

SISTEMATIZACIÓN DE UNA EXPERIENCIA DE EVALUACIÓN CUALITATIVA. HACIA UNA NUEVA CONCEPCIÓN DEL PROCESO EDUCATIVO

Felipe Tirado Segura*
Alejandro Miranda**
y Ana Elena del Bosque***

* Profesor Titular "C", de Tiempo Completo Definitivo en la UNAM-FES-IZTACALA, jefe del programa de Investigación Psicoeducativa, miembro del Sistema Nacional de Investigadores.

Correo e: ftirado@unam.

** Académico en la UNAM-FES-IZTACALA, psicólogo dedicado a los estudios de la educación en línea y los procesos culturales en internet.

Correo e: correo@alejandromiranda.org

*** Profesor Asociado "B" de Tiempo Completo en la UNAM-FES-Iztacala, en la Carrera de Psicología. Colaboradora en diversos programas de investigación. Correo: anedelbosque@gmail.com

REVISTA DE LA EDUCACIÓN SUPERIOR

ISSN: 0185-2760

Vol. XL (3), No. 159

Julio - Septiembre de 2011, pp. 9 - 28

Ingreso: 08/09/10 • Aprobado: 08/07/11

Resumen

En este trabajo se reporta una investigación en la que se sistematiza un procedimiento de evaluación cualitativa a partir de la valoración hecha por los propios estudiantes de ensayos realizados por sus compañeros. El propósito es inducir el aprendizaje de competencias complejas, tales como la reflexión, el análisis crítico, la argumentación, la formación de juicios sustentados, a partir de la elaboración y evaluación de ensayos colectivos.

Palabras clave:

- Evaluación cualitativa
- Competencias complejas
- Aprendizaje

Abstract

This paper reports on the research of a systematized qualitative assessment evaluation process carried out by the students themselves on their peers' essays, as a way to induce the learning of complex skills, such as reflection, critical analysis, arguing and issuing supported judgments, through collective essay writing and evaluating.

Key words:

- Qualitative evaluation
- Complex competences
- Learning

Marco conceptual

Los sistemas escolarizados están altamente estructurados y formalizados, por lo cual resultan ser rígidos para transformarse y poder afrontar los continuos y acelerados cambios que se presentan en la actualidad. La transformación resulta necesaria e impostergable tanto para resolver problemas como para obtener ventajas de las nuevas circunstancias. Un problema es la enorme demanda desatendida, por lo que se requiere encontrar vías que permitan ampliar la matrícula y a su vez mejorar la calidad educativa. Se considera que esto será posible si se saben aprovechar las múltiples ventajas que ahora ofrecen las tecnologías digitales para *distribuir la enseñanza* y desarrollar nuevas concepciones del proceso educativo.

El modelo educativo no tiene porqué restringirse al esquema presencial escolarizado, está la oportunidad para aprovechar la interconexión en línea y construir modalidades de enseñanza mixta. Los modelos mixtos o mezclados, comúnmente referidos en inglés como *blended learning* o *b-learning*, combinan la enseñanza a distancia y/o en línea con la presencial. Se asume que al vincular estas modalidades es posible aprovechar las ventajas propias de cada una de ellas. Es cierto que los modelos mixtos han coexistido en la educación a distancia desde hace mucho tiempo, fue recurrente que los contenidos y ejercicios educativos se ofrecieran por vía postal y se asistiera de manera regular, esporádica o voluntariamente a tutorías presenciales. Sin embargo, hoy se han transformado profundamente las circunstancias de la educación a distancia debido a las posibilidades que ofrece la enseñanza asistida por cómputo, particularmente cuando se tiene conexión a internet (García-Aretio, Ruiz y Domínguez, 2007).

En las décadas de los años sesenta a ochenta se desarrolló todo un campo de experiencias orientado a valerse de la instrucción asistida por computadora (*CAI-Computer Aided Instruction*), que bajo los planteamientos de las *máquinas de enseñanza* sustentados en el *conductismo* de la década de los cincuenta, se concibieron maneras para personalizar la enseñanza al ofrecer a cada alumno la posibilidad de avanzar a su propio ritmo en el tiempo y sitios que mejor le convinieran.

Para los años noventa la velocidad de los nuevos procesadores en cómputo, la expansión exponencial en la capacidad de memoria, el aumento de la calidad de los recursos multimedia (imágenes-audio-video), los vínculos de hipertexto e hipermedia, las interconexión vía internet, producen una revolución en la producción y divulgación de la información que rápidamente se trata de aprovechar para mejorar las estrategias educativas, lo que da lugar a planteamientos diversos como *Computer Based Education*, *Computer Based Training*, *Computer-Based Learning*, *E-learning*, *Teaching-Learning Environments*, etcétera.

En particular, la comunicación vía internet es vista como una gran ventaja para los sistemas de *educación a distancia*, en tanto otorga la posibilidad de comunicarse de manera *asincrónica*, es decir, sin tener que interactuar de

forma simultánea; además de abatir el problema de las distancias para entablar comunicación de manera expedita y con múltiples interlocutores, lo cual permite el desarrollo de sistemas gestores de *redes sociales*. Estas propiedades dan lugar a nuevos planteamientos educativos como el *Computer Supported Collaborative Learning* (Bartolomé, 2004; Stegmann, Weinberger y Fischer, 2007; Laurillard, 2009; Ligorio y Ritella, 2010; Mihaela y Stoicescu, 2011; Whittaker y Bonakdarian, 2011) los cuales pueden ser combinados (mezclados) con las modalidades de *educación presencial*. Bajo este modelo se sustenta la experiencia a la que se refiere este trabajo.

Los sistemas escolarizados suelen estar concebidos bajo un paradigma presencial convergente, en el cual un grupo de alumnos en un salón de clases es requerido a centrar su atención en la exposición de los argumentos sustentados por un parlante, que suele ser el profesor. Con el tiempo se ha observado que es un sistema que tiende a inhibir la divergencia de ideas por las dificultades naturales que hay para operar sincrónicamente con la diversidad, ya que no se puede atender diversos puntos de vista al mismo tiempo, motivo por el cual no es un paradigma que aliente la creatividad y la crítica de los estudiantes. Las más de las veces los argumentos son sustentados por el profesor, por lo cual el modelo se centra en la función docente y no en las actividades o aprendizajes de los propios alumnos.

En un proceso formativo lo más relevante es lo que hacen los estudiantes y no los profesores (Biggs, 2003). El desarrollo de la comprensión en la apropiación del conocimiento, así como el pensamiento crítico y creativo, se tiene que dar en cada alumno a partir de sus propias reflexiones, de aquí que se requiere de un modelo activo centrado en el alumno, que le permita construir y desarrollar sus conocimientos, sus creencias, habilidades y actitudes, tal y como se sustenta en el modelo *constructivista* de la educación (Estes, 2004). Por esto se requiere disponer de un *modelo divergente* en que cada alumno tenga la oportunidad de desarrollar sus propias reflexiones y creatividad, en lugar de uno centrado en la función docente. El planteamiento es construir modelos de *enseñanza divergente y distribuida* entre los propios estudiantes (Tirado, Bustos y Miranda, 2007).

La multiplicidad de planteamientos que pueden surgir cuando cada alumno tiene oportunidad de desarrollar sus propias ideas, hace imposible que un solo profesor pueda atenderlas, de aquí que una dinámica alternativa puede ser la *distribución* de tareas y la integración de los alumnos en equipos de trabajo. Estos equipos pueden ser convocados para elaborar ensayos por colaboración en torno a los temas curriculares del curso, seleccionados los temas de acuerdo con la preferencia de los alumnos, lo cual abre la posibilidad de tener un *modelo divergente* para que cada estudiante pueda desarrollar su creatividad bajo sus propios argumentos y analizarlos en colaboración con sus compañeros, lo que hace un modelo de *enseñanza distribuida* centrado en la *co-construcción* (construcción por colaboración con pares). La *co-construcción* de ensayos puede ser regulada e inducida por medio de criterios (*rúbricas*) que se presenten al inicio del curso y que se definan como los criterios tanto para el trabajo como para la evaluación, de manera que se constituyen en los

parámetros guía y normativos de las actividades del curso, porque guían la ejecución de los alumnos además de ser las metas esperadas a alcanzarse en el curso y a ser evaluadas.

La elaboración de ensayos es un recurso muy favorable para promover la reflexión, el juicio crítico y la creatividad, en tanto implica la manifestación de las ideas por escrito. La escritura es el mejor instrumento del pensamiento, es un medio ideal para desarrollar, precisar, expresar y comunicar las ideas propias (Olson, 1998). La escritura es un recurso que permite repensar lo escrito, lo cual admite que sea ajustado el texto tantas veces como se desee, hasta poder quedar a la plena satisfacción de quien lo escribe, pudiendo así lograr su mejor expresión. Si la expresión oral de ciertas ideas es transcrita textualmente y se compara con el mismo planteamiento pero ahora expresado a través de un escrito realizado con cierto esmero, se hacen evidentes las grandes diferencias que se logran a favor de la expresión escrita.

La elaboración de ensayos por *colaboración* opera en un proceso de *cognición distribuida*, donde los procesos cognitivos se distribuyen en entorno a la colaboración, el comportamiento del grupo unido por fines comunes no operan como la suma de un conjunto de individuos aislados, sus acciones se dan dentro de un el entorno sociocultural en el que son influidas y potenciadas por el grupo, dadas las interacciones que se dan entre las personas y los recursos disponibles (propósitos, contenidos, materiales, artefactos, normas, tareas, ideas, argumentos, contraargumentos), llevados éstos a acciones (elaborar un ensayo) donde aprender y hacer son inseparables (“*learning by doing*”). Se trata de un ejercicio propicio para promover el contrapunto de ideas y criterios, la reflexión dialógica, el uso de la argumentación para sustentar propuestas, lo que posibilita la producción de trabajos más sólidos en tanto es el producto de la suma de ideas, de la conjunción de esfuerzos que permite alcanzar mejores resultados (Bearison y Dorval, 2002). Es claro que dos personas piensan más que una y tres más que dos.

El trabajo por colaboración también permite fomentar la interlocución respetuosa, asumir compromisos y promover la responsabilidad compartida, con lo cual cambia el enfoque de la evaluación que suele centrar la atención sólo en los resultados de contenidos académicos, por otra orientada a considerar el proceso de construcción por colaboración como proceso formativo, también se promuevan principios sustantivos para contribuir en la edificación de una ética social promotora de prácticas de rendición de cuentas y convivencia cívica (Tirado, 2007).

Baker (2009) hace una serie de señalamientos sobre la importancia de la *evaluación formativa* y la *rendición de cuentas* como procesos que pueden mejorar la instrucción, al tomar medidas que incidan en la enseñanza dentro del aula. Señala también Baker que un sistema de *rendición de cuentas* implica hacer más responsable a las personas y puede contribuir a la confiabilidad y validez de la evaluación. La evaluación puede servir a distintos propósitos, debe haber “sensibilidad instructiva”, ya que resulta razonable crear oportunidades para aprender de los materiales que son medidos en las pruebas.

Se requiere de evaluaciones sensibles en las que haya oportunidades para que los estudiantes realicen tareas significativas de alta calidad, que reflejen el dominio en torno a los contenidos complejos de un curso, que permita la evaluación de las prácticas en las aulas, las tareas y trabajos realizados por el estudiante.

No sin razón han señalado Canales y Gilio (2008) que la evaluación en México se ha concentrado en propiedades cuantitativas del sistema educativo, en factores asociados al rendimiento, tales como número de reprobados, número de aciertos en las pruebas *PISA*, número de estudiantes por profesor, número de profesores de tiempo completo, número de profesores con posgrado, número de publicaciones, etcétera, y se ha descuidado la evaluación en sí misma del proceso de enseñanza–aprendizaje, de manera que este tipo de evaluación no permite incidir en el mejoramiento del proceso.

El concepto de evaluación ha pasado por diferentes momentos que lo han ido modificando, durante varios siglos la evaluación se centró en el acto de juzgar, gradualmente se reemplazó por nociones de medición más técnicas en la que fue ganando una visión psicométrica, la cual hoy en día sigue vigente aunque ahora se complementa con otros elementos que se centran en la evaluación del proceso formativo con categorías más de índole cualitativa. De esta forma se han construido diversas comparaciones, una centradas en la norma o distribución de poblaciones y otras referidas a criterios de logro.

Ahora se intenta dar otro paso hacia una evaluación centrada en la ejecución de respuestas construidas, donde el alumno demuestre sus *competencias* para hacer las tareas que se espera pueda ejecutar, en los contextos donde se requieren realizar, por lo que se le ha referido a este planteamiento como *evaluación auténtica*. En esta nueva visión se busca que quienes evalúan dejen de ver al proceso evaluativo como algo independiente a la enseñanza y el aprendizaje, se enfatiza que la evaluación se realice como parte integral del proceso educativo (Ahumada, 2005; Díaz Barriga, 2006; Castillo y Cabrerizo, 2007).

La evaluación tiene como una de sus funciones sustantivas conocer la manera en que se realiza el proceso educativo con el fin de poder mejorar, de aquí que Ahumada (2005) sugiere concebir una enseñanza y evaluación en términos de dominio de procesos, de manera personalizada y diferenciada; dar mayor peso a la función diagnóstica que a la función administrativa de la acreditación, considerar el carácter productivo y valerse de técnicas e instrumentos evaluativos múltiples que permitan la *evaluación auténtica* como alternativa, de manera tal que se evalúe la puesta en práctica de programas basados en *competencias*, en el que los alumnos comienzan desde un principio afrontando tareas asignadas con la intención de resolver problemas auténticos, haciendo que el aprendizaje sea más funcional, aplicado y significativo. De esta manera se está centrado en los procesos del aprendizaje, en el quehacer del alumno y no en el del profesor, por lo que constituye una evaluación formativa en lugar de sumativa.

Una evaluación de este tipo debe ofrecer la información a los alumnos sobre su proceso de aprendizaje, de tal manera que pueda realizar las correc-

ciones que sean pertinentes durante el proceso. Es importante que el alumno asuma la responsabilidad de su propio aprendizaje y utilice esta valoración como un medio que le permita alcanzar los conocimientos y competencias propuestas, por lo que la participación de los alumnos es central en esta modalidad evaluativa, enfatizando así la posibilidad de la *autoevaluación* y la *co-evaluación*.

La evaluación generalmente está a cargo de los profesores, quienes además son los responsables de traducirla a una calificación para la acreditación del curso; sin embargo, este proceso puede socializarse, incorporando a los alumnos para contribuir en ella. La *co-evaluación* conlleva varias ventajas como son: tener una participación más activa por parte de los estudiantes, hacer horizontal el proceso de evaluación, que los alumnos aprendan a valorar temas que previamente han estudiado y preparado, lo que promueve la reflexión crítica y responsable al requerir afinar y sustentar sus juicios, además de constituir un repaso de lo previamente estudiado. Esto favorece la autorregulación, meta-cognoscitiva de su aprendizaje, al requerir percatarse de aciertos y errores, además puede aumentar los elementos de confiabilidad y validez de la evaluación como indica Baker (2009).

Castañeda (2006) señala que en una comunidad de aprendizaje y práctica en la que se compartan metas, tareas y responsabilidades, la *co-evaluación* representa una herramienta importante para apoyar a los estudiantes a que se familiaricen con los criterios, valores y metas del aprendizaje, así como para desarrollar habilidades de discusión y argumentación, permite centrarse en la forma de realizar las actividades asignadas para el aprendizaje a lo largo del curso, y deja de ser una simple forma de asignar una *calificación* al finalizar el curso.

Muchos profesores se resisten a la *co-evaluación* por suponer que es riesgosa, consideran que los juicios que tomen los alumnos pueden no estar bien sustentados, carecer de elementos, tener sesgos debidos a simpatías o antipatías, o bien pactar beneficios por mutua conveniencia. Para que la *co-evaluación* sea efectiva, es importante solicitar al alumno que exponga argumentos fundamentados y señale la evidencia empírica en la que sustenta sus apreciaciones. Esto implica que el alumno tiene que reflexionar sobre lo que hace y la manera en que lo hace, así como en los resultados de su actividad para valorar, requiere tener claro que no se evalúa a la persona, sino a las actividades, los trabajos y los productos realizados.

Para hacer la *co-evaluación* los estudiantes deben contar con los criterios correspondientes, éstos deben ser explícitos y puestos a su consideración desde el inicio del curso. Para ello el profesor debe previamente planear y redactar dichos criterios a través de *rúbricas*, que constituyen guías con escalas de evaluación en las que se establecen los niveles de dominio relativos al desempeño que un alumno debe mostrar con respecto al proceso y productos determinados. Las *rúbricas* representan una evaluación formativa en tanto permiten determinar la calidad de la ejecución basada en el ejercicio de la reflexión crítica, al evaluar tareas que no implican respuestas correctas o incorrectas, sino la apreciación dentro de un amplio rango de criterios que

qualifican de modo progresivo el tránsito de un desempeño que va de *incipiente* y que sería típico de un *novato*, hasta el nivel de un *experto* que tiene pleno dominio, los cuales son referidos dentro de un rango de criterios de apreciación (muy mal–mal–regular–bien–muy bien), que pueden traducirse a una relación de magnitud numérica ordinal (1–2–3–4–5) que permita la operatividad matemática.

Los criterios responden a preguntas como qué aspectos caracterizan la ejecución de un especialista o experto, cuáles son las características que distinguen entre una ejecución excelente, buena, promedio y deficiente. Es importante que se revisen los criterios junto con los estudiantes antes de llevar a cabo la tarea o actividad que será evaluada a través de este instrumento, de tal manera que se conozcan con antelación los criterios con los que serán evaluados los desempeños; lo que permite reducir la subjetividad en la *co-evaluación* (Díaz Barriga y Hernández, 2000; Quesada, 2003; Ahumada, 2005; Díaz Barriga, 2006).

Procedimiento

El desarrollo de la experiencia que se sustenta en este trabajo está basado en las consideraciones antes referidas a partir de tres planteamientos. El primero es que la evaluación no sólo puede dar cuenta del proceso educativo, sino también puede inducirlo. El segundo es que el proceso se puede inducir a partir de la elaboración de ensayos colectivos y la *co-evaluación* de estos ensayos que implican tareas complejas, por ser la expresión escrita de ideas donde se requieren aportaciones críticas y creativas expresadas bajo argumentos bien sustentados. El tercer planteamiento es que los estudiantes son capaces de hacer la *co-evaluación* de manera responsable, sustentada, válida y confiable.

Evaluar la calidad de un ensayo es una tarea compleja, pues es ineludible que se basa en juicios de apreciación que son subjetivos, en tanto se trata de valorar la calidad de un cuerpo complejo de ideas. Sin embargo, es posible hacer una evaluación menos subjetiva si se acota con criterios claros previamente definidos (*rúbricas*) y se basa en la apreciación colegiada expresada por juicios diversos, lo que no puede ocurrir cuando sólo evalúa un profesor. La evaluación colegiada permite neutralizar, si no al menos atenuar, parte de la subjetividad si se ciñe a criterios y se expresa el sustento en que se fundamentan los juicios, lo cual además también puede ser evaluado (*meta-evaluación* = evaluación de la evaluación).

En nuestra concepción se parte de que la *co-evaluación* realizada por los pares ofrece tres grandes ventajas, una es que se juzga a partir de circunstancias semejantes y compartidas (*pares*), la otra consiste en que al ser colegiada se obtiene una pluralidad de juicios, y la más relevante es que incide en la propia formación de los alumnos. Es bien sabido que lo que es evaluado cobra importancia, por lo tanto se le presta atención y se procura realizar las acciones esperadas, para así obtener una buena nota. De aquí que la evalua-

ción determina en buena medida el quehacer de los estudiantes, por lo mismo la estrategia de evaluación puede ser diseñada y utilizada para inducir el aprendizaje, dando a conocer previamente los criterios y procedimientos que serán utilizados. Por esto es conveniente proporcionar a los estudiantes los criterios normativos que se usarán en las *rúbricas*, con el fin de que desde el inicio del curso traten de cumplir con estos preceptos en el proceso de elaboración de sus ensayos. Por ejemplo, a nuestros estudiantes se les indicó que en el ensayo se calificaría la *Organización*, la *Claridad*, la *Suficiencia*, la *Bibliografía*, etcétera.

La evaluación se ha limitado a valorar el logro del aprendizaje y no a inducir el aprendizaje, cuando las dos funciones pueden ser compatibles y complementarias. La capacidad de los alumnos para participar en el proceso de evaluación (*co-evaluación*) ha sido un recurso educativo poco aprovechado, porque, como ya se indicó, se tiene la creencia de que si los alumnos se evalúan serán parciales e injustos, que se regalarán calificaciones altas en un acto de complicidad, o por el contrario, que otorgarán notas reprobatorias como represalia cuando hay rivalidades. Este estudio demuestra que esto no es así, o no tiene porque ser así, ya que se pueden desarrollar procedimientos bien sistematizados para que la *co-evaluación* sea realizada con responsabilidad y de manera sustentada.

El transferir funciones de la evaluación a los estudiantes constituye un principio formativo de relevancia educativa, en tanto se deposita la confianza en las capacidades de los alumnos para hacer juicios sustentados y de manera responsable, promueve su capacidad para ejercer una valoración justa. Al hacer la evaluación horizontal se hace transparente y se atenúa el poder del profesor para asentar calificaciones a su sólo arbitrio, lo que también es relevante en tanto contribuye a la formación cívica del estudiante.

En nuestro procedimiento, los equipos para hacer un ensayo colectivo se integraron a partir de la elección de un tema del curso que fuera de su interés. Todos fueron expuestos en los seminarios presenciales a lo largo del semestre. Para cada tema se forman al menos tres equipos, de manera que el primer equipo pueda evaluar al segundo, el segundo al tercero y el tercero al primero, de este modo se pueden hacer evaluaciones cruzadas y preservar el anonimato. Además, al evaluar los ensayos de equipos y no de individuos, se desvanece aún más el posible efecto perverso de otorgar calificaciones bajas basadas en rivalidades personales, o poner calificaciones altas para recibir en retribución buenas notas. Pero lo más relevante de este procedimiento es que todos los alumnos prepararon un tema para hacer su propio ensayo colectivo a lo largo de un semestre, tema que es igual al del ensayo que le toca evaluar, de manera que está particularmente familiarizado con la temática, incluso es muy probable que lleguen a tener más conocimientos puntuales y actualizados sobre el tema que el propio profesor del curso. De esta manera la *co-evaluación* también constituye un repaso del tema trabajado durante el semestre, pero ahora visto desde otra perspectiva, a partir de la lectura crítica y reflexiva que les demanda formarse juicios sustentados propios en torno

a los diferentes aspectos que son sometidos a su evaluación por medio de *rúbricas*. De acuerdo con la taxonomía de Bloom *et al.* (1956) y de su revisión (Anderson *et al.*, 2001), el hacer juicios (evaluar) constituye una de las competencias cognitivas de mayor relevancia, por lo que se debe inducir en el proceso formativo.

Gracias a los recursos computarizados mediados en línea es posible, en un aula virtual, promover la *co-evaluación* de manera fácil y expedita, lo que además forma un ambiente que contribuye a que los alumnos ganen competencias digitales. De este modo se puede llevar el registro pormenorizado de cada acción en bases de datos relacionadas que permiten transparentar todo el proceso, lo que en sí es de gran relevancia para la legitimidad y credibilidad social de la evaluación, pero además se pueden utilizar formatos electrónicos que son fáciles para responder y que corresponden al ambiente digital con el que ahora están familiarizados los estudiantes.

Por ejemplo - Contribución: Se ofrecieron aportaciones originales en el trabajo

Totalmente de acuerdo De acuerdo En desacuerdo Totalmente en desacuerdo

Argumento que sustente tu apreciación: _____

También se pueden valorar aspectos importantes para la colaboración y así inducir en los estudiantes el trabajo colectivo:

Por ejemplo - Respeto: Se usaron adjetivos calificativos despectivos frente a los argumentos expuestos

Totalmente de acuerdo De acuerdo En desacuerdo Totalmente en desacuerdo

Argumento que sustente tu apreciación: _____

El trabajo por colaboración y la *co-evaluación* constituyen parte central de los objetivos del curso, en tanto se busca desarrollar dos principios formativos en los estudiantes, uno es el sentido de responsabilidad frente a los compromisos que se deben asumir, y el otro el respeto basado en el aprecio y consideración de sus compañeros, para inducir la predisposición para construir acuerdos por consenso.

Hipótesis:

Si se sistematiza la evaluación a partir de equipos colegiados (opiniones diversas), bajo criterios específicos bien definidos (*rúbricas*), con un procedimiento en el que se preserva el anonimato (imparcialidad), se pueden obtener evaluaciones cualitativas bien sustentadas sobre tareas complejas (elaboración de ensayos) a partir de los juicios de los propios estudiantes.

Escenario:

La investigación se realizó con estudiantes de dos grupos de licenciatura que cursaban el quinto semestre de la carrera de psicología durante el ciclo escolar 2010-1, en la UNAM (FES Iztacala). El curso se diseñó bajo una modalidad mixta que contempla tanto seminarios presenciales en los que se revisan todos los contenidos curriculares del curso, así como actividades que se realizan en un *aula virtual* en línea (se utilizó una plataforma Moodle) donde se construyen y evalúan los ensayos realizados por colaboración.

Con el propósito de desarrollar competencias de trabajo en grupo, los ensayos son elaborados en equipo (de 4 a 5 integrantes) utilizando una plataforma de escritura por colaboración (*Wiki*) dispuesta en el aula virtual, en la que hay también foros y *chats* para apoyar la deliberación en torno al desarrollo del ensayo.

Se proporciona a los alumnos, en el aula virtual, el listado de los temas curriculares del curso y se les pide que les otorgue números ordinales de acuerdo a su predilección, de forma tal que se le pueda asignar a un equipo de trabajo que realizará un ensayo en torno a una de sus preferencias temáticas. Se integran al menos tres equipos (A, B, C,...) por cada tema, de manera que el equipo A califica el ensayo del equipo B, el B el de C y éste el de A, de modo que nunca se evalúan recíprocamente los mismos equipos, evitándose posibles sesgos por complicidad.

Los alumnos califican electrónicamente en el aula virtual, donde se tiene a la vista las *rúbricas*, las cuales consideran 7 propiedades del ensayo a evaluar. Los rubros son *Organización*: El orden o estructura de la exposición de las ideas es adecuado; *Claridad*: Está bien escrito, se comprende con facilidad el texto; *Suficiencia*: Se tocaron los puntos más relevantes del tema; *Tratamiento del caso*: Se abordó y analizó apropiadamente un caso práctico desde el planteamiento del marco conceptual del trabajo; *Contribución*: Hay aportaciones originales de los autores; *Conclusiones*: Hay una buena síntesis y se resaltan las aportaciones del trabajo apoyadas en el caso; *Bibliografía*: Las referencias bibliográficas son pertinentes al tema y actualizadas. La apreciación en cada rubro se dio con base en una escala tipo *Likert* con *respuesta sustentada*, donde se expresa el grado de acuerdo que se tiene con la afirmación que se proporciona y se escribe la razón en que se sustenta la apreciación.

Por ejemplo - Contribución: Hay aportaciones originales de los autores

Totalmente de acuerdo De acuerdo En desacuerdo Totalmente en desacuerdo

Argumento que sustente tu apreciación: _____

Finalmente, al concluir el curso se invita a los alumnos a que evalúen diferentes atributos de la experiencia del curso utilizando una escala tipo *Likert* como la antes descrita. Uno de los atributos es que expresen sus opiniones con respecto a la experiencia de la *co-evaluación*, para lo cual se formulan 8 preguntas específicas.

Por ejemplo - Favoritismo: Algunos compañeros califican con favoritismo

Totalmente de acuerdo De acuerdo En desacuerdo Totalmente en desacuerdo

Resultados

El análisis de la metaevaluación (evaluación de la evaluación) que se reporta en esta investigación se hizo bajo el principio de la correspondencia y congruencia de las evaluaciones emitidas por los alumnos, con base en dos momentos y condiciones. La primera fue valorar la consistencia de las calificaciones asentadas por los estudiantes. En la segunda se hizo la valoración de estas calificaciones con base en los argumentos expresados como sustento, a partir de las estimaciones realizadas de manera independiente por tres profesores del curso (juez A, juez B y juez C).

Sujetos

Se trabajó con dos grupos (a y b) de la carrera de licenciatura en psicología, uno integrado por 36 alumnos (a) y el otro por 27 (b). De los 63 alumnos inscritos 12 desertaron (19%), los 51 restantes quedaron integrados en 14 equipos para elaborar los ensayos temáticos, procurando un arreglo en que quedaran integrados más o menos el mismo número de alumnos del grupo “a” y del grupo “b”.

Rigor de la evaluación de los alumnos

El primer análisis que se realizó fue valorar la consistencia de las evaluaciones que hicieron los alumnos con respecto a los ensayos de sus compañeros que les tocó evaluar. La primera observación fue apreciar si las calificaciones otorgadas eran diferenciadas (que hubiera varianza), que no tuvieran efectos de “techo” al conceder sólo calificaciones extremas altas, o a la inversa, efectos de “piso” al otorgar únicamente calificaciones muy bajas. Sólo tres alumnos no tuvieron varianza ya que en todos los rubros asentaron la misma calificación. Otros tres hicieron una autoevaluación de su ensayo y no del elaborado por sus compañeros, por lo cual se descartaron y no fueron considerados en estos análisis.

El promedio de calificaciones de todos los ensayos fue de 49.9 en una escala del 1 al 100, es decir que en promedio no son altas ni bajas calificaciones, porque oscilan en un valor intermedio. El ensayo que recibió la calificación más alta fue de 57.7 puntos y la más baja 40.2, lo que no hace una gran diferencia (17.5 puntos). La varianza promedio fue de 0.43, lo que refleja que hubo heterogeneidad en las calificaciones al reconocer diferencias en la calidad de los componentes de los ensayos evaluados. Esto rompe el prejuicio que suele tenerse en el sentido de que si los alumnos evalúan se regalarán las calificaciones, o por el contrario, que reprobarán a su compañeros dadas las posibles rivalidades. En ningún alumno se observó este tipo de comportamiento, y se puede afirmar por las calificaciones otorgadas que evaluaron con rigor.

La *confiabilidad* es una condición sustantiva en todo procedimiento de evaluación. Se refiere a qué tanto se puede confiar en una evaluación, qué tan certero es el diagnóstico. Por ejemplo, si se tienen dos jueces que evalúan una misma prueba, sería de esperar que los dos hagan el mismo diagnóstico, será confiable si otorgan la misma calificación, o al menos una muy parecida. Un procedimiento muy recurrido es contar con tres jueces y apreciar si éstos coinciden en su evaluación, o al menos dos de ellos, de aquí que el tener equipos de estudiantes para evaluar ofrece una condición muy favorable para estimar la confiabilidad.

Para valorar el nivel de *confiabilidad* de las calificaciones otorgadas por los diferentes alumnos se hicieron las comparaciones de sus evaluaciones. De las 322 (46 por 7) calificaciones otorgadas por los 46 alumnos en los 7 rubros (1- Organización, 2- Claridad, 3- Suficiencia, 4- Tratamiento del caso, 5- Contribución, 6- Conclusiones, 7- Bibliografía), coincidieron en 203 ocasiones, es decir en el 63% de los casos hay acuerdo. Para tener un indicador de la consistencia interna de las calificaciones se obtuvo el valor Alpha de Cronbach, que fue de 0.668. Aunque este valor es aceptable, no es muy alto, pero se trata de 14 ensayos diferentes, de aquí que se obtuvieron las correlaciones para cada equipo que evaluó el mismo ensayo, correlacionando sus calificaciones y las calificaciones promedio en cada uno de los 7 rubros, asumiendo que el promedio corresponde a la mejor estimación del consenso del grupo. El pro-

medio de todas las correlaciones fue de 0.59, el alumno que obtuvo la mejor correlación fue de 0.97, 19 alumnos (37%) obtuvieron muy buenas correlaciones por ser superiores a 0.70, 17 (33%) estuvieron entre 0.50 y 0.70, los restantes 15 estudiantes (29%) obtuvieron correlaciones relativamente bajas (entre 0.50 y 0.20).

Estas estimaciones permiten apreciar que en el 70% de las evaluaciones realizadas por los alumnos tienen una alta consistencia, fueron hechas con responsabilidad y rigor, por lo que se consideraron válidas y confiables.

Valoración de la evaluación de los alumnos

Para hacer una estimación más cuidadosa del rigor de las evaluaciones realizadas por los estudiantes, tres jueces, los profesores del curso, revisaron de manera independiente uno por uno los argumentos que ofrecieron los alumnos para sustentar la calificación que habían otorgado. Esta evaluación se hizo apreciando si se habían apegado los estudiantes a los criterios explicitados en las *rúbricas*. La estimación se asentó con una calificación en base a una escala tipo *Likert* de 4 valores (0 – muy mal, 1 – mal, 2 – bien y 3 – muy bien).

A la hora de revisar los argumentos de sustento de la evaluación, como ya se indicó, se apreció que tres alumnos se habían autoevaluado, no comprendieron que se trataba de evaluar el ensayo realizado por sus compañeros y no el suyo, de manera que fueron descartados en estos análisis, por lo que el número de evaluaciones valoradas corresponde a 48 estudiantes en 7 rubros, lo cual hace un total de 336 juicios emitidos, número suficientemente amplio para poder valorar la calidad de las evaluaciones realizadas por los estudiantes.

Una forma de expresar el propósito de este análisis de manera sintética es demostrar si los profesores podían reconocer la calidad de las evaluaciones realizadas por los estudiantes, al coincidir (confiabilidad) en reconocer quiénes eran buenos, regulares o malos evaluadores, por apegarse a los criterios expuestos en las *rúbricas* para cada rubro.

La coincidencia entre las apreciaciones de los profesores se estimó a partir del grado de correspondencia, medido con base en el número de *acuerdos* (coincidencias en las calificaciones otorgadas), por las *correlaciones* con el promedio generado por los 3 jueces (profesores) y las calificaciones de cada profesor, así como por la heterogeneidad de calificaciones asentadas medida por la *varianza*, y finalmente, por las *diferencias entre las medias* de las calificaciones otorgadas por cada profesor con base en un análisis de varianza (correspondiente a la hipótesis nula: hay diferencias en los juicios de cada evaluador).

Número de acuerdos:

Con base en el procedimiento tradicional en psicometría para estimar la confiabilidad a partir de tres jueces, se contaron el número de acuerdos cuando los tres o dos de ellos otorgan la misma calificación. De los 336 argumentos

que sustentaban las calificaciones otorgadas por los alumnos (48 alumnos por 7 rubros = 336), los tres jueces otorgaron la misma valoración en 32 ocasiones (9.5%), coincidieron al menos dos de ellos en 224 veces (66.7%) y hubo desacuerdo entre los tres en 80 casos (23.8%), lo que permite apreciar que en la mayoría de las ocasiones (76.2%) al menos dos profesores estuvieron de acuerdo en otorgar la misma calificación. De las 256 (32 + 224) ocasiones en que al menos dos de los jueces estuvieron de acuerdo con la misma valoración, el juez A coincidió en el 69.5% de las ocasiones (178 veces), el juez B coincidió en el 60.9% de las veces (156), y el juez C en el 82% (210). Considerando la escala acordada (0 – muy mal, 1 – mal, 2 – bien y 3 – muy bien), de los 48 estudiantes ninguno resultó ser “muy mal” evaluador a juicio de los 3 jueces, 5 (10.4%) fueron considerados “malos”, 39 (81.2%) “buenos” y 4 (8.3%) fueron valorados como “muy buenos”. Es decir que prácticamente el 90% (89.6%) de los alumnos fueron considerados entre buenos y muy buenos evaluadores, habiendo discrepancias en el 10% restante de los casos.

La media de las evaluaciones emitidas por los tres jueces en los 7 rubros fue de 1.99, lo que de acuerdo con la escala nuevamente corresponde a “bien”. En el rango de variación se aprecia que la calificación promedio que menos se distanció con respecto a la media de los tres jueces fue el juez C (- 0.15), y la que más la del juez A (0.53). Considerando que los valores plausibles están entre 0 y 3, significa que esta desviación (0.53) con respecto a la media fue de sólo el 17.6%.

Correlaciones

Se correlacionaron las calificaciones de cada juez con el promedio de las 3 calificaciones otorgadas a cada alumno (48) en cada uno de los 7 rubros. Todas las correlaciones fueron positivas y la mayoría por arriba de 0.70, lo que implica que hubo una relación de correspondencia muy alta. El juez A obtuvo una correlación promedio de 0.78, la más baja fue de 0.57 en el rubro de *Conclusiones* y la más alta de 0.87 en *Bibliografía*. El juez B obtuvo una correlación promedio de 0.81, la más baja fue de 0.73 en el rubro de *Conclusiones* y la más alta de 0.87 en *Suficiencia*. El promedio del juez C fue de 0.69, siendo la más alta de 0.87 en *Tratamiento del caso* y la más baja de 0.38 en el rubro de *Claridad*. Esta correlación hace apreciar que este juez en este rubro se distanció notoriamente de la evaluación promedio, pero fue la única correlación que obtuvo un valor menor a 0.50. Todas las correlaciones son estadísticamente significativas al nivel de $p < 0.01$. En este análisis el juez B fue el mejor, en tanto es el que se acercó más a la varianza de las calificaciones promedio.

Varianza

Se puede plantear que un juez que no genere varianza en sus apreciaciones resulta ser un mal juez, en tanto todo le parece igual y por lo mismo no dis-

crimina. Entre menor es la varianza significa que el juez otorgó calificaciones más homogéneas, y a la inversa, entre mayor es el valor de la varianza las calificaciones son más heterogéneas. Se calculó la varianza para cada uno de los tres jueces en los 7 rubros, así como la varianza promedio para cada uno de ellos. El juez A obtuvo una varianza promedio de 0.51, en el caso del juez B fue de 0.50 y en el del juez C de 0.51, es decir la heterogeneidad de las calificaciones fue prácticamente la misma en los tres casos.

La varianza debe ser valorada sistémicamente tomando en cuenta otros indicadores de ponderación como se ha hecho en este estudio (número de acuerdos, correlaciones, tamaño de las discrepancias, diferencias con respecto a la media), ya que la homogeneidad en las calificaciones no necesariamente indica que el juez no tiene capacidad para discriminar, porque sería factible, aunque poco probable, que haya un desempeño muy bueno o muy malo en todos los rubros, por lo tanto un buen juez podría otorgar una calificación homogénea y no significa necesariamente que ha calificado mal.

Diferencia entre las medias

Finalmente, otra manera de valorar el nivel de acuerdo que hay entre dos jueces es hacer una prueba para estimar si la diferencia entre las medias y la varianza de sus calificaciones resulta ser estadísticamente significativas con respecto a la del otro juez, de manera que si no hay diferencias significativas se puede decir que las calificaciones son equivalentes, que la diferencia entre las calificaciones de un juez y el otro resultan estar dentro de un rango de varianza que no es estadísticamente significativo, se podría decir que no importaría qué juez evaluó porque el resultado sería prácticamente el mismo. Esto se estimó corriendo un análisis de varianza de un factor, con comparaciones múltiples *post hoc*, Scheffe, con un nivel de significación de $p < 0.05$ (Silva, 2004).

Para hacer este análisis se utilizó el procesador estadístico SPSS, comparando en cada uno de los 7 rubros las medias del juez A con las del juez B y C, así como las del juez B y C. En las 14 comparaciones posibles para el juez A (7 rubros con juez B y 7 con juez C), sólo en un caso no hubo diferencias significativas con el juez C. El juez B y C no tuvieron diferencias significativas en 6 de los 7 rubros, es decir en el 85.7% de las ocasiones sus juicios fueron equivalentes. En este análisis el juez C fue el mejor, dado que en 7 rubros sus calificaciones promedio fueron estadísticamente iguales con las de alguno de los otros dos jueces.

Este análisis resulta muy riguroso porque asume que el juicio de apreciación está dado en una misma escala, lo que resulta en una visión de objetividad excesiva, en tanto hay elementos de la varianza que corresponde a una estimación diferenciada en la valoración de cada juez, por ejemplo, en nuestro caso se observa que el juez B y C prácticamente no tuvieron diferencia al coincidir en el 86% de los casos, en cambio el juez A sólo coincidió con el juez C en un caso. Al ver las medias de las calificaciones el juez A tiene un valor de 2.51, el juez B de 1.61 y el juez C de 1.84, se puede decir que el juez B tuvo un criterio más riguroso en sus apreciaciones, que es cercano al del

juez C, en cambio en la apreciación del juez A le parecieron mejor hechas las evaluaciones de los alumnos. Es como utilizar dos reglas al medir una misma altura, una pautada en centímetros y la otra en pulgadas, de modo que siempre habrá menos pulgadas que centímetros, pero las evaluaciones podrían ser totalmente equivalentes. Otra manera de interpretar es como si un juez estima que 6 centímetros es lo suficientemente alto y otro estima que esto ocurre con 8 centímetros. Esta intersubjetividad es una fuente de varianza incontrolada en la valoración de cualidades en tareas complejas. Lo que uno debe esperar es que hay al menos correspondencia ordinal: estos son malos evaluadores, estos regulares, estos buenos y estos muy buenos, reflejado por las correlaciones.

Opinión de los alumnos de la co-evaluación

Para finalizar se analizan las 8 preguntas en las que se solicitó a los estudiantes opinaran sobre la experiencia de la *co-evaluación*. Esta valoración se hizo de manera abierta y voluntaria una vez concluido el curso, el número de estudiantes que ingresó al aula virtual fue ya reducido ($N = 25$), no obstante esta información nos permite tener una apreciación de las opiniones prevalentes en torno a la *co-evaluación*.

Las preguntas y respuestas fueron las siguientes. Primera pregunta). Si consideraban que algunos compañeros califican con favoritismo. La mayoría (40%) se abstuvo contestando la opción “No sé”, que quizás es la mejor respuesta pues ciertamente no lo sabían, el 32% manifestó que no y 24% que sí. Segunda pregunta). Si era conveniente que los alumnos contribuyan a la co-evaluación. El 68% manifestó su acuerdo, el 16% expresó su desacuerdo y el resto se abstuvo (18%). Tercera pregunta): Si la manera de co-evaluar era la apropiada. 52% indicó que sí, 36% que no y el resto se abstuvo. Cuarta pregunta): Si era conveniente hacer la co-evaluación. 64% manifestó su acuerdo, 24% su desacuerdo, las demás fueron abstenciones. Quinta pregunta): Si consideraba que era objetiva. La mayoría (52%) consideró que sí, 12% que no y el restante 36% indicó “No sé” o se abstuvo de responder. Sexta pregunta): Si algunos compañeros califican con mala fe. 52% indicó que no, 12% contestó que sí y el restante 36% se abstuvieron. Esta pregunta es complementaria con la primera, es interesante apreciar que hay una opinión más inclinada a pensar que se llega a calificar con favoritismo, que con mala fe. Séptima pregunta): Si los alumnos con supervisión de los profesores deberían ser los responsables de la evaluación. Aquí las opiniones están divididas, 44% está de acuerdo y en el mismo porcentaje están en desacuerdo. Octava pregunta): Si los alumnos deberían ser siempre los únicos responsables de la evaluación. En esta pregunta el 88% está en desacuerdo, sólo un 12% opinó que sí, llama la atención que no hay abstenciones.

Es interesante apreciar con estas opiniones que los alumnos tienen incertidumbre sobre el arbitraje de sus compañeros, y que claramente prefieren que la evaluación sea supervisada por los profesores, lo que nos parece una indicación relevante.

Al final del cuestionario de salida, se invita de manera abierta a expresar observaciones, consideraciones críticas o sugerencias en torno al curso. En esta sección sólo dos alumnos hicieron una alusión a la *co-evaluación*, uno expresó sus reservas a que se les califique sobre su participación en la elaboración de un ensayo colectivo y que no sea una evaluación de su trabajo personal. El otro señalamiento apunta a que se debería combinar la *co-evaluación* con la del profesor. De hecho esto es así, la *co-evaluación* de los ensayos es un puntaje base el cual es ponderado con muchos otros elementos, tales como participación y contribuciones personales, cumplimiento con las lecturas del curso, participación en los seminarios, asistencia y puntualidad, entre otros indicadores de desempeño.

Conclusiones

Hay tres problemas de gran relevancia en la educación superior en México. Uno es el alto número de jóvenes en edad escolar que no estudian, muchos de ellos por deserción y otro tanto por ser rechazados dadas las limitaciones de la oferta disponible en las universidades. Parte del problema es que los modelos escolarizados son costosos, demandan mucho personal especializado (profesores), las edificaciones son difíciles de reproducir, la demanda siempre creciente los sobrepasan, hay muestras de insuficiencia para afrontar los cambios acelerados de la actualidad. De aquí que se requiera encontrar nuevas vías que nos ofrezcan un futuro más viable, como las de distribuir la enseñanza, al involucrar de una manera más activa a los propios estudiantes en el proceso formativo, transformar las funciones del profesor, hacer más horizontales las decisiones que afectan a los estudiantes al hacerlos copartícipes.

Otro problema es la calidad del proceso educativo, ya que de manera predominante sigue bajo un modelo tradicional centrado en un profesor que instruye, donde la atención es convergente en lo que expone, no se promueven actitudes emprendedoras en los estudiantes, quienes están a la espera de recibir instrucciones, con una actitud pasiva, reactiva sólo para atender las indicaciones recibidas, no se enseña a ser proactivo, a tener *autorregulación*, desarrollar iniciativas, creatividad, originalidad. De aquí que se considere necesario orientarse a la búsqueda de modelos de *enseñanza divergentes* que se centren en el alumno, que promuevan sus iniciativas, así como su pensamiento crítico y capacidades creativas.

El tercer problema es promover la formación ética de los estudiantes, modelar la convivencia cívica, desarrollar valores de ética social a partir del respeto por el otro, del respeto al principio de la reciprocidad, de la responsabilidad ante compromisos compartidos y asumidos, de la colaboración para la elaboración de proyectos comunes, de la *co-evaluación* responsable del desempeño académico de los propios compañeros a partir de la formación de juicios críticos, ponderados, imparciales y bien sustentados.

De aquí que este trabajo se concibe como una contribución en busca de una nueva concepción del proceso educativo, al ubicar en el centro del proceso de enseñanza a los propios estudiantes, con base en tres planteamientos: la *enseñanza distribuida*, la *enseñanza divergente* y la *co-evaluación*, integrados a partir de una estrategia *mixta* de enseñanza *presencial* y *virtual*. Los resultados de esta investigación apuntan a que la *co-evaluación* realizada por los alumnos fue hecha con responsabilidad, se otorgaron calificaciones promedio que no son altas ni bajas, se aprecia que son consistentes y por lo mismo confiables.

La valoración que realizaron los tres profesores del curso sobre los argumentos expresados por los alumnos para justificar las evaluaciones que emitieron los estudiantes, coinciden en su mayoría, lo que implica que los criterios de las *rúbricas* fueron claros, las medias de calificación fueron muy parecidas en los tres jueces, las correlaciones resultaron bastante altas, los niveles de varianza fueron prácticamente el mismo, lo que permite concluir que cerca del 90% de los estudiantes pueden ser considerados buenos evaluadores, que tienen rigor y son confiables para evaluar una tarea compleja como lo es calificar un ensayo, lo cual implica la apreciación de procesos cognoscitivos complejos a partir de la expresión escrita.

Quedan retos pendientes en la *enseñanza distribuida*, porque no todos los alumnos tienen las mismas competencias ni se comprometen por igual, lo que genera desigualdades disruptivas. Se tienen que crear arreglos que permitan inducir y regular mejor el trabajo por colaboración, donde las responsabilidades compartidas y la distribución de tareas sean equitativas. Un gran problema se presenta cuando debido a la falta de cumplimiento ante los compromisos asumidos por parte de alguno o algunos estudiantes, se empobrece el trabajo colectivo o tiene que ser compensado con el trabajo de otro. Esto produce de manera justificada una gran inconformidad y hace difícil la evaluación de ensayos colectivos. El recurso empleado para atenuar este problema ha sido dar seguimiento en el *aula virtual* al desempeño individual y ponderar la calificación con base en estos indicadores, pero habrá que buscar nuevos caminos para atender esta problemática.

Referencias

- Ahumada, P. (2005). *Hacia una evaluación auténtica del aprendizaje*. México, Paidós.
- Anderson, L. W. y Krathwohl, D. R. (Editors); Airasian, P. W.; Cruikshank, K. A.; Mayer, R. E.; Pintrich, P. R.; Raths, J. and Wittrock, M. C. (Contributors). (2001). *A Taxonomy for Learning, Teaching, and Assessing A Revision of Bloom's: Taxonomy of Educational Objectives*. New York. Addison Wesley Longman.
- Baker, E. (2009). "Consideraciones de validez prioritaria para la evaluación formativa y de rendición de cuentas", en *Revista de Educación*, 384, pp. 91-109.
- Bartolomé, Antonio (2004). "Blended Learning. Conceptos básicos", en *Revista de Medios y Educación*, 23, pp. 7-20.
- Bloom, B. S. (ed.), M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. (1956) *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain*. New York: David McKay.
- Bearison, D. J. and Dorval, B. (2002). *Collaborative cognition*. Westport: Albex.

- Biggs, J. (2003). *Teaching for Quality Learning at University*. Open University Press. McGraw Hill.
- Castañeda, S. (2006). *Evaluación del aprendizaje en el nivel universitario*. México, UNAM.
- Castillo, S. y Cabrerizo, J. (2007). *Evaluación educativa y promoción escolar*. Madrid, Pearson Prentice Hall.
- Canales, A y Gilio, M. (2008). “La actividad docente en el nivel superior: ¿diferir el desafío?”, en Rueda. M. (coordinador) *La evaluación de los profesores como recurso para mejorar su práctica*, México, UNAM, Plaza y Valdés Editores.
- Díaz Barriga, F. (2006). *Enseñanza situada: vínculo entre la escuela y la vida*, México: McGraw Hill.
- Díaz Barriga, F y Hernández, G. (2002). *Estrategias docentes para un aprendizaje significativo; una interpretación constructivista*, México, McGraw Hill.
- Estes, Ch. (2004). “Promoting Student-Centered Learning in Experiential Education”, en *Journal of Experiential Education*, 27(2), pp. 141-161.
- García-Aretio L., Ruiz M. y Domínguez D. (2007). *De la educación a distancia a la educación virtual*, Barcelona, España, Ariel.
- Laurillard, D. (2009). “The pedagogical challenges to collaborative technologies”, en *International Journal of Computer-Supported Collaborative Learning*. 4 (1), pp. 5-20
- Ligorio, M. B. & Ritella, G. (2010). “The collaborative construction of chronotopes during computer-supported collaborative professional tasks”, en *International Journal of Computer-Supported Collaborative Learning*. 5 (4), pp. 433-452
- Mihaela Sgera F y Stoicescu D. (2011). “Using blended learning as a tool to strengthen teaching competences”, en *Procedia Computer Science*. 3, pp. 1527-1531.
- Olson, D. R. (1998). *El mundo sobre el papel. El impacto de la escritura y la lectura en la estructura del conocimiento*. Barcelona, Gedisa.
- Quesada, R. (2003). *Cómo planear la enseñanza estratégica*, México, Limusa.
- Silva, A. (2004). *Métodos Cuantitativos en Psicología, Un enfoque Metodológico*, México, Trillas.
- Stegmann, K., Weinberger, A. y Fischer, F. (2007). “Facilitating argumentative knowledge construction with computer-supported collaboration scripts”, en *International Journal of Computer-Supported Collaborative Learning*. 2 (4), pp. 421-447
- Tirado, S.F. (2007). “La formación ciudadana y la ética social”, en Vidales, I y Maggi, R. (Compiladores) *La democracia en la escuela. Un sueño posible*, México, Colegio de Estudios Científicos y Tecnológicos del Estado de Nuevo León.
- Tirado F, Bustos A. y Miranda A. (2007). “Enseñanza divergente, diferenciada y distribuida a partir de un aula virtual”, en *Memorias: IX Congreso Nacional de Investigación Educativa*. Consejo Mexicano de Investigación Educativa, Facultad de Educación, Universidad Autónoma de Yucatán, Mérida, Yucatán.
- Whittaker, T. y Bonakdarian, E. (2011). “Face-to-face experiences for online students: effective, efficient, and engaging hybrid classes”, en *The Journal of Computing Sciences in Colleges*. Volume 26 (4), pp.140-148.