

EVALUACION POR COMPETENCIAS UTILIZANDO LOGICA DIFUSA**Diana Patricia Bedoya Ruiz^a, Carlos Mario Sierra Duque^b**^a Estudiante Maestría en Ingeniería UdeA, dianabedoya570@gmail.com ^b Docente Facultad de Ingeniería U de A, csierrad@gmail.com

Abstract

The evaluation process of learning competences requires a comprehensive assessment of student that includes affective, cognitive and psychomotor aspects. Learning assessment systems have carried out the estimation of reached learning level by a student in a specific time and, as assessment dimension, only the cognitive aspect. This form of assessment has as shortcomings that ignores history and not takes into account learning process assessment of another dimensions of the training of apprentices.

One of the principles of the competences assessment is the search to articulate qualitative and quantitative assessment, for this there are different criteria of evaluation and to each of them a weight is assigned in the final evaluation. Evaluation criteria must be oriented to assess the development of the student in the different psychological domains: cognitive, affective, and psychomotor (CHURCES, 2009) allowing set the level of the competence. This paper presents a proposal for construction of a fuzzy rubric for the evaluation which aims to use the potential of traditional assessment rubric

Keywords: *Evaluation by competences, assessment, fuzzy inference rubrics.*

Resumen

El proceso de evaluación del aprendizaje por competencias debe incluir diferentes dominios psicológicos: cognitivo, afectivo y psicomotor. Los sistemas para la evaluación del aprendizaje, en su mayoría, efectúan una valoración del nivel de aprendizaje alcanzado por el estudiante en un momento específico y sólo sobre el aspecto cognitivo, omitiendo las otras dos dimensiones de la formación, el aspecto afectivo y el psicomotor.

Uno de los propósitos de la evaluación por competencias es la búsqueda de la articulación entre valoración cualitativa y cuantitativa; precisada una competencia, se plantean diferentes criterios de evaluación de desempeño y a cada uno de ellos se les asigna un peso en la evaluación final, dichos criterios deben estar orientados a valorar el desarrollo del estudiante en los diferentes dominios psicológicos (CHUCES, 2009), la valoración de estos puede hacerse de forma cuantitativa y/o cualitativa indistintamente, siempre y cuando se defina un mecanismo que permita establecer el nivel de desarrollo de la competencia. Este trabajo presenta una propuesta para la construcción de una rúbrica difusa de valoración cuyo objetivo es utilizar el potencial de las rúbricas de valoración tradicionales fusionándolas con las características de los sistemas de inferencia difusa.

Palabras clave: *Evaluación por competencias, Rúbricas de valoración, Inferencia difusa.*

Introducción

Los resultados de las pruebas Pisa para Colombia (Programme for International Student Assessment), han evidenciado que la ampliación de cobertura propuesta mediante el programa La revolución educativa (MEN 2010), no es suficiente para instaurar un adecuado sistema educativo. Con el propósito de aportar al mejoramiento de la formación de los estudiantes colombianos, el MEN, en dicho documento ha planteado la necesidad de “consolidar una cultura de la evaluación que se refleje en un sistema de evaluación centrado en el desarrollo de competencias” como parte esencial del ciclo de calidad.

La evaluación por competencias es un proceso que contempla una definición de resultados de aprendizaje, la identificación de criterios de desempeño, una medición –valoración- de saberes, un análisis de resultados, y una retroalimentación y orientación al estudiante durante el proceso de enseñanza aprendizaje. La evaluación por competencias plantea que la valoración de saberes debe realizarse de forma continua, lo que incide en la dificultad experimentada por los docentes al realizar el registro y procesamiento de la cantidad y variedad de aspectos – competencias, resultados de aprendizaje, criterio de evaluación, criterio de desempeño, entre otros- considerados en la evaluación de los estudiantes. La complejidad del proceso de valoración aumenta cuando a) el proceso debe realizarse mediante 3 subprocesos - a las que, en adelante, también llamaremos autovaloración, heterovaloración y covaloración -, b) considerar cada uno de los dominios psicológicos del aprendizaje -afectivo, cognitivo y psicomotor- y c) poder realizarse de forma cualitativa o cuantitativa indistintamente. En particular, la valoración hecha tanto en términos cualitativos como cuantitativos dificulta llevar a cabo una evaluación clara y eficaz de los estudiantes, pues en la aritmética tradicional no se dispone de herramientas –formales ni informáticas- para operar conjuntamente valoraciones lingüísticas y valoraciones numéricas.

El producto de investigación reportado en este artículo, pretende proveer al sistema de educación colombiano de un instrumento formal, en cuanto tiene un fundamento matemático riguroso, para realizar la etapa de valoración de saberes en un proceso de evaluación por competencias; mediante lo que denominaremos rúbrica difusa. La rúbrica difusa permite llevar a cabo valoraciones tanto cuantitativas como cualitativas, en los 3 dominios psicológicos del aprendizaje -afectivo, cognitivo y psicomotor- y efectuadas por 3 tipos de evaluadores - hetero, co y autoevaluadores.

1 Objetivo

Construir una rúbrica de evaluación borrosa que permita llevar a cabo el proceso de evaluación por competencias utilizando lógica difusa.

2 Desarrollo de la innovación

2.1 Aspectos de la evaluación por competencias

La evaluación por competencias es considerado un proceso de retroalimentación entre educador y educando en el cual se pretende determinar la suficiencia de los segundos, según se alcancen unos niveles de logro mínimos en los criterios de desempeño que conforman las competencias de referencia, cuando mediante estas competencias se exhiben las capacidades de los estudiantes para resolver tareas y problemas pertinentes. En ese sentido se debe tener una total claridad sobre para quién, para qué, por qué y cómo es la evaluación que deben tener los educandos (Parra, 2009)

Para diseñar un proceso de evaluación por competencias se deben definir los resultados de aprendizaje y los criterios de desempeño para la competencia a evaluar (Parra, 2009). “*Un resultado de aprendizaje es un enunciado escrito acerca de lo que se espera de un estudiante o aprendiz sea capaz de hacer al finalizar una unidad de un módulo/curso*” (Adam, 2004). Un criterio de desempeño es la descripción que expresa el tipo y grado de aprendizaje que se espera haya alcanzado el estudiante con respecto a las competencias expresadas como objetivos del módulo.

Junto con los resultados de aprendizaje, los criterios de desempeño especifican cuáles son los requerimientos mínimos para la obtención de la competencia, los cuales se establecen a través de los saberes esenciales: los requeridos para que una persona logre los resultados a evidenciar en el proceso. Estos saberes se clasifican en saber ser, saber conocer y saber hacer; deben ser tenidos en cuenta en el momento de diseñar los instrumentos de evaluación que le permitirán obtener las evidencias de desempeño de los alumnos; estas evidencias de desempeño son “pruebas claras y manifiestas de los conocimientos, habilidades, destrezas y actitudes que una persona posee y que determinan su competencia” (Tejada, 2011)

Tradicionalmente, para la valoración de competencias se ha recurrido a la construcción y uso de *matrices de valoración* o *rúbricas*, pues ellas facilitan la calificación del desempeño de los estudiantes a través de conjuntos de criterios de desempeño. Las rúbricas son escalas de medición usadas para valorar el desempeño, definidas como guías de puntaje; consisten en un conjunto de criterios preestablecidos, empleados en la valoración del trabajo del estudiante o productos resultantes de las tareas desempeñadas por ellos (Mertler, 2001).

Durante la valoración por competencias, en el proceso de evaluación, además del docente y el estudiante, se pueden ver involucrados actores como los compañeros de curso, otros docentes, entidades de control o representantes de organizaciones de naturaleza distinta a la educativa; esto provoca que la valoración por competencias requiera y pueda integrar las mediciones obtenidas en evaluaciones catalogadas como: autoevaluación, que busca fomentar en los estudiantes el autoconocimiento y la autorregulación; coevaluación que permite que los estudiantes se valoren entre sí, con el fin de posibilitar que reciban retroalimentación de sus pares; heteroevaluación que posibilita que los estudiantes reciban retroalimentación de personas diferentes a sus pares, como el docente de la asignatura, la institución, las empresas, o el Estado.

2.2 La incertidumbre en el proceso de valoración

Cada nivel de desempeño que puede asignarse como indicador de logro de un criterio de desempeño, manifiesta cierto grado de incertidumbre. Determinar valores numéricos únicos en muchas ocasiones impide aceptar que este tipo de procesos no es por naturaleza exacto. ¿Cómo saber, por ejemplo, que un estudiante ha alcanzado un nivel de aprendizaje que le merece una nota de 4.3, y no una de 4.1 o una de 4.2 en una escala 0-5? ¿Tendría más sentido emplear una valoración cualitativa como *alrededor de 4.3?* o incluso, *¿aplazado, aprobado, o excelente?*

Uno de los propósitos de la rúbrica borrosa propuesta es sustentar el procesamiento formal, con miras a superar el proceder intuitivo- de valoraciones cuantitativas y cualitativas; en tal sentido, se acudió al estudio y aplicación del marco teórico de los sistemas difusos. Aunque se han propuesto variadas maneras de incorporar el concepto de borrosidad en los procesos de evaluación educativos (Klir, Yuan, 1995), (Biswas, 1995), (Chen y Lee, 1997), (Weon y Kim, 2001), (Bai y Chen, 2008); (Saleh y Kim, 2009), (Yadav, Soni, Pal, 2014), éstas no permiten tener en cuenta la multidimensionalidad de los procesos evaluativos por competencias, ni la integración de los diferentes actores que participan en ella.

El instrumento de valoración propuesto, y descrito mediante este artículo, se enfoca en el uso de sistemas de inferencia difusa o borrosa. Un sistema de inferencia difusa se entiende como una articulación de elementos que pretenden realizar un proceso de razonamiento; el sistema se activa cuando de un conjunto de proposiciones borrosas requiere inferirse una conclusión que es, generalmente, difusa. Para el caso que nos atañe, una proposición borrosa-hace referencia a la valoración cualitativa o cuantitativa efectuada por uno de los tipos de actores, respecto del nivel de desempeño obtenido en una de las dimensiones de formación, para un criterio-de desempeño asociado con una competencia específica.

La comprensión de este tipo de sistemas requiere una fundamentación en la noción de conjunto difuso. La teoría de conjuntos difusos, introducida por Zadeh (Zadeh, 1965), permite representar de forma matemática *conceptos difusos*, entendidos como aquellos cuya delimitación de su significado es vaga.

Un conjunto difuso A en un universo de discurso X se determina por medio de una función de pertenencia μ_A , que toma valores en el intervalo $[0,1]$; $\mu_A(x)$ es el grado de pertenencia del elemento x al conjunto A . Formalmente, un conjunto borroso A se puede representar como

$$A = \{(x, \mu_A(x)) \mid x \in X\}. \quad (1)$$

2.3 Lineamientos para el diseño de la rúbrica difusa

Para el desarrollo del instrumento de valoración por competencias empleando un sistema borroso, al que llamaremos Rúbrica borrosa, es apropiado exponer algunas guías o lineamientos que establecen el desenvolvimiernto que debe seguir la rúbrica en un proceso de valoración.

El estado inicial del proceso de valoración se establece con el ingreso de un valor numérico o lingüístico – uno por cada una de las dimensiones de valoración (saber, ser y hacer) – que efectúan cada uno de los 3 tipos de evaluadores (auto, co y heteroevaluador) por cada criterio establecido para una competencia. A partir de acá se define lo siguiente:

El sistema debe posibilitar la especificación, en la escala de valoración, de valores cualitativos mediante conjuntos difusos; una especificación posible se muestra en la Figura 1. Cada institución educativa tiene la facultad de definir: la cantidad de conjuntos difusos (granularidad) por cada criterio de desempeño, la función de pertenencia correspondiente en la escala de valoración numérica, así como la cantidad de niveles de la escala numérica de valoración (su resolución). Las funciones de pertenencia de los conjuntos a representar son usadas durante todo el proceso, es decir, las diferentes valoraciones y los resultados obtenidos son dados en la misma escala. El sistema, para evitar inconsistencias, no acepta variación de ninguna escala de valoración en un mismo proceso valorativo.

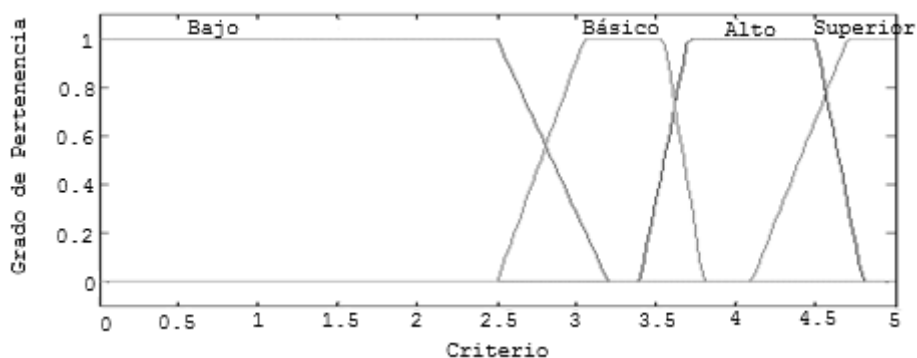


Fig 1. Posibles Funciones de pertenencia, niveles de la escala de valoración nacional en el rango de la escala numérica

El proceso de construcción de rúbricas de valoración, respeta la taxonomía de Bloom. Para esto se permite seleccionar el nivel de desarrollo de la competencia que desea evaluar y, según esa decisión, elegir el verbo más adecuado para la redacción del criterio de desempeño que valora la competencia. Si el verbo seleccionado no fue cargado en el repositorio del sistema, éste se marca como equívoco. La lista disponible de verbos a emplear en la rúbrica comprende una categorización planteada por Andrew Churches (Churches, 2009), que es acorde con la taxonomía.

El motor de inferencia permite hacer el cálculo difuso de la valoración final de un estudiante; este proceso requiere la definición previa de las escalas de valoración, y la definición de las bases de reglas difusas correspondientes a cada módulo. Para llevar a cabo el proceso de inferencia se debe:

Paso 1: Seleccionar la base de reglas con la cual se desea hacer la valoración.

Paso 2: Ingresar la valoración, numérica o lingüística, del responsable (actor, o tipo de valoración) para cada criterio de desempeño. Cuando un valor asignado es numérico, el módulo lo fusifica, es decir, se le establece el grado de pertenencia en el conjunto difuso estipulado para el criterio de desempeño en cada regla borrosa. Si la valoración ingresada es textual o lingüística, debe corresponder exactamente a una de las establecidas en el sistema, apreciables en la base de reglas, y el paso de fusificación no se lleva a cabo.

Paso 3: Evaluar cada regla una vez efectuada la fusificación: se identifica si para todos los criterios involucrados en el antecedente de la regla, el valor de pertenencia del elemento fusificado es mayor a cero; de ser así, se aplica el operador de implicación. El valor resultante se denominará valor de activación (o grado de cumplimiento);

Paso 4. Obtener la salida borrosa de cada regla. El grado de pertenencia de cada valor de la escala numérica del consecuente al nuevo conjunto borroso es el grado original cuando no supere el valor de activación de la regla, o es este último en caso contrario.

Paso 5: Agregar las salidas producidas por las reglas. Se estipula para cada elemento en la escala de valoración definido en salida, el máximo grado de pertenencia en los conjuntos borrosos resultantes de cada regla; este valor será el grado de pertenencia del elemento en el conjunto borroso final o agregado.

Paso 6. Calcular un valor concreto del conjunto borroso obtenido en el paso anterior.

La valoración del proceso de aprendizaje de un educando puede ser realizado por diferentes entes (estudiantes, docentes, entidades); cada valoración requiere la configuración de una rúbrica de valoración, la cual es procesada en el motor de inferencia, obteniendo para cada una de las rúbricas una valoración resultante. Para que el proceso de valoración final sea llevado a cabo, es decir para conocer cuál es la valoración final obtenida por el estudiante para un momento valorativo, se deben integrar los resultados de las valoraciones de los distintos entes; en este proceso se toman los valores obtenidos del procesamiento de las rúbricas específicas, una por actor, y se lleva a un nuevo proceso de inferencia, en el cual se define una base de reglas de 3 entradas, cada una de estas entradas está dada por los resultados de los procesos de valoración anteriores.

Finalmente, el método para la caracterización lingüística de los resultados establece cuál es la etiqueta lingüística –elemento del conjunto de términos- a la que más se acerca el valor numérico o cualitativo resultante del sistema de inferencia; para esto se calcula la distancia existente entre el valor resultante en el proceso de inferencia y los centroides de los conjuntos borrosos asociados con las etiquetas, y se selecciona aquella con la que se presente menor distancia.

No se plantean resultados lingüísticamente complejos porque los sistemas de valoración deben entregar los resultados en la escala de valoración lingüística estándar, definida por el ente rector de cada sistema educativo para facilitar la movilidad de los estudiantes a otras instituciones educativas.

2.4 Descripción del sistema de valoración propuesto: rúbrica borrosa

El modelo de valoración propuesto en este trabajo definirá, para una competencia cualesquiera c , un conjunto de dominios de valoración o dimensiones de formación $D_F = \{s, h, b\}$, donde s simboliza la dimensión del ser en la formación; h representa la del hacer, y b la dimensión del saber; un conjunto de actores de valoración $D_A = \{H, C, A\}$, heterovaloración (H), covaloración (C) y autovaloración (A); y 3 colecciones de criterios de desempeño, una para cada tipo de actor, $DC_i = \{1, \dots, d_i\}$; y $i \in D_A$.

Observe en la Fig un esquema que describe la arquitectura del sistema propuesto, cada módulo correspondiente a un tipo de valoración se divide en varios tipos de sub-módulos: el primer tipo de sub-módulo, **Integrador de las Valoraciones Asignadas en las Dimensiones de Formación**, recoge las valoraciones c_{ij_k} ; éstas corresponden a las efectuadas por el tipo de actor $i \in D_A$ para el criterio de desempeño $j \in DC_i$ en la dimensión de formación $k \in D_F$. De este sub-módulo se infiere la valoración para el criterio de desempeño j realizado por el evaluador i , c_{ij} . Un segundo tipo de sub-módulo, **Integrador de las Valoraciones Resultantes en los Criterios de Desempeño**, reúne las valoraciones c_{ij} y arroja el valor de la competencia procedente de cada tipo de actor c_i . El último tipo de sub-módulo, **Integrador de las Valoraciones Resultantes de cada tipo de Actor de Valoración**, agrega las valoraciones procedentes de los distintos actores de valoración c_i , para obtener la valoración final de la competencia c .

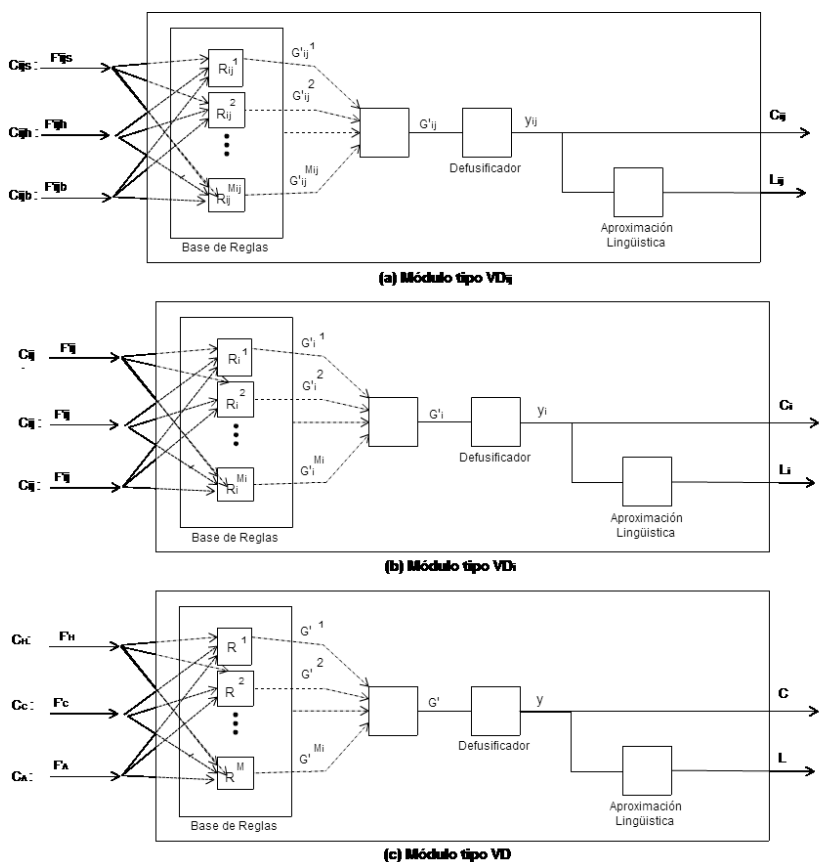


Fig 2. Representación gráfica del modelo matemático planteado

3 Resultados

Para la realizar las pruebas necesarias al funcionamiento de la rubrica borrosa se implemento una herramienta informática que permitiera llevar la validación y verificación con los diferentes grupos de docentes.

El proceso de validación y verificación del modelo se realizó en dos fases: i) Determinar que la lógica de valoración utilizada por los docentes puede ser modelada como un sistema difuso, ii) Verificar que el modelo difuso configurado de forma completa permite evaluar competencias, obteniendo resultados aceptados por expertos en valoración y evaluación.

Al interior de cada una de las fases fueron identificadas dos etapas: en la primera fase, la etapa 1 busca generar una configuración inicial, con la participación de un grupo de docentes, de los parámetros necesarios para el funcionamiento del motor de inferencia, estos son los diferentes conjuntos difusos con sus funciones de pertenencia y las bases de reglas. La etapa 2 pretende validar dicha configuración a través de la realización de una encuesta a docentes. Para la segunda fase, su primera etapa establece que el grupo de docentes que validó la configuración inicial-etapa 2, fase 1-, deberá seleccionar los resultados que consideran más adecuados para una etapa de un proceso valorativo, mientras que en la segunda etapa se analiza los resultados entregados por la rubrica borrosa, para un proceso valorativo completo, con la ayuda de un grupo de expertos en evaluación.

Fase 1. Modelado del método de valoración docente.

Etapa 1. Configuración de los parámetros de entrada del motor de inferencia. La validación del modelo permitió llevar a cabo una definición de funciones de pertenencia de cada uno de los conjuntos borrosos asociados con las etiquetas lingüísticas a emplear. Con ese objeto se llevó a cabo una reunión inicial con un grupo de 5 docentes, quienes por consenso determinaron las siguientes funciones de pertenencia: bajo= $\{0/0; 0/0; 0/2.5; 0/3.2\}$; básico = $\{0/2.5; 1/3.05; 1/3.55; 0/3.8\}$; alto= $\{0/3.4; 1/3.7; 1/4.5; 0/4.8\}$; superior= $\{0/4.1; 1/4.7; 1/5; 0/5\}$. Estas funciones serán utilizadas a lo largo de todo el proceso de valoración en los diferentes módulos. Sin embargo, la herramienta informática le permite a las diferentes instituciones educativas establecer sus propias funciones de pertenencia de ser necesario.

Una vez definidos los conjuntos difusos, es importante realizar la definición de la base de reglas que utiliza el motor de inferencia. Para esto se llevó a cabo una definición inicial, la cual fue presentada al grupo de 5 docentes anteriormente mencionado, y se llevaron a cabo los ajustes propuestos para la configuración del motor de inferencia.

Para la primera fase del proceso de validación del modelo propuesto se empleó un módulo encargado de integrar las valoraciones de los criterios de desempeño generadas por un único actor, dado que tiene como principal objetivo verificar la pertinencia de los resultados obtenidos con la configuración resultante del trabajo realizado con los docentes para el sistema de inferencia. La razón para esto es que la escala de valoración a usar debe ser la misma a lo largo de todo el proceso valorativo, y la forma de construcción de la base de reglas se mantiene para los módulos restantes del sistema, siempre y cuando no haya una política interna que determine una escala de valoración variable para la institución. De presentarse este caso, se debe llevar a cabo el mismo proceso de validación para todas las escalas y bases de reglas definidas.

En la etapa 2 de la fase 1 se realizó una encuesta a 52 docentes de colegios públicos del área metropolitana del Valle de Aburra que imparten asignaturas en las áreas de matemáticas, español, ética, y tecnología para los grados 10 y/o 11. Las Instituciones Educativas en las cuales trabajan los docentes encuestados hacen parte de la Alianza Futuro Digital Medellín y por esto están familiarizados con un método para la construcción de sus rubricas de valoración. Esta precisión es importante pues no existe ninguna directiva gubernamental que precise la forma de llevar a cabo la valoración de los estudiantes al interior de los programas académicos, lo que dificulta encontrar una

población de docentes que utilice las mismas formas de evaluación. Se pretende, entonces, identificar posibles variaciones en las funciones de pertenencia y la base de reglas definidas previamente.

Se solicitó a los docentes realizar una nueva definición, de las funciones de pertenencia de los diferentes conjuntos difusos pertenecientes a la escala de valoración a utilizar; para llevarla a cabo, se determinan en cada conjunto los distintos puntos de referencia (o vértices), obteniendo los siguientes resultados al utilizar el estadístico de la media: bajo={0/0.53; 1/0.9;1/2.51; 0/3.12}; básico={0/2.86; 1/3.17; 1/3.58; 0/3.85}; alto={0/3.75; 1/3.98;1/4.46;0/4.71}; superior={0/4.47; 1/4.7;1/4.96;0/5}; dichos resultados permiten evidenciar que para los docentes encuestados la escala valoración no debe iniciar en 0. Ésta valoración normalmente se relaciona con la inasistencia del estudiante a la actividad valorativa. Sin embargo, exceptuando el conjunto bajo, las funciones de pertenencia para los otros conjuntos no sufren cambios significativos.

Con el objeto de validar o de mejorar una base de reglas (columnas 1, 2 y 3 en la Tabla 1: 16 combinaciones distintas de valoraciones en 3 criterios de desempeño) que pretende determinar el nivel de desempeño en una competencia cualquiera –el consecuente de cada regla- se le presentó al grupo de docentes previamente mencionados la siguiente pregunta: “Suponga que se encuentra en un momento valorativo en el que se busca evidenciar el resultado de aprendizaje de una competencia específica, y para ello se emplean tres criterios de desempeño de igual importancia. Por cada combinación presentada en las 3 primeras columnas, seleccione la valoración resultante para la competencia”. Para la prueba fueron seleccionados un conjunto de 16 combinaciones que hicieran una muestra representativa de la forma de razonamiento difuso, pues la base de regla sería muy extensa para la validación exhaustiva. Una vez recogidos los resultados del nivel de desempeño asignados por los docentes para cada combinación o antecedente de regla, se contrastaron con los que se habían "asumido" inicialmente (ver columna 4, Tabla1); el porcentaje de concordancia se muestra en la columna 5, Tabla 1.

Se optó como acción de mejoramiento de la base de reglas, mantenerlas en su estructura original, cuando lograron una concordancia de al menos un 60%; de lo contrario, el nivel de desempeño se modificó de manera que el nuevo valor resultante de la evaluación de la regla fuera aquel con mayor representación de la lógica de valoración utilizada por los docentes, los resultados muestran que el proceso de valoración del sistema de inferencia difuso, presenta aceptación en 11/16 opciones presentadas en términos de valoraciones lingüísticas. Sin embargo, la base de reglas puede ser definida por la institución educativa de acuerdo a sus políticas de valoración.

La disparidad de los resultados (mayor concordancia 94.92%, menor concordancia 28.81%, concordancia media: 63.24; desviación estándar: 21.59), evidencia el nivel de subjetividad que tiene el proceso de valoración; se puede esperar, incluso, que de implementarse el sistema para llevar a cabo la valoración de estudiantes reales, el nivel de subjetividad del proceso valorativo docente no disminuya.

Fase 2. Verificar que el modelo difuso permite evaluar competencias.

Etapa 1. Selección de los resultados más adecuados para una etapa del proceso valorativo. Se lleva a cabo el análisis de los resultados obtenidos por el sistema difuso al ser usado para el procesamiento de valoraciones cuantitativas, se presentaron a los profesores 16 combinaciones de valores numéricos para 3 criterios de desempeño (columnas 1, 2 y 3, Tabla 2) y los resultados obtenidos utilizando los métodos de valoración rubrica difusa, promedios y rubrica tradicional, (columnas 4, 6, 8, Tabla2), para la valoración numérica lograda de una competencia cualesquiera asociada

Tabla 1 Resultados obtenidos del proceso de validación de la base de reglas inicial.

Valoración Criterio de desempeño 1	Valoración Criterio de desempeño 2	Valoración Criterio de desempeño 3	Resultado Sistema de inferencia	
			resultado base de reglas 1	Porcentaje Concordancia
Superior	Alto	bajo	Básico	61,02
Superior	Superior	bajo	Alto	71,19
Superior	Bajo	bajo	Básico	42,37
Superior	Alto	básico	Alto	71,19
Superior	básico	básico	Básico	71,19
Superior	básico	superior	Alto	81,36
Superior	básico	bajo	Básico	86,44
Superior	Alto	alto	Alto	81,36
Superior	superior	alto	Alto	28,81
Alto	Alto	bajo	Básico	66,10
Alto	básico	bajo	Básico	84,75
Alto	Alto	básico	Básico	35,59
Alto	Bajo	bajo	Bajo	33,90
Alto	básico	básico	Básico	94,92
Básico	Básico	bajo	Bajo	33,90
Básico	Bajo	bajo	Bajo	67,80

En las rubricas de valoración que se ha denominado en este artículo como “tradicional” (correspondiente a “rúbrica no borrosa”), los docentes, para cada uno de los criterios de valoración, asignan un puntaje de acuerdo con el desempeño de los estudiantes, luego se realiza la sumatoria de los puntos obtenidos y estos se llevan a una nota numérica, para el caso puntual de la rúbrica de referencia, el cálculo de la nota final se realiza así:

$$n = \begin{cases} 0 & p < c \\ 3 + \frac{(p-(c*2))*2}{c} & p \geq c \end{cases} \quad (2)$$

Donde

n, es la nota obtenida por el estudiante.

p, es el total de puntos obtenidos por el estudiante

c, es el número de criterios definidos en la rúbrica

(ROSERO, 2012)

El método de valoración que se ha denominado en este artículo como “promedio ponderado”, busca encontrar una valoración definitiva operando valoraciones de criterios de desempeño de distinta importancia o peso, así:

$$n = \sum_i w_i * c_i \quad (3)$$

Donde

n , es la nota obtenida por el estudiante.

w_i , es el peso (valor en el intervalo [0,1]) asignado al criterio de valoración i . Además, deben cumplir con la siguiente restricción: $\sum w_i = 1$

c_i , es el valor numérico asignado al criterio de valoración i

Finalmente, se ha denominado “rubrica difusa” al método expuesto a lo largo del artículo. Para la pregunta “Suponga que se encuentra en un momento valorativo en el cual se busca evidenciar el resultado de aprendizaje de una competencia específica, y para ello se emplean tres criterios de desempeño de igual importancia. Por cada combinación presentada en las 3 primeras columnas de la Tabla 2, seleccione dos resultados (por fila) de las columnas 4, 6, 8; Tabla2, que considere cercanos a la valoración numérica que asignaría Ud. al resultado de aprendizaje de la competencia.”

Una vez tabuladas las respuestas se obtienen por cada combinación los resultados presentados en las columnas 5, 7, 9, tabla 2. En general cuando a los docentes se les pide seleccionar dos notas como posibles resultados de la valoración en resultados numéricos, tienden a promediar las notas dadas, por tanto es de esperarse que el resultado seleccionado por la mayoría de los docentes fuese el valor obtenido por el método del promedio, Promedio/Fuzzy: 17/0; Una vez seleccionado este valor, los docentes deben analizar los dos resultados restantes, para establecer cuál de ellos representa mejor, según su criterio, el nivel de logro de la competencia; los resultados obtenidos son: Fuzzy/Rubrica: 16/1, se puede inferir, que para el procesamiento de valoraciones numéricas, el resultado entregado por el método difuso, es más acorde a la forma de razonamiento utilizada por la mayoría de los docentes.

Etapa 2, analizar los resultados entregados por la herramienta, para un proceso valorativo completo. Los resultados obtenidos del trabajo realizado con los docentes para la validación de la base de reglas y de los conjuntos difusos, fueron sometidos a la inspección de un grupo de docentes de pedagogía involucrados en la investigación de herramientas y técnicas para la evaluación. Se les hace la presentación el prototipo informático que implementa el modelo matemático planteado, para integración de los diferentes actores del proceso valorativo, permitiéndoles hacer diferentes simulaciones. Con ellos se obtiene la siguiente retroalimentación: “El modelo encontrado puede ser llevado a la Escuela con facilidad porque permite que cada profesor o colectivo de profesores configuren su sistema de evaluación de acuerdo con la normatividad existente, con la propuesta curricular de la institución y con la concepción que se tenga de la evaluación”. (Zapata, 2014)

“El modelo matemático encontrado exige que el profesor o colectivo de profesores que lo utilice reflexione sobre la evaluación en el contexto del currículo, esto representa una gran ventaja porque ayuda a generar una cultura de la Evaluación con la que aún no se cuenta en la mayoría de las instituciones educativas.” (Zapata, 2014)

Tabla 2 Resultado de selección de valoraciones obtenidas con los diferentes métodos difusos, promedio, rúbrica

Valoración numérica Criterio de desempeño 1	Valoración numérica Criterio de desempeño 2	Valoración numérica Criterio de desempeño 3	Valoración numérica Resultado de aprendizaje 1: Rúbrica borrosa		Valoración numérica Resultado de aprendizaje 2: Promedio		Valoración numérica Resultado de aprendizaje 3: Rúbrica tradicional	
			nota	Votos	nota	Votos	nota	Votos
1	2,9	3,5	3,2	31	2,47	38	0	3
4,7	4,5	3,1	4,22	30	4,1	38	3,67	8
3,1	4,1	4,7	4,1	34	3,97	38	3	4
3,1	4,5	4,3	4,16	34	3,97	37	3	2
4,5	2,7	4,9	4,18	30	4,03	38	3	6
4,3	0,5	4,3	4,1	22	3,03	38	1,67	12
3,5	3,5	4,5	4,16	13	3,83	38	3,67	23
3,3	4,9	2,9	4,1	30	3,7	37	1,67	5
0,5	3,9	3,5	3,49	23	2,63	38	1,67	14
4,1	4,7	3,1	4,1	31	3,97	38	3	7
1,5	3,3	4,5	3,63	33	3,1	38	1,67	5
3,5	3,5	3,9	4,1	21	3,63	38	3	15
2,5	4,9	4,3	4,1	33	3,9	38	2,33	4
4,5	2	2,7	3,55	34	3,07	38	0	2
4,3	2	4,5	4,1	28	3,6	38	2,33	7
2,7	4,7	2	3,55	33	3,13	38	0	3
4,7	3,7	4,3	4,41	37	4,23	38	0	1

4 Conclusiones

Los resultados obtenidos en la primera fase de la segunda etapa de validación permiten identificar una marcada diferencia en los resultados obtenidos entre los diferentes métodos de valoración utilizados, incluso existe una diferencia muy significativa entre los dos métodos de valoración tradicionalmente utilizados: valoración por promedios y rúbricas no borrosas, lo que significa que la selección de cualquiera de los métodos puede presentar una variación importante en la valoración final que obtendrá el estudiante. Como los métodos de valoración tradicionales son indistintamente utilizados en las diferentes instituciones y en muchos casos son mezclados en el proceso de valoración no es viable asumir ninguno de los métodos como el más adecuado para la valoración del aprendizaje de los estudiantes, esto dependerá de las exigencias y las características deseadas por cada docente u institución educativa.

De acuerdo con los resultados obtenidos en el proceso de validación con docentes y la retroalimentación dada por el grupo de expertos en evaluación una vez utilizada la herramienta, se puede establecer que el modelo de valoración difuso planteado hace una

representación aceptable del proceso de valoración por competencias y facilita la valoración de los estudiantes en términos cualitativos.

Dado que el modelo planteado pretende hacer una representación de la forma de razonamiento docente, el correcto funcionamiento de éste se encuentra condicionado por la correcta construcción de las funciones de pertenencia que representarán la escala de valoración lingüística, al igual que la definición de la base de reglas a utilizar; en ella se deben contemplar todos los posibles casos pertinentes y debe plasmar el nivel de exigencia deseado por la institución, esto determinará el nivel de exigencia a tener en cuenta en el modelo de valoración difuso.

El modelo para valoración por competencias utilizando lógica difusa, empleando una herramienta informática, permite al docente mejorar la eficiencia del proceso de valoración, pues solo requiere determinar cuáles son los criterios de valoración a tener en cuenta en cada momento y le posibilitará tener un conocimiento discriminado de los resultados obtenidos por el estudiante por cada uno de sus dominios. Si bien disponer de una herramienta informática como la propuesta en este trabajo presenta facilidades desde lo pedagógico para efectuar los procesos valorativos, y estandariza los resultados de la medición para los estudiantes en cuanto a la valoración cualitativa, el proceso es altamente costoso en términos computacionales, debido al crecimiento de la base de reglas al aumentar el número de criterios de valoración.

4.1 Trabajo futuro

El presente proyecto abre las puertas para ahondar en diferentes aspectos de la valoración utilizando métodos de inferencia difusa, sería de gran interés diseñar el modelo matemático que permita asignar diferentes pesos a los criterios de evaluación que son utilizados durante el proceso de valoración, así como la posibilidad de definir los pesos de las valoraciones emitidas por los diferentes actores involucrados.

Dado el costo computacional del proceso de valoración, llevar a cabo un seguimiento del estudiante no solo requiere el diseño e implementación de una arquitectura que soporte computación distribuida para el procesamiento, sino que demanda el diseño de un repositorio de almacenamiento de información bien estructurado para permitir el procesamiento de los resultados obtenidos por el estudiante en cada momento valorativo para llevar a cabo la valoración final. Una vez establecidos estos mecanismos se puede pensar en integrar un proceso semi-evaluativo al sistema de valoración.

REFERENCIAS

ADAM, S. (2004) "Using Learning Outcomes: A consideration of the nature, role, application and implications for European education of employing learning outcomes at the local, national and international levels." . United Kingdom Bologna Seminar, Herriot-Watt University

CHURCES, A. *Taxonomía de Bloom para la era digital*.<
<http://www.eduteka.org/articulos/TaxonomiaBloomDigital>> [Julio, 2012]

MERTLER, C. A. *Designing scoring rubrics for your classroom. Practical Assessment, Research & Evaluation*, 7(25). <http://PAREonline.net/getvn.asp?v=7&n=25> [Consulta: Abril 19, 2014]

PARRA A. C. *EVALUACION POR COMPETENCIAS PROPUESTA DESDE EL DECRETO 1290*. Alianza Medellín Futuro Digital. 2009

ZADEH L. A. (1965): Fuzzy Sets. *Information and Control*, 8, pp. 338-353

ZAPATA, D. Validación Pedagógica del Modelo para evaluación por competencias utilizando lógica difusa. [Realizado: 14 de Agosto de 2014.]

BAI, S.-M; Chen, S.-M. "Evaluating students' learning achievement using fuzzy membership functions and fuzzy rules". *Expert Systems with Applications*, 34, 399–410

BISWAS, R. (1995) "An application of fuzzy sets in students' evaluation", *Fuzzy Sets and Systems*

Chen, S. M., & Lee, C. H. (1999). "New methods for students' evaluating using fuzzy sets". *Fuzzy Sets and Systems*.

KLIR, George, YUAN, Bo (1995). *Fuzzy Sets and Fuzzy Logic*. Prentice Hall.

SALEH, I; KIM, S. (2009) "A fuzzy system for evaluating students' learning achievement". Elsevier

TEJADA, J. (2011) "La evaluación de las competencias en contextos no formales, dispositivos e instrumentos" *Revista de Educación* N° 354. Ministerio de Educación. España

WEON, S. KIM, J. (2001) "Learning achievement evaluation strategy using fuzzy membership function". *31st ASEE/IEEE Frontiers in Education Conference*. 2001

ROSERO, C. *Estrategia Metodológica Para El Seguimiento Al Desarrollo De Competencias En Estudiantes De La Media Técnica Del Proyecto Alianza Futuro Digital Medellín (Afdm)*. Trabajo Final de Máster, Universidad EAFIT, Escuela de Posgrado, Medellín, Colombia. 2012

YADAV, R. SONI, A. PAL, S. (2014) "A study of academic performance evaluation using Fuzzy Logic techniques", *IEEE Xplore*