente por Russell y Norvig[19].

3. Diseño del Algoritmo ODISEO

La definición de algoritmo más generalizada es el conjunto de instrucciones o acciones ordenadas y sistemáticas que dan solución a un problema. En el desarrollo de ODISEO, se llevara a cabo teniendo en cuenta los test adaptativos informatizados TAI , los sistemas de tutor inteligente STI , con el modelo de desarrollo de software en cascada de tal forma que el inicio de cada etapa debe esperar a la finalización de la etapa anterior; la metodología de desarrollo de software será ADDIE según Strickland [20] siendo esta la que mejor se adapta para el modelo instruccional el cual estará enfocado al diseño instruccional de Gagné y se desarrollara en ambiente web.

3.1 Análisis y Definición de Requerimientos

Teniendo en cuenta los temas tratados anteriormente se deben tomar los requerimientos para la e-evaluación, los elementos del TAI, y DE STI , y los componentes de evaluación para ingeniería de sistemas, cada uno de estos aspectos presenta requerimientos funcionales y no funcionales, así como externos e internos que encierran una complejidad en el desarrollo de ODISEO, En esta primera etapa serán analizados los requerimientos básicos para el funcionamiento del algoritmo, se definirá el objetivo del software. Si bien esta es una primera aproximación sirve de límite para las siguientes etapas.

3.2 Diseño del Algoritmo

Se definirá el diseño del algoritmo y el aplicativo con un máximo nivel de detalle, generando modelos de aspecto gráfico, del contenido, y del funcionamiento, estos modelos son prototipos del sitio software y serán reflejados con exactitud, se trabajará con usuarios evolucionándolos hasta que quede satisfecho con todos los aspectos.

3.2.1 Modelamiento del algoritmo

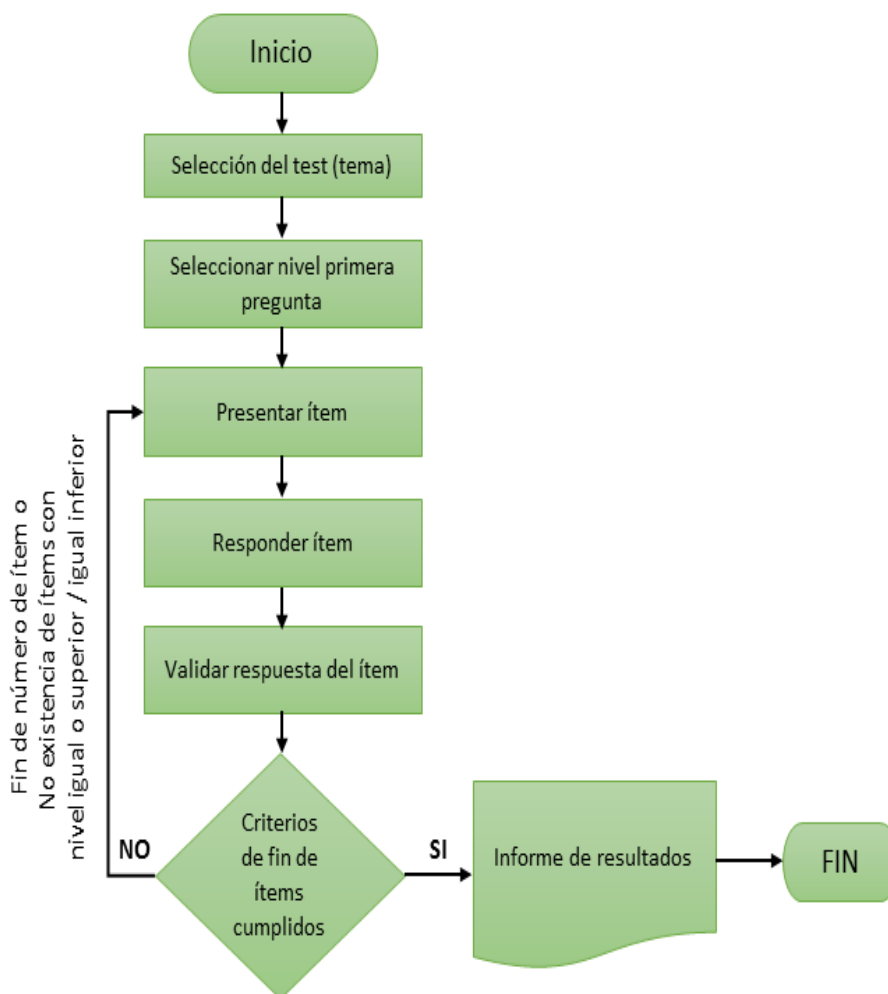


Imagen 8: Estructura de un STI.

Identificación de los stakeholders en donde se reconocen todas aquellas personas u organizaciones impactadas directa o indirectamente.

Actor	Caso de Uso
Administrador	Registrar asignatura Autenticar Hacer Backup BD. Restaurar BD Autenticar profesores
Docente	Administrar Asignaturas Introducir temas Elaborar preguntas Realizar Test Resultados
Estudiante	Autenticar Realizar entrenamiento Realizar simulacro Ver resultados

Tabla 3 Diagrama de Caso de Uso; Fuente: Elaboración propia

3.2.2 Modelamiento del comportamiento del sistema

Rol del administrador

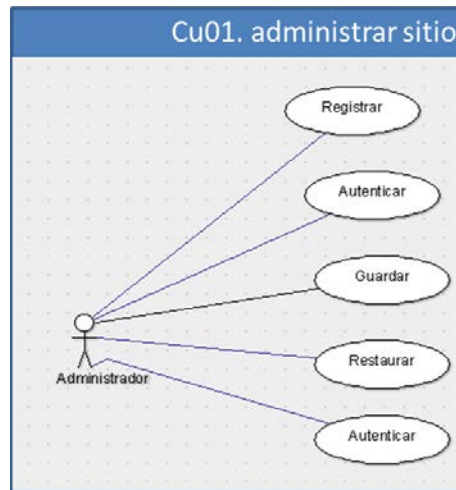


Diagrama de Caso de Uso: Administrador de la Aplicación; Fuente: Elaboración propia

Rol del docente

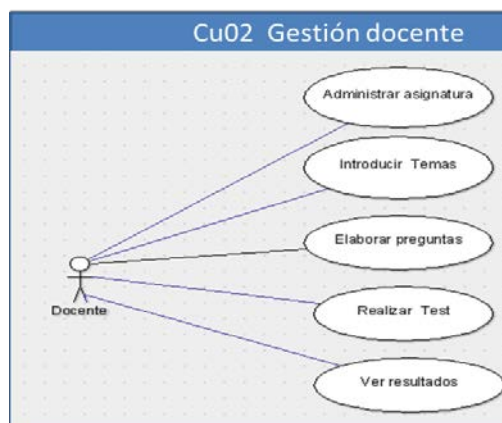


Diagrama de Caso de Uso: Docente de la Aplicación; Fuente: Elaboración propia

Rol del estudiante

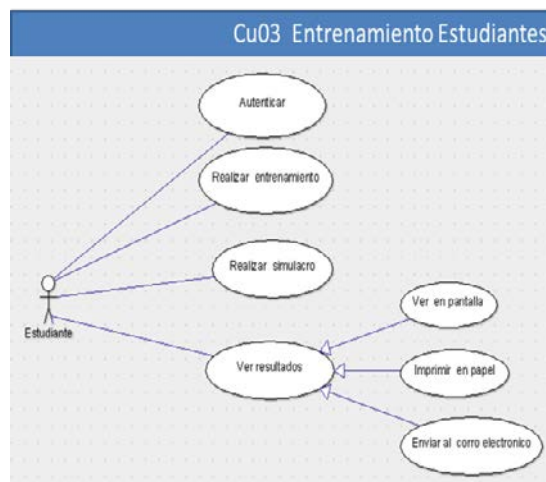


Diagrama de Caso de Uso: Estudiante de la Aplicación; Fuente: Elaboración propia

Diagrama de Clases

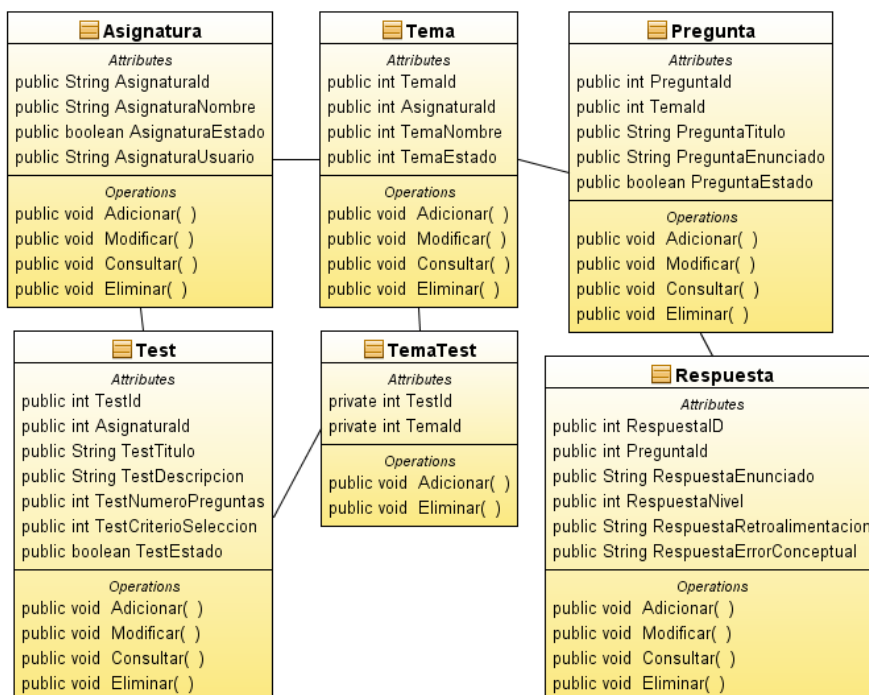


Figura 7 Diseño de clases; Fuente: elaboración propia

3.3 Implementación y Prueba de Unidades

Se hace realidad el modelo de diseño, es decir que se elaborarán y adaptarán los elementos gráficos y multimediales, se codificarán las páginas, los programas y scripts, se definen y preparan las bases de datos para que la herramienta quede en funcionamiento.

En esta etapa hay que realizar pruebas exhaustivas para asegurar el perfecto funcionamiento de la misma, todo esto se hace primero en el ambiente de desarrollo y luego en el servidor de Internet en el que realmente funcionará.

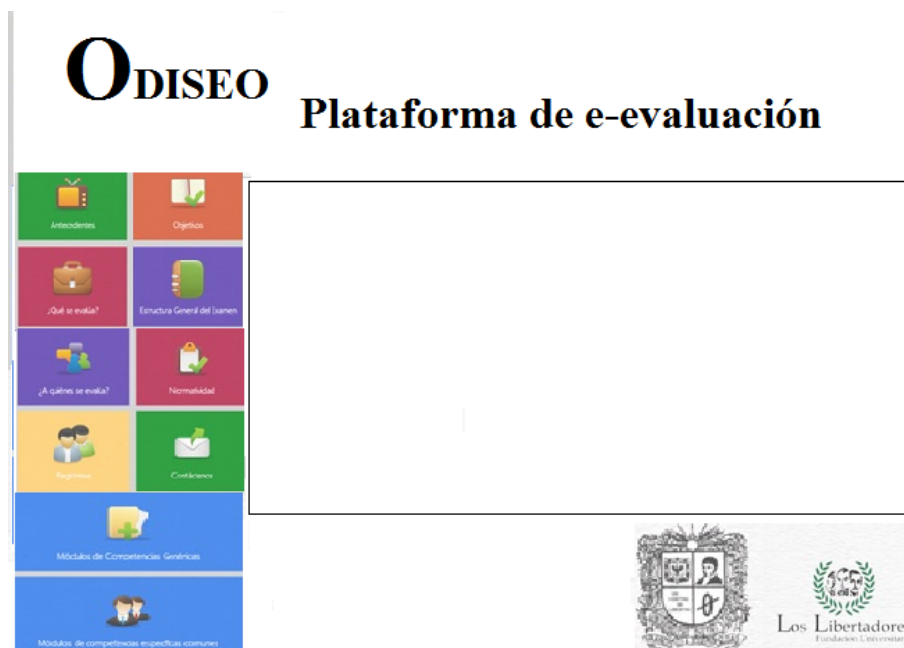
3.4 ODISEO: Algoritmo Evaluación Inteligente

Es una plataforma web que se empleara para administrar, gestionar, recursos, materiales, actividades, generar informes, distribuir y controlar los procesos de evaluación virtual.

Prototipo Algoritmo Evaluación Inteligente



ODISEO, está compuesto por tres roles: administrador de la plataforma, docente que puede administrar los test y ver estadísticas y estudiantes quienes pueden realizar los test y estar al tanto de su desempeño.



Conclusiones

Se asevera que los resultados de las Pruebas Saber Pro constituyen actualmente un instrumento para medir la *calidad* que da la institución por ende toda estrategia que se haga para potencializar los resultados obtenidos por los estudiantes se debe convertir en una herramienta que tenga todo el apoyo institucional.

Se comprende que es necesario aprovechar el empleo de las *TIC* como estrategia en los procesos a realizar para la preparación de las pruebas Saber Pro.

La E-evaluación es una alternativa que rompe barreras de tiempo y espacio, ya que está disponible 24 horas por 365 días.

En cuanto a la teoría del Ítem existen aplicaciones que se adaptan fácilmente para desarrollar y aplicar los test. Entre estas aplicaciones se encuentran el test adaptativo Informatizado TAI, el cual proporciona una evaluación de habilidades y conocimientos, de manera concreta, eficaz y efectiva, como también haciendo más rápido, eficiente el proceso de producción, aplicación y calificación de pruebas, sirviendo como sistema de información para que el docente conozca el nivel en que se encuentran los estudiantes y así pueda generar estrategias para guiarlos a alcanzar niveles superiores.

De esta manera se hace énfasis en la necesidad que las directivas institucionales apoyen proyectos encaminados a la investigación sobre los procesos de evaluación virtual, ya que falta indagar, analizar, diseñar, implementar, realizar pruebas, estandarizar, entre otros.

La plataforma ODISEO, Es una propuesta intuitiva de fácil manejo, en ella los roles de cada personaje involucrado pueden encontrar sus respectivos procesos, y ver estadísticamente el nivel de desempeño.

En la aplicación del pre test y post test de esta investigación se determina que los procesos realizados a nivel virtual son mucho más motivantes para los estudiantes ya que aumentan la autonomía e incorporan el trabajo colaborativo y responsabilidad, a los estudiantes les gusta trabajar herramientas virtuales por ende todas las evidencias confirman que los estudiantes participantes en la plataforma ODISEO desarrollaron significativamente más sus competencias, optimizaron y potencializaron los conocimientos, además los estudiantes presentan un alto grado de satisfacción en el manejo de la plataforma.

Referencias Bibliografica

[1] Reategui, M .Arakaki y C. Florez (2001) El reto de la evaluación. Lima; Plancad-GTZ, ministeriode educación.

[2] Orientaciones para el examen de Estado de calidad de la educación superior SABER PRO (ECAES) , Ingeniería de Sistemas 2010, ICFES

[http://www.acofi.edu.co/portal/documentos/Guia%20Ingenieria%20de%20sistemas\[1\].pdf](http://www.acofi.edu.co/portal/documentos/Guia%20Ingenieria%20de%20sistemas[1].pdf)

[3] Churchill, A. (2004). Ensuring quality in online higher education courses. University of Massachussets: Center for Education Policy. Disponible en www.umass.edu/senate/adhoc/online_report_full.pdf.

[4] Bunderson, C.V., Inouye, D.K. y Olsen, J.B. (1993). The Four Generations of Computerized Educational Measurement. En: R.L. Linn (Ed.) Educational Measurement (3rd ed.). New York: MacMillan Publishing Co., ps. 367-403

[5] Barberà, E. (2006, Julio). Aportaciones de la tecnología a la e-Evaluación. RED. Revista de Educación a Distancia, número especial VI. Revisado el 15/17/2013 en <http://www.um.es/ead/red/M6>

[6] Rodríguez Gómez, G., e Ibarra Sáiz, Ma. S. (Coord.) (2011). *e-Evaluación orientada al e-aprendizaje estratégico en Educación Superior*. Madrid: Narcea.

[7] Rocío Yuste, L. A. y Florentino Blázquez, C. (2012). La e-evaluación de aprendizajes en educación superior a través de aulas virtuales síncronas. *Revista Científica de Educomunicación, XX* (39).

[8] Lord, F.M. (1955). Esúmatmgtestreliabüity. *Educational and Psychological Measurement*, 15, 325-336.

[9] Rizopoulos, D. (2006). Irm: An R package for latent variables modeling and item response theory analysis. *Journal of statistical software*.

[10] Barrada, J. R., Olea, J., Ponsoda, V., y Abad, F. J. (2010). A method for the comparison of item selection rules in computerized adaptive testing. *Applied Psychological Measurement*, 34, 438-452

[11] Cattell, R. B. (1986,a). Scales and the Meaning of Standardized Scores. Johnson. *Functional Psychological Testing: Principles and Instruments*. New York: Brunner/Mazel, Publishers

[12] Arbero, M^a I. (1999). Gestión informatizada de bancos de ítems. En J. Olea, V. Ponsoda, y G. Prieto (ed.): *Tests Informatizados: Fundamentos y Aplicaciones*. Madrid: Pirámide

[13] [10] Julio Olea Díaz, Vicente Ponsoda Gil. (2013). *Test Adaptativos*, Universidad Nacional de Madrid, España, ISBN 978-84-362-6667-2

[14] Parshall, Cynthia G.(1995). *New algorithms for Item Seleccition and Exposure Control With Computerizes Adaptive Testing*, Universiti Utara Malaysia

[15] Giraffa, L.M.M.; Nunes, M. A.; Viccari, R.M. (1997). *Multi-Ecological: an Learning Environment using Multi-Agent Architecture MASTA'97: Multi-Agent System: Theory and Applications*, Proceedings Coimbra, Universidad de Coimbra

[16] Ataldi, Z; Salgueiro, F. y Lage, F. (2007). *Fundamentos para el Submódulo Evaluador en Sistemas Tutores Inteligentes: Diagnóstico, predicción y autoevaluación*. CACIC 2007.

[17] Garner, H. (2000) *La educación de la mente y el conocimiento de las disciplinas*, Paidós

[18] Gardner, H. (2002) *La inteligencia reformulada*. Paidós

[19] Russell, S. J. Y Norvig, P. (2003). *Artificial Intelligence: A Modern Approach* (2nd Edition). Prentice Hall

[20] Strickland, A.W (2006). "ADDIE". Idaho State University College of Education, Science, Math & Technology Education