

## **EL JUICIO DEL ORÁCULO. LA RED COMO RECURSO PARA LA MEJORA DE LA EVALUACIÓN EDUCATIVA.**

**Daniel David Martínez Romera**  
*Universidad de Málaga*

### **Resumen**

La evaluación puede llegar a ser fuente de dificultades para los docentes preocupados por aplicarla de forma útil, y un condicionante externo mal entendido por parte del alumnado. En este sentido, el calado social de Internet ha transformado no sólo nuestra forma de ver el mundo, sino también cómo aprendemos y comprobamos la veracidad de los hechos. Por ello, utilizarla de forma eficiente parece una opción más interesante que oponérsele esgrimiendo la bandera de sus efectos negativos, muy reales por otro lado: debilitamiento del pensamiento crítico, exaltación de los resultados en detrimento de los procesos, trivialización de la importancia del saber propio... Y es que también nos permite refinar procesos diagnósticos como la evaluación de pares de una forma antes impensable para contextos distintos al pequeño grupo.

Sin embargo, y a diferencia de lo que se argumenta tradicionalmente, este tipo de evaluación es común a casi todas las etapas de la vida: cuando se trabaja o colabora en grupo se construyen juicios basados en el desempeño individual y colectivo de los pares, independientemente de si estamos en educación infantil, primaria, secundaria, universitaria, un grupo de investigación o de técnicos. Como ejemplo de su aceptación y utilidad han quedado, en el ámbito académico, estándares como el de la evaluación de artículos por doble ciego. Para ilustrar el papel que puede llegar a tener la Web en estos procesos, así como los paralelismos trazables entre la teoría de redes como instrumento para estudiar Internet y las relaciones interpersonales, se utilizarán de base experiencias llevadas a cabo en el marco del Proyecto de Innovación Educativa (PIE13-171) – La evaluación de 360º para evaluar las competencias profesionales en la enseñanza universitaria, y también durante el presente curso académico, en la Universidad de Málaga.

**Palabras clave:** Educación, Innovación, Evaluación, Web, Didáctica, Ciencias Sociales.

## 1. Introducción

El interés por aplicar nuevas metodologías en el ámbito educativo superior lleva manifestándose como un punto de investigación interesante en los últimos años (Pérez, E. B., Gómez-Linares, A., y Martín, R. P., 2015; Panadero-Calderón, E. y Alonso-Tapia, J., 2013); algo lógico por otra parte, ya que la naturaleza de las relaciones sociales actuales está cada vez más mediatizada por la tecnología (Casero-Ripollés, 2012; Castells, 2012 y 2011) en la ya conocida como sociedad del conocimiento. Por ello, los docentes se encuentran en una encrucijada educativa, entre los modos y prácticas válidos hasta ayer y la vorágine de interdependencias y conectividades de hoy, entre concepciones centradas en el conocimiento como eje curricular y concepciones articuladas en las destrezas de aprendizaje por otro; lo que, al margen de las sempiternas instrumentalizaciones políticas, sigue siendo motivo de debate educativo (Casanova, 2012).

La evaluación de 360° es un instrumento de interés plenamente vigente y extendido (Arzate, 2013; Cano, 2008; Alzina, Olmo, Soler y Escoda, 2006) que permite desarrollar un proceso de evaluación desde todas las ópticas implicadas en el aula: tanto alumnado como individuo y conjunto como profesorado implicado. Con ello se pretende desarrollar una aproximación integrada al proceso de aprendizaje y su mejora, objeto en última instancia de la evaluación, en el que la perfectibilidad de los aprendizajes se produce tanto por motivaciones intrínsecas como valoraciones extrínsecas, de modo que el autoconcepto se desarrolle de forma positiva y no ilusoria (Fernández, Ramos-Díaz, Zuazagoitia, 2016), fomentando así actitudes asertivas y resilientes que permitan crear espacios de trabajo en equipo constructivos.

El Proyecto de Innovación Educativa que ha constituido el soporte de la presente investigación fue aplicado, por el conjunto del profesorado implicado, en varias asignaturas de grado y posgrado de las Facultades de Educación, Ciencias, Informática y Periodismo de la Universidad de Málaga. Además, la naturaleza de las mismas, claramente heterogéneas dada su procedencia, se vio acentuada por singularidades como el curso académico en el que se imparte y su relevancia curricular (troncal, obligatoria, optativa). La aplicación de un instrumento de evaluación en dicho contexto supone, de facto, un reto. Sin embargo, también subraya su capacidad para ser aplicada en una gran variedad de contextos educativos, aspecto que la hace especialmente interesante.

Dentro de este conjunto de experiencias, el desarrollo de la presente propuesta de innovación educativa se realizó en el grupo C de la asignatura Educación para la Ciudadanía y los Derechos Humanos, formada por 78 estudiantes, perteneciente al Grado de Educación Primaria de la Facultad de Educación, durante el curso 2014/15. Además, y con objeto de establecer una referencia comparativa, se realizó un modesto ejercicio análogo en el Máster de Profesorado y Doble Máster de Profesorado, Filosofía y Ciudadanía dentro de la asignatura de Innovación Docente e Iniciación a la Investigación Educativa de la especialidad de Ciencias Sociales, formada por 42 estudiantes, a partir de una única tarea de evaluación entre pares con valoración externa del docente, durante el curso anterior, 2013/14.

## **2. Método**

El objetivo del presente estudio fue comprobar la utilidad de los recursos tecnológicos en el proceso de evaluación del alumnado durante la realización de trabajos grupales, para lo que se utilizó como instrumento tecnológico la rúbrica digital (*rubric*, <https://gteavirtual.org/rubric/>) desarrollada por el Grupo de Investigación en Globalización, tecnología, educación y aprendizaje de la Universidad de Málaga.

El trabajo principal, sobre la asignatura de Educación para la Ciudadanía y los Derechos Humanos, analizó longitudinalmente uno de los 14 grupos, de 5 miembros, en que se organizó la clase para realizar las distintas tareas presenciales y no presenciales. La selección del mismo se planteó de forma transparente bajo dos premisas: su carácter voluntario y la no repercusión en la evaluación final, lo que podría condicionar claramente su actitud y dedicación, enmascarando así el verdadero comportamiento del método. Varios fueron los grupos que se ofrecieron a participar, siendo el elegido el que presentaba un mayor número de miembros sin colaboración previa en anteriores cursos, con lo que se pretendía evitar la vacuidad de evaluar grupos consolidados de trabajo eficientes.

El carácter longitudinal del estudio quedó garantizado con el establecimiento de tres puntos de control sobre otras tantas tareas, en diferentes momentos de la asignatura:

- 2.1. Inicio: Construcción de un modelo democrático. Tarea grupal realizada al inicio de la asignatura, y que pretendía acercar al alumnado a la complejidad de la democracia contemporánea a partir de un texto de Robert Dahl (2004). El mismo se ofreció con antelación para su lectura individual, de modo que los miembros pudieran tener un debate basado en posturas meditadas, que facilitaran la

exigencia de una respuesta redactada colectiva a las cinco preguntas fundamentales que formula el autor sobre el tema.

- 2.2. Ecuador: Evaluación de pares mediante rúbricas. Tarea realizada en el Ecuador de la asignatura, que tenía por finalidad acercar al alumnado a la dificultad de la toma de decisiones y la emisión de juicios sobre los demás. Cada grupo debía establecer una rúbrica de evaluación para la exposición oral de los trabajos finales de la asignatura, en la que se especificara: criterios, medición y ponderación. Tras el debate final en gran grupo, la matriz resultante sería la que cada alumno utilizaría para evaluar las exposiciones de sus compañeros.
- 2.3. Final: Trabajo Final de Grupo. Aunque finalizada con la asignatura, se trata de una tarea de largo recorrido. Desde el punto de vista de la investigación metodológica, es la que ha permitido conocer con mayor detalle la evolución de las relaciones internas y el compromiso individual de cada uno de los miembros, ya que es la única tarea que se abordó de forma recurrente.

Tras la finalización y entrega de cada tarea, se realizó una reunión de valoración global a dos niveles: individual y grupal, basada en la realización y análisis de las tres rúbricas on-line que articulaban la propuesta (autoevaluación, heteroevaluación y evaluación del profesor), véase Tabla I. Por tanto, las rúbricas debían ser cumplimentadas dos veces en cada tarea por parte del alumnado (autoevaluación y evaluación de pares) y una por parte del profesor, lo que en conjunto supuso la utilización de 9 rúbricas digitales. En este sentido, es importante subrayar el carácter anónimo de las evaluaciones de pares, en tanto que los miembros evaluados podían ver su calificación media pero no la estimación individual realizada por sus pares. A su vez, la observación de la docente, necesaria para emitir su juicio, se basó en la observación del trabajo de clase en las sesiones presenciales correspondientes.

**Tabla I. Transcripción del modelo de rúbrica digital utilizado.**

<b>Ítems (Ponderación)</b>	<b>Puntuación Media</b>
(25.00%) Coordina al grupo, organiza y distribuye tareas, toma iniciativas, resuelve problemas	%
(15.00%) Colabora y ayuda a sus compañeros, comprometido con la tarea	%
(10.00%) Acepta, negocia e integra las diversas ideas. Genera confianza en el grupo	%
(10.00%) Presenta una actitud pasiva en el desarrollo de la actividad, descansa sobre los demás	%
(15.00%) Expresa su opinión más que escuchar a los demás	%
(25.00%) Hostil, agresivo, plantea problemas, disruptivo, entorpece el desarrollo de la actividad	%
<i>(100.00%) Trabajo en equipo (autoevaluación/heteroevaluación/profesor)</i>	%

La rúbrica digital se centró en la evaluación de una única competencia, el trabajo en equipo, con objeto de medir un aspecto concreto del proceso de aprendizaje. Su elección se basó en la relevancia que tiene dicha competencia en la formación del profesorado, el desarrollo de la sociabilización del alumnado y su utilidad en los entornos laborales (Poy-Castro, Mendaña-Cuervo, González, 2015; Vadillo, 2013; van-der Hofstadt Roman y Gras, 2013). En la Ilustración 1 se recoge la estructura inicial de la eRúbrica para el docente, con la lista de miembros a evaluar, el número de evaluaciones realizadas y la calificación media recibida, pulsando sobre cada miembro se accede a la eRúbrica de evaluación sobre dicho participante, de forma análoga a la descrita en la Tabla I. Esta estructura es semejante para los dos modelos restantes, cambiando únicamente la lista de participantes a evaluar (1 en el caso de la autoevaluación o x-1 en el caso de la evaluación de pares).

The screenshot shows the 'eRúbrica' web application interface. At the top, there is a navigation bar with the logo and the text 'Herramienta para la evaluación por competencias'. Below this, the user's personal area is identified as 'Área personal de Daniel David Martínez Romera'. The main content is a table with the following columns: 'Nombre', 'Grupo', 'Mi Evaluación', 'Aprueba por mi Evaluación', 'Nº Evaluadores', 'Puntuación media de todos los evaluadores', and 'Aprueba por todos los Evaluadores'. The table contains several rows of data, with the last row showing a score of 53.00% and a green checkmark.

Nombre	Grupo	Mi Evaluación	Aprueba por mi Evaluación	Nº Evaluadores	Puntuación media de todos los evaluadores	Aprueba por todos los Evaluadores
Daniel David Martínez Romera		-	-	-	-	-
[Redacted]		-	-	1	44.00%	✘
[Redacted]		-	-	1	44.00%	✘
[Redacted]		-	-	1	56.00%	✔
[Redacted]		-	-	1	50.00%	✔
[Redacted]		-	-	1	53.00%	✔

**Ilustración 3. Captura de pantalla de la eRúbrica (docente)**

Al efecto se definieron 6 ítems con diferentes pesos en función de su relevancia, más un ítem sintético que resumiera cuantitativamente el significado de la rúbrica. La aplicación de la misma dependía del objeto de evaluación (uno mismo, los compañeros o la equivalente realizada por el profesor), para facilitar esto se definieron 3 rúbricas digitales en cada tarea con sus correspondientes enlaces URL. Así, en cada caso se accedía a una eRúbrica semejante, pero con una intención distinta:

- De autoevaluación: aquí cada alumna accedía para evaluarse únicamente a sí misma.
- De heteroevaluación (o pares): aquí cada alumna evaluaba, individualmente, al resto de compañeras, excluyéndose ella misma.
- Docente: sin participación del alumnado. Se evaluaba individualmente la competencia de trabajo en equipo de cada miembro, obteniéndose al final del proceso la valoración docente de grupo.

La experiencia muestra, no obstante, una clara debilidad, la dificultad de contextualizar los posibles resultados ante la ausencia de grupos de control comparables, ya que la naturaleza de las experiencias metodológicas realizadas por parte del resto del profesorado implicado en el PIE se llevaron en ramas de conocimiento con trasfondos diferentes (Física Aplicada, o Educación Primaria, por ejemplo). Con todo se intentaron cotejar las experiencias para discernir posibles elementos comunes a investigar, discrepancias o aspectos ambiguos en los que seguir profundizando.

Afortunadamente, en el curso anterior se pudo llevar a cabo un primer ensayo basado en la evaluación de 360° a través de la plataforma disponible en el Máster de Profesorado, y que se utilizó para estructurar una única tarea de forma semejante, aunque no idéntica, a las anteriores; ya que en última instancia se trataba de una primera tentativa de evaluación del método. Entendemos que pese a la diferente trayectoria pedagógica de ambos colectivos, la idea evaluación entre pares como experiencia de colectivo docente es tan válida como necesaria en el Profesorado de Primaria y Secundaria, motivo por el que se la incluye, con las debidas reservas, aquí.

El grupo de clase de Innovación, formado por 42 personas debía realizar una lectura individual de dos textos de opinión: la conocida disputa entre Benejam Arguimbau, 2007 y Rodríguez Adrados, 2007, sobre formación del profesorado. Y entregar, tras ello, un texto en el que se pudiera evaluar claramente tanto elementos menos subjetivos, como la identificación de los elementos clave de la confrontación mediante un resumen, como la valoración de las posturas y el posicionamiento al respecto, claramente subjetivos. Con ello se pretendía poner sobre la mesa la dificultad de evaluar en Ciencias Sociales, ya que los educadores deben ayudar a construir pensamiento crítico autónomo, y no semejante algo a lo que nuestras propias convicciones nos hacen ciegos con frecuencia. El cotejo, como se ilustrará más adelante, arrojó algunos elementos interesantes para la reflexión.

### **3. Resultados**

El seguimiento presentó una primera incidencia con la baja en la asignatura de uno de los componentes, por motivos ajenos a la dinámica objeto de estudio. No obstante se optó por mantener el grupo asignado, como mejor candidato de los posibles.

El *análisis de la primera tarea*, tanto en las entrevistas individuales como en el intercambio grupal, dejó patente el recelo con el que se veía la innovación metodológica, que podría resumirse en una suerte de silogismo: debido a una clara reticencia basada en su percepción como un instrumento de delación, ninguno de los miembros mostró el mínimo afán de protagonismo ni discrepancia, probablemente en un intento por demostrar cohesión de grupo, lo que se reflejó en resultados de rúbricas bastante homogéneas durante la primera parte de la investigación.

No obstante, la percepción que poco a poco íbamos construyendo del grupo basada en nuestra observación arrojaba algunos matices:

- Solicitud de un sistema de recompensa extrínseca, pese a que explícitamente se indicó lo contrario al inicio de la asignatura. En un intento por favorecer su rendimiento académico a través de la participación en la experiencia.
- Defectos en el diseño de la rúbrica. Tanto en las entrevistas privadas como en el debate colectivo surgieron algunas puntualizaciones agrupables en dos categorías:
  - Construcción grosera, en cuanto a la secuencia de preguntas. Subrayaron, de forma unánime, la semántica utilizada como excesivamente compleja o ambiciosa, demasiados verbos concatenados, en algunas de las preguntas formuladas, lo que dificultó la valoración.
  - Ausencia de elementos cualitativos. Probablemente sea una cuestión característica del colectivo implicado en la experiencia, ya que en las facultades de educación los aspectos cualitativos suelen tener una posición central en el seguimiento de los procesos de enseñanza-aprendizaje. La propuesta ofrecida basada en un sistema de medición tipo Likert dio pie a un intenso debate, que a nuestros ojos puso en evidencia cómo método y sujeto se condicionan mutuamente, en este caso debido a la formación de partida del segundo.

Nuestra observación había puesto de relieve algunos perfiles-tipo dentro del grupo, pero en esta fase aún no se habían manifestado como posible fuente de conflicto: claramente al menos dos personas del grupo poseían una personalidad fuerte y una tercera exhibía una actitud pasiva durante el tiempo de trabajo, la cuarta persona se ubicaba en una posición intermedia, con una virtud que a la larga demostró ser fundamental, la asertividad.

Si la autoevaluación mostró escasa utilidad en este momento, la heteroevaluación ya permitía observar elementos que corroboraban la visión docente que íbamos construyendo, especialmente sobre el posible advenimiento de conflictos: los propios miembros del grupo iban detectando quiénes aglutinaban la mayor parte del tiempo en sus intervenciones y además no eran permeables a las aportaciones ajenas. El resultado final de la primera evaluación entre pares fue un claro reflejo de ello, con dos personas levemente por debajo del umbral de valoración del 50% y otras dos de forma análoga por encima de la misma marca.

El *análisis de la segunda tarea* añadía a las dificultades de consenso el hecho de estar construyendo un instrumento de medición entre pares. Por tanto, no manifestar la postura personal tendría, en este caso, una repercusión posterior inexistente en la tarea anterior. Y entendemos que ese añadido, junto con el devenir de las sesiones, ayudó a que se hicieran patentes tanto las fortalezas como las debilidades del grupo de trabajo.

Por primera vez las entrevistas individuales señalaban, claramente, problemas de comunicación interna que podrían resumirse en una reorganización de roles que a nadie dejó satisfecho: una de las personas con más iniciativa decidió adoptar una actitud pasiva y conformista, mientras que la otra experimentó un incremento en su responsabilidad al frente del grupo, situación en la que ambas personas demostraron finalmente no encontrarse cómodas. Esta situación era muy bien descrita por el miembro asertivo del grupo, que se quejaba de un cierto agotamiento psicológico ante la situación de mediación constante en la que se encontraba inmersa. La persona pasiva del grupo reconoció su actitud, y la valoró como un mal menor ante la situación de tensión e incapacidad para alcanzar consensos que estaban caracterizando las dinámicas de trabajo. Se produjo, además, una reivindicación grupal llamativa: el de la consideración de la nota individual sobre la colectiva, en un intento por salvar la situación con el menor perjuicio posible. Mucho menos interesante se mostró la reunión grupal previa, haciéndose evidente que las formas se guardaban más como una cuestión de protocolo entre pares que como una estrategia coordinada frente a la innovación, o el docente.

De esta experiencia obtuvimos tres elementos de análisis:

- Diálogo interno no fluido. La distancia entre caracteres o bien era excesiva, o demasiado exigua. Se ofrecieron pautas para mejorar la comunicación y la escucha activa, así como se abordó la importancia de la asertividad tanto desde una perspectiva puramente ciudadana, como específicamente educativa.
- Dificultad para alcanzar acuerdos. Esto partía, a nuestro entender, de un idea equivocada de la dialéctica, o su confusión con la retórica. Algo especialmente cierto entre las dos personas dominantes del grupo, ya que su objetivo, según pudimos dilucidar en las entrevistas, era hacer ver su postura como la válida o correcta, no proponerla para debatirla, aceptar modificarla o incluso reconocer mejores alternativas a tener como referencia de trabajo, en su caso. *Verba docent, exempla trahunt*, dice el proverbio latino, pero no siempre es fácil encontrarlos,

por lo que se analizaron algunas situaciones de clase como referencia, especialmente aquellas en las que el docente aceptó, modificó o renunció una propuesta inicial suya en pos del consenso final.

La rúbrica de autoevaluación reflejó la situación de frustración que embargaba, en mayor o menor grado, a cada uno de los miembros del grupo. Siendo la puntuación media inferior al 50%, no obstante no se reconocieron actitudes causantes de conflicto, aunque sí hubo miembros que reconocieron su dificultad para escuchar, su necesidad de expresarse, o incluso de pasar desapercibidos en un intento por rehuir todo conflicto. Curiosamente, la práctica totalidad del grupo, y en contra de lo observado en las entrevistas, se consideraban solventes en cuanto aceptación, negociación e integración de diversas ideas en el grupo.

El anterior es un elemento de reflexión muy interesante en cuanto a la introducción de este tipo de innovaciones, ya que la rúbrica había permitido constatar una distancia clara entre el autoconcepto y la percepción ajena. Prueba de ello, por primera vez desde el inicio de la experiencia, la rúbrica de evaluación de pares registró la identificación de un miembro del grupo como hostil o disruptiva.

El *análisis de la tercera tarea* fue sin duda el que permitió confirmar las relaciones detectadas, pero también el que puso sobre la mesa nuestro escaso éxito, relativo en el mejor de los casos, para reconducir la dinámica de trabajo por un sendero más productivo. Dada la envergadura de la misma, y el tiempo de trabajo fuera de clase que exigió, las rúbricas de este trabajo eran por descontado las que mejor podían mostrar, ya en la recta final de la asignatura, cómo había sido el funcionamiento interno del grupo.

Lo primero que hay que reconocer al respecto es la necesidad de un trabajo más profundo por nuestra parte para hacer entender el verdadero sentido de la autoevaluación: de su revisión deducimos que las respuestas a este cuestionario se asocian más con la expresión del yo ideal que con la autocrítica, pues nadie se consideró disruptivo, nadie se consideró pasivo y las puntuaciones en cuanto a colaboración diálogo y coordinación fueron perfectas (100%). Dado el peso de la tarea en la asignatura, entendemos que son variables personales como el recelo, o la necesidad de una buena evaluación, las que explican estos resultados. Por ello, aunque no se trate de datos totalmente veraces, sí que nos pueden ayudar a reflexionar y mejorar la práctica docente, al descubrir una estrategia por parte del alumnado que nada tiene que ver con la educación, y que encuentra su origen más que

probable en prácticas docentes y sistemas de evaluación pretéritos que han alentado dichos comportamientos.

La evaluación de pares sí mostró de forma más nítida la disensión interna que se venía observando desde la tarea previa. De entrada, la valoración grupal media del trabajo en equipo se mostró insuficiente (48,86%), si bien el comportamiento de las diferentes variables entra en clara contradicción con los datos y manifestaciones recogidos en las entrevistas personales:

El trabajo global en equipo alcanza una valoración de 88,57%, a pesar de que se constataron actitudes manifiestas de pasividad y enfrentamiento que llegaron incluso a dificultar la reunión del grupo en las sesiones de trabajo no presenciales, según testimonio de todos los miembros, por supuesto, desde enfoques casi antagónicos. Por tanto, que la colaboración y la negociación sean aspectos ambos valorados con un 100% carece de sentido a nuestro juicio, si bien una acción elemental para resolver esta contradicción fue expresada de forma tácita por el alumnado en las entrevistas individuales: se sabía que los resultados, aún a grosso modo, se iban a comentar, y dado el reducido tamaño del grupo, era perfectamente identificable el comportamiento en las respuestas de sus miembros: la capacidad de inferencia jugó en nuestra contra. Presentan valores casi marginales actitud pasiva (6.67%), expresa su opinión más que escuchar a los demás (2.86%) y actitudes no deseables (2.86%).

Desde un punto de vista comparativo, llama poderosamente la atención cómo el profesorado de Primaria se mostró claramente más interesado por la metodología que el de Secundaria, éste veía en los elementos de heteroevaluación una invasión de la autonomía docente y el derecho de cátedra, manifestándose en dichos términos, estos argumentos ponen de relieve que el primer grupo concibe la educación como un proceso continuo y subjetivo por naturaleza, mientras que el segundo asumen un modelo basado en las recompensas extrínsecas puntuales en las que el docente es el referente intelectual con el que se debe medir, de modo que incluir actores secundarios no termina de entenderse como un ejercicio de aprendizaje colectivo y sí como un cuestionamiento peyorativo de su juicio.

Sin embargo, el desarrollo de la tarea deparó algo no esperado: el profesorado de Secundaria, que partía de una distancia intelectual clara respecto a la utilidad de la técnica, acabó abrazándola y argumentando sus críticas, de cara a sostener sus evaluaciones,

cuando de partida esto no parecía una meta alcanzable dadas sus motivaciones. Además, se observó una cierta convergencia de las calificaciones al alza, con una media de 8.1 para una muestra final de 36 alumnos, y con una  $\sigma$  de 1.2, en la que sólo 3 alumnos obtuvieron calificaciones de pares inferiores a 7. Encontramos, así, un elemento claro de convergencia pedagógica alcanzado no a través de la retórica magistral, sino de la práctica instrumental. Lo que de partida sólo supo ser visto por un colectivo ya habituado a la innovación educativa, acabó siendo más valorado por otro que acabó comprendiendo la importancia de poder construir conocimiento y criterios de manera colectiva.

#### **4. Conclusiones**

En general, el alumnado no mostró una conducta especialmente crítica, tanto en la autoevaluación como en la evaluación de pares, realizando puntuaciones altas en ambos casos. Depurando datos y realizando un análisis más fino encontramos algunas diferencias cuando se comparan los diferentes puntos de control establecidos. Al realizar la comparación, con otro grupo sometido a una metodología semejante, encontramos un comportamiento similar en cuanto a la generosidad de las evaluaciones, lo que nos abre interrogantes que será necesario explorar, en tanto el rango de explicaciones posibles oscila entre la estrategia discente para superar la prueba y la mejora intrínseca debida a la utilización de la nueva metodología. En todo caso, el alumnado voluntario mostró un claro interés, una vez terminada cada fase, y el propio procedimiento, para aportar sus ideas de mejora, lo que nos indica claramente el interés despertado en ellos y la utilidad que ven en este tipo de metodologías de cara a su futura docencia.

Desde un punto de vista cualitativo la evaluación ayudó a dinamizar el grupo, tanto de forma interna como con el docente como actor de apoyo y orientación, no obstante estos aspectos quedaron fuera de los elementos a medir (frecuencia y profundidad de las consultas respecto a un grupo de control). Es necesario mencionar aquí que el alumnado de la investigación principal mostró desde un comienzo sus reservas por la excesiva cuantificación de la evaluación que a su criterio suponía la metodología; y aunque tras las diversas aclaraciones éstas fueron matizándose, resulta evidente que la compensación entre los modelos cuantitativo y cualitativo de la metodología es otro aspecto a seguir investigando.

La evaluación de 360° presenta unos requerimientos de investigación que es necesario reconocer como restrictivos, ya que puede realizarse siempre y cuando el docente

disponga de un número reducido de estudiantes, pues atender a una clase a la vez que estamos concentrados observando el comportamiento de un equipo resulta complicado, y cuestiona el necesario rigor metodológico, ya que dicha observación constituye la base de un juicio de efectos innegables en la evaluación. Por ello, sería más eficiente si, por ejemplo, la clase la comparten dos docentes, realizando solo uno de los dos la observación mientras el otro está coordinando el trabajo. Aunque de entrada esto puede resultar ilusorio, esto es bastante factible puesto que podría realizarse en momentos puntuales del curso y no durante todo el mismo, obviamente intercambiando los papeles entre los docentes y sus propias asignaturas para hacer posible la experiencia en ambos casos.

Finalmente, creemos que esta metodología es susceptible de ser integrada con otras, como la Open Video Annotation (Blomberg, Renkl, Sherin, Borko y Seidel, 2013; Es et al., 2014; Gallego Arrufat y Díaz Martín, 2015; Nielsen, 2015), con objeto de compensar las debilidades señaladas, y fortalecer sus bondades. Así, unido a estos análisis, recomendamos algunas otras posibilidades de mejora como sería grabar las secuencias de vídeo y su posterior análisis entre docentes para compartir criterios de evaluación. Otra posibilidad, que deberemos explorar a partir de aquí, es la de permitir que los estudiantes pudieran realizar comentarios sobre conductas propias y ajenas en la clase, mediante videos grabados, intercambiando así el sentido y significado de los comportamientos entre estudiantes y docentes, siguiendo al menos en parte el concepto de clase invertida (Bishop y Verleger, 2013).

## 5. Referencias bibliográficas

- Alzina, R. B., Olmo, F. M., Soler, M. O., y Escoda, N. P. (2006). Evaluación de 360º: Una aplicación a la educación emocional. *Revista de Investigación Educativa*, 24(1), 187-203.
- Arzate, O. (2013). Coaching Educativo: Una Propuesta Metodológica para Innovar en el Aula. *Ra Ximhai*, 9(4), 177-185.
- Benejam Arguimbau, Pilar (2007). Falso dilema en la formación del profesorado de secundaria. *Periódico El País, Sección Tribuna*, 30 de Mayo de 2007.
- Bishop, J., y Verleger, M. (2013). The flipped classroom: A survey of the research. ASEE Annual Conference [www.asee.org](http://www.asee.org). Obtenido de <http://goo.gl/ZpIdxi> (revisado 16-02-2016).
- Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., y Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online / Journal Für Bildungsforschung Online*, 5(1), 90-114.
- Cano, E. (2008). La evaluación por competencias en la educación superior. *Profesorado: revista de curriculum y formación del profesorado*, 12(3), 11.

- Casanova, María Antonia (2012). El diseño curricular como factor de calidad educativa. *REICE*, nº 10.4, 6-20.
- Casero-Ripollés, Andreu (2012). Más allá de los diarios: el consumo de noticias de los jóvenes en la era digital. *Comunicar* nº 39, vol. 20.
- Castells, M. (2012). *Redes de indignación y esperanza: Los movimientos sociales en la era internet*. Madrid: Alianza Editorial.
- Castells, M. (2011). Democracy in the age of the Internet. *Transfer: journal of contemporary culture*, nº 6, 96-103.
- Dahl, Robert (2004). La democracia. *POSTData*, nº 10, 11-55. Diciembre 2004.
- Es, E. A. v., Tunney, J., Goldsmith, L. T., y Seago, N. (2014). A framework for the facilitation of teachers' analysis of video. *Journal of Teacher Education*, 65(4), 340-356. doi:10.1177/0022487114534266.
- Fernández, A. R., Ramos-Díaz, E., Ros, I., & Zuazagoitia, A. (2016). Implicación escolar de estudiantes de secundaria: la influencia de la resiliencia, el autoconcepto y el apoyo social percibido. *Educación XXI: en prensa*.
- Gallego Arrufat, M., y Díaz Martín, C. (2015). Actitud del alumnado hacia la investigación en educación: Trabajando con vídeos en estudios de grado. *Universitas Tarraconensis. Revista De Ciències De l'Educació*, 1(1), 8-29.
- Panadero-Calderón, E. y Alonso-Tapia, J. (2013). Revisión sobre autoevaluación educativa: evidencia empírica de su implementación a través de la autocalificación sin criterios de evaluación, rúbricas y guiones. *Revista de Investigación en Educación*, 11(2), 172-197.
- Pérez, E. B., Gómez-Linares, A., y Martín, R. P. (2015). La formación en valores en el EEES: propuesta de instrumento, medio y técnicas de evaluación. *Tendencias emergentes en evaluación formativa y compartida en docencia: El papel de las nuevas tecnologías*, 208, 215.
- Poy-Castro, R., Mendaña-Cuervo, C., y González, B. (2015). Diseño y evaluación de un juego serio para la formación de estudiantes universitarios en habilidades de trabajo en equipo. *RISTI-Revista Ibérica de Sistemas e Tecnologías de Informação*, (SPE3), 71-83.
- Rodríguez Adrados, Francisco (2007). Enseñanza y conocimiento. *Periódico El País, Sección Tribuna, 19 de Marzo de 2007*.
- Manolov, R., Leiva, D., Però, M., Solanas, A., Guàrdia, J., y Losada, J. L. (2015). La evaluación individual, grupal y por pares, ¿Qué aporta en “Técnicas de investigación”? *Revista del Congrés Internacional de Docència Universitària i Innovació (CIDUI)*, (2).
- Nielsen, B. L. (2015). Pre-service teachers' meaning-making when collaboratively analysing video from school practice for the bachelor project at college. *European Journal of Teacher Education*, 38(3), 341-357. doi:10.1080/02619768.2014.983066.
- Vadillo, M. T. P. (2013). *Liderazgo y motivación de equipos de trabajo*. ESIC Editorial.
- van-der Hofstadt Roman, C., y Gras, J. M. G. (2013). *Competencias y habilidades profesionales para universitarios*. Ediciones Díaz de Santos.