



## Sistematización de la evaluación cualitativa

### Un procedimiento de valoración de la calidad de la *co-evaluación*

Felipe Tirado

Alejandro Miranda

Ana Elena Del Bosque

[ftirado@unam.mx](mailto:ftirado@unam.mx)

Universidad Nacional Autónoma de México FES Iztacala

## RESUMEN

En este estudio se reporta una investigación que se llevó a cabo con dos grupos de estudiantes de licenciatura, en la que se les invitó a evaluar las propiedades cualitativas de un ensayo realizado por sus compañeros (*co-evaluación / evaluación por pares*). El propósito del estudio fue recabar evidencia empírica sobre los niveles de consistencia y la validez de los juicios de sustento que ofrecieron los estudiantes sobre las evaluaciones que realizaron, lo cual fue valorado de manera independiente por tres profesores.

La conclusión del estudio es que los estudiantes evaluaron con rigor y sustento.

## INTRODUCCIÓN

Los sistemas escolarizados tradicionales dan claras muestras de agotamiento, su eficacia para alcanzar los objetivos curriculares es muy limitada, como se puede apreciar en los múltiples indicadores publicados por el *Instituto Nacional para la Evaluación de la Educación* (Backhoff, Andrade, Peón, Sánchez y Bouzas, 2006. Backhoff, Andrade, Sánchez y Peón, 2008. Sánchez y Andrade; 2009), resulta muy poco eficiente por los altos costos que involucra tanto la infraestructura (edificios) como el personal requerido (docentes y administrativos). La eficiencia terminal



promedio en las instituciones de educación superior es muy baja, al alrededor del 32% (Pérez, 2006) no tiene capacidad para absorber la demanda, “Actualmente, de la población de 18 a 23 años de edad, en términos de cobertura neta, sólo el 18% está incorporado al sistema de educación superior” (SEP, 2007), tan sólo la *UNAM* rechaza a más de 150 mil aspirantes cada año.

El modelo escolarizado sigue anclado en una concepción decimonónica donde un grupo de alumnos se concentra en un salón de clases dotado de bancos y un pizarrón, cuya actividad principal se concentra en el discurso, a veces apoyado por notaciones, lo que corresponde a un modelo convergente donde la atención se centra en quien tiene el uso de la palabra, que las más de las veces resulta ser el profesor. Pareciera que lo importante es qué hace el profesor y no los alumnos.

Hay un amplio consenso en que la educación debe ser formativa y no informativa, que lo relevante es desarrollar en cada alumno las competencias que le permitan lograr una vida de calidad y plenitud, para lo cual se requiere formar personas con capacidad de reflexión, que sean críticas, creativas y emprendedoras.

Para lograr estos propósitos se requiere promover modelos de aprendizaje activo y proactivo, no pasivo y receptivo como son los modelos instruccionales dominantes, hay que pasar del dominio del aula y el profesor a la actividad del estudiante. Se necesita sistematizar modelos divergentes que permitan a cada estudiante desarrollar sus propias ideas (Tirado, Bustos y Miranda, 2007), favorecer procesos constructivos (Estes, 2004), que estén centrados en la actividad de los estudiantes (Biggs, 2003).

La evaluación puede ser un instrumento que puede regular los procesos formativos, porque es un recurso que permite no sólo la función tradicional de valorar los logros escolares alcanzados, sino también puede modelar el proceso en tanto lo que se evalúa toma el carácter de importante, por lo mismo se atiende y procura cumplir con ello.

## PLANTEAMIENTO

El planteamiento de este estudio es promover la formación de estudiantes universitarios a partir de su participación en tareas complejas, como son la



elaboración de ensayos colectivos. Se plantea a la elaboración de un ensayo por colaboración como una tarea compleja que implica múltiples componentes, como la definición de un tema a tratar, valerse de la localización de la literatura especializada y actualizada de mayor relevancia e impacto del tema a tratar, para así poderse ubicar en los conocimientos de frontera, lo que permite conocer las tesis y contrastarlas de vanguardia que sustentan los autores de mayor relevancia, de aquí reflexionar en relación con las ideas encontradas para discernir críticamente y elaborar un pensamiento con reflexiones propias, hacer la exposición del entramado de ideas por escrito, con claridad, para llegar a una conclusión en la que de manera sintética se destacan los planteamientos más relevantes, las ideas originales y los juicios propios.

Un ensayo se realiza por medio de la escritura, que es el mejor instrumento del pensamiento, en tanto es un medio ideal para precisar y expresar las ideas (Olson, 1998), dado que genera registros, lo cual posibilita revisar y reflexionar sobre lo ya escrito, permite que el texto sea ajustado tantas veces se desee hasta llegar a la plena satisfacción de quien lo escribe, pudiendo lograr de este modo su mejor expresión.

Para que un estudiante gane competencias en la escritura de ensayos requiere ejercitar esta tarea, y comprender con puntualidad cuáles son los elementos constitutivos de un buen ensayo, lo cual se puede lograr a partir de elaborar ensayos y realizar ejercicios de evaluación sistematizada.

La evaluación de ensayos es una tarea compleja, porque constituyen estructuras de ideas interrelacionadas en las que se derivan afirmaciones y conclusiones bajo determinados fundamentos, lo cual tiene que ser valorado con base en juicios de apreciación, en los que necesariamente hay una visión subjetiva de lo que se considera “debe de ser”, lo que se estima que “está bien hecho”. Aquí no hay una métrica, una posible medición de magnitudes que permita establecer puntajes objetivos.

Lo que sí puede haber es juicios de opiniones afines o disidentes en torno a si está bien o mal hecho un ensayo considerando diferentes aspectos, y en este sentido las apreciaciones pueden ser concurrentes, iguales, parecidas o diferentes, lo que puede



ser valorado sistemáticamente de manera puntual y objetiva. De aquí que el planteamiento sea colegiar las evaluaciones (constituir jurados), de manera que se obtenga una pluralidad de apreciaciones acotadas, atendiendo determinados aspectos (*rubros*), definidos bajo ciertos criterios (*rubricas*), y dado cierto parámetro (*escala*), de tal forma que haya una correspondencia en cuanto a los aspectos del ensayo que se evalúan, los criterios con los que se juzga y la manera en cómo se expresa la apreciación, lo que permite acotar la subjetividad. Además se puede enriquecer estas evaluaciones al solicitar que se refiera cuáles fueron las razones que llevó a otorgar la apreciación expresada en la valoración. Bajo este procedimiento se plantea que es posible reconocer y diferenciar la calidad de un ensayo, dentro de cierto rango de precisión.

En el planteamiento de este estudio se considera que los estudiantes pueden integrar los jurados para evaluar los ensayos realizados por sus propios compañeros, lo que aquí se refiere como la *co-evaluación*. La *co-evaluación* puede ofrecer múltiples ventajas, la primera es contribuir en la formación de los estudiantes, al requerirles la lectura activa y cuidadosa de un ensayo elaborado por sus compañeros, en el que se solicita se forme un juicio de qué tan bien está elaborado revisando diferentes aspectos (*rubros*), siguiendo ciertos criterios establecidos en las *rubricas*, con apego a una *escala* y expresando las razones en que sustenta su juicio.

Para enriquecer la evaluación colegiada de ensayos conviene que los evaluadores estén familiarizados con el tema, de manera que puedan reconocer si la bibliografía es suficiente, relevante y actualizada, si se tocan los puntos más relevantes del tema, si hay aportaciones originales, etcétera. Para esto los alumnos trabajan elaborando su propio ensayo a lo largo del semestre sobre el mismo tema del ensayo que van a evaluar al final de curso.

Los estudiantes son una fuente de opiniones muy rica para colegiar la evaluación de tareas complejas, además el hacer los procesos de evaluación más horizontales, donde todos los participantes se involucren, crea condiciones de convivencia más participativa en la que las decisiones tomadas las más de las veces de manera unilateral por los docentes queden acotadas, para eliminar o atenuar de este modo las arbitrariedades que a veces comenten algunos profesores.



### *Hipótesis de trabajo*

Los estudiantes pueden ser muy buenos evaluadores para juzgar propiedades cualitativas de una tarea compleja, si se establece para ello un procedimiento sistemático en el que los alumnos se preparen en el tema a evaluar, se tengan los aspectos a evaluar bien puntualizados (*rubros*), con criterios bien definidos (*rubricas*) y con gradientes específicos (*estandarizados*), en el que se crucen las evaluaciones (*co-evaluación*) y prevalezca el anonimato.

### *Procedimiento*

Se utilizó un *aula virtual en línea* que operó bajo una plataforma *Moodle*, lo cual ofrece un medio muy operativo, en tanto facilita todos los arreglos requeridos para llevar a cabo el diseño de operación requerido y se crean registros automatizados en bases de datos de todas las actividades desplegadas dentro del *aula*.

Se forman equipos de trabajo para hacer ensayos por colaboración durante todo el semestre escolar, a partir de la selección de un tema curricular que sea de la preferencia del alumno. Para esto se solicita que del temario asignen un número ordinal de acuerdo con sus preferencias, y se le otorga el tema más cercano a su primera elección, siempre y cuando esta elección permita la formación de los grupos requeridos para que se pueda desarrollar un proceso de co-evaluación a partir de un diseño en que se pueda circular la evaluación. Por ejemplo, el grupo A califica al grupo B, el B al C y el C al A, de manera que nunca se evalúan los mismos grupos.

Para esto se establecen los aspectos a evaluar, en nuestro caso siete *rubros*, se definen una serie de *rubricas* que corresponden a los criterios y un gradiente *estandarizado* (0 – muy mal, 1- mal, 2 – bien, 3 – muy bien). Las evaluaciones se hacen preservando el anonimato doble (*doble ciego*), es decir no se sabe a quién se está evaluando ni se sabe quien evaluó, de manera tal que se promueva la imparcialidad, evitando que se obsequien calificaciones altas en busca de complicidad o se castigue con notas bajas por posibles antipatías o rivalidades existentes.



7, 8 y 9 de Octubre de 2010

En la primera fase de nuestro estudio, cada alumno evalúa de manera independiente uno de los ensayos, del mismo tema que abordó, siguiendo los criterios especificados en las rubricas, con el gradiente convenido, y sustentando el juicio cualitativo de apreciación que le llevó a la calificación que otorgó.

En la segunda fase, con el propósito de valorar la calidad de la co-evaluación realizada por los estudiantes, los profesores del curso (tres) juzgan qué tan bien están realizadas las evaluaciones, esto a partir de los juicios de sustento que ofrecieron cada alumno en cada *rubro* para fundamentar su valoración. Se trata de una meta evaluación, en la que se realiza la evaluación de la evaluación.

### *Instrumento*

Los alumnos calificaron electrónicamente en un *aula virtual* a partir de las *rubricas* que consideraban los siete rubros en función de las propiedades a evaluar, los cuales son: 1- la **Organización** (El orden o estructura de la exposición de las ideas es adecuado); 2- la **Claridad** (Está bien escrito, se comprende con facilidad el texto); 3- la **Suficiencia** (Se tocaron los puntos más relevantes del tema); 4- el **Tratamiento del caso** (Se abordó y analizó apropiadamente un caso práctico desde el planteamiento del marco conceptual del trabajo); 5- la **Contribución** (Hay aportaciones originales de los autores); 6- las **Conclusiones** (Hay una buen síntesis y se resaltan las aportaciones del trabajo apoyadas en el caso) y 7- la **Bibliografía** (Las referencias bibliográficas son pertinentes al tema y actualizadas). La apreciación se dio con base en una escala tipo *Likert* donde se expresa el grado de acuerdo que se tiene con la afirmación que se expone en el criterio de cada *rubro* y se solicita escribir la razón en que basa su 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: \_\_\_\_\_



## ANÁLISIS DE RESULTADOS

### *Sujeto*

Se trabajó con dos grupos escolares en los que estaban inscritos 63 alumnos que cursaban el quinto semestre de la licenciatura en psicología, de los cuales 12 desertaron. Con los 51 alumnos restantes se formaron 14 grupos de trabajo para la elaboración de los ensayos temáticos por colaboración.

Para hacer una valoración sistemática de las evaluaciones, se procedió a estimar el grado de consistencia interna y congruencia entre las calificaciones que otorgan los alumnos, a partir de los niveles de acuerdo, las correlaciones entre las calificaciones, la varianza, y los análisis de varianza de las distribuciones con referencia a las medias de calificación otorgadas.

### *Primera fase*

La calificación promedio que los alumnos otorgaron a los ensayos fue de 49.9 en una escala del 0 al 100, es decir, que su calificación no fue alta ni baja. Este resultado está en contra del prejuicio que sostiene que los alumnos irresponsablemente se regalaran calificaciones altas o por el contrario las utilizarán para castigar.

Hubo una varianza promedio de 0.43, lo que indica que no se otorgaron calificaciones homogéneas, sino diferenciadas, es decir, hubo discriminación de la calidad del ensayo en sus diferentes rubros, es decir, no fue siempre la misma valoración: muy mal o muy bien.

Para valorar la confiabilidad, se contabilizó de las 322 calificaciones otorgadas, en cuántas hubo coincidencia, encontrando que en el 63% de las ocasiones (203) hubo correspondencia. Esto no quiere decir que en el restante 37% hubiera un total desacuerdo, ya que unos podrían señalar muy mal y otros simplemente mal, y aunque no hay correspondencia tampoco hay un completo desacuerdo.

También se corrió el análisis de confiabilidad (Alpha de Cronbach) en que se obtuvo un valor de 0.668, que es un valor aceptable, sobre todo porque se debe



tener en cuenta que en este análisis se trata de 14 ensayos y por lo mismo de 14 grupos, por lo cual no está mal este valor de consistencia.

Dado que son 14 grupos, se corrieron las correlaciones de las calificaciones de cada alumno con las del promedio de su propio grupo. Sólo en el 29% se obtuvieron correlaciones bajas (menores a 0.50), por lo que se puede indicar que hubo un buen nivel de correspondencia en el 70% de las apreciaciones de los estudiantes.

### *Segunda fase*

Los tres profesores participantes valoraron los 336 juicios que dieron los alumnos para sustentar cada una de sus calificaciones. Los profesores coinciden entre ellos en el 76% (256 veces) de las ocasiones, sólo en 80 caso hubo discrepancias, pero nuevamente esto no quiere decir que hubiera un total desacuerdo, unos podrían marcar bien y otros muy bien, que aunque no son iguales sí son similares.

Al correlacionar las calificaciones promedio de los tres profesores con las de cada uno de ellos, se obtuvieron correlaciones muy altas, el promedio fue de 0.76. Sólo en un profesor, un caso (en el rubro de *Claridad*), obtuvo una correlación por debajo de 0.50 (0.38), lo que indica que la apreciación de los profesores fue muy similar.

La varianza de calificaciones de los tres profesores fue de 0.51, 0.50 y 0.51, es decir que la variación promedio de sus calificaciones fue prácticamente la misma. También se realizó un análisis de varianza de las calificaciones otorgadas por cada profesor para apreciar si había diferencias estadísticamente significativas, lo cual se estimó corriendo un análisis de varianza de un factor, con comparaciones múltiples post hoc, Scheffe, con un nivel de significación 0.05, siguiendo el procedimiento que señala Silva (2004). Sólo en 1 de los 7 rubros el profesor **A** no obtuvo diferencias significativas con el profesor **C**, con el profesor **B** tuvo diferencias en los 7 rubros. El profesor **B** y **C** no tuvieron diferencias significativas en 6 de los 7 rubros, sólo en uno la diferencia entre las medias fue significativa. Se puede decir que si no hay diferencias significativas entre las apreciaciones de un profesor y otro, significa que daría lo mismo quién de ellos califica, en tanto llegan a la misma valoración.

De la valoración de los juicios ofrecidos por los estudiantes ninguno de los estudiantes fue considerado por los profesores como “muy mal” evaluador, 10% (5)





fueron valorados como “malos” y el restante 90% como “buenos” (82%) o “muy buenos” (8%).

### Opiniones

Al finalizar el curso se les aplicó a los estudiantes un cuestionario de 8 preguntas para valorar sus opiniones en torno al ejercicio de *co-evaluación* realizada. En términos generales el 72% manifestó que no consideraban que sus compañeros calificaban con favoritismo, sólo un 12% cree que hay mala fe en sus compañeros a la hora de evaluar. La mayoría de los estudiantes (64%) considera conveniente que se realice la *co-evaluación*, aunque el 88% está de acuerdo en que se haga supervisada por el profesor.

## CONCLUSIÓN

Con base en los resultados encontrados se puede contestar afirmativamente la hipótesis de trabajo, señalando que si se siguen procedimientos bien sistematizados, en los que los alumnos tengan la oportunidad de prepararse en el tema que van a evaluar, se les proporcionan los *rubros* de lo que deben evaluar, se ofrecen *rúbricas* con escalas *estandarizadas*, se hacen evaluaciones cruzadas preservando el anonimato, entonces es posible obtener evaluaciones bien ponderadas que valoren las propiedades cualitativas de una tarea compleja, como son la elaboración de un ensayo.

Si los estudiantes son conminados a realizar un ensayo, se podrá tener un modelo divergente de educación proactiva, en la que cada uno de ellos tenga oportunidad de construir sus propias ideas a partir de una visión crítica y creativa. Integrar a los estudiantes a los proceso de evaluación, permite hacer más horizontal la formación, permitiendo que el alumno asuma una responsabilidad en su proceso formativo que le permita reflexionar sobre el mismo y con esto caminar hacia modelos de responsabilidad y enseñanza distribuida, en los que se transformen las funciones de las labores docentes, ofreciendo de este modo la oportunidad de una nueva gestión educativa que permita ampliar la matrícula y así poder atender mejor a la gran demanda que hoy deja a cientos de miles de jóvenes sin ejercer su derecho a la educación.



## REFERENCIAS

- Backhoff, E., Andrade, E., Peón, M., Sánchez, A. y Bouzas, A. (2006). *El aprendizaje del Español, las Matemáticas y la Expresión Escrita en la educación básica en México: sexto de primaria y tercero de secundaria*. México, D.F.: Instituto Nacional para la Evaluación de la Educación.
- Backhoff, E., Andrade, E., Sánchez, A. y Peón, M. (2008). *El aprendizaje en tercero de preescolar en México: Lenguaje y comunicación y Pensamiento matemático*. México, D.F.: Instituto Nacional para la Evaluación de la Educación.
- Biggs, J. (2003). *Teaching for Quality Learning al University*. Open University Press. McGraw Hill.
- Estes, Ch. (2004). Promoting Student-Centered Learning in Experiential Education. *Journal of Experiential Education*, 27(2), pp. 141-161.
- Olson, D. R. (1998). *El mundo sobre el papel. El impacto de la escritura y la lectura en la estructura del conocimiento*. Barcelona: Gedisa.
- Pérez, J. (2006). La eficiencia terminal en programas de licenciatura y su relación con la calidad educativa. En *Revista Electrónica Iberoamericana sobre Calidad, Eficiencia y Cambio en Educación*. Vol.4, No 1, pp. 130-148. <http://redalyc.uaemex.mx/pdf/551/55140110.pdf>
- SEP (2007). *Plan Nacional de Educación 2007 -2012*. México: Secretaría de Educación Pública.
- Sánchez, A. y Andrade, E. (2009). *El aprendizaje en tercero de secundaria en México. Informe sobre los resultados de Excale-09, aplicados en 2008*. México, D.F.: Instituto Nacional para la Evaluación de la Educación.
- Silva, A. (2004). *Métodos Cuantitativos en Psicología, Un enfoque Metodológico*. México: Trillas.
- Tirado F., Bustos A. y Miranda A. (2007). Enseñanza divergente, diferenciada y distribuida a partir de un aula virtual. Memorias: *IX Congreso Nacional de Investigación Educativa*. Consejo Mexicano de Investigación Educativa, Facultad de Educación, Universidad Autónoma de Yucatán, Mérida, Yucatán