

Diferentes métodos de evaluación a través del uso de los clickers. Comparación de la utilidad percibida por los estudiantes

Different evaluation methods through by the use of clickers. Comparison of students' perceived utility

Estrella Fernández, Rebeca Cerezo, Marta Méndez, Natalia Suárez y Alejandra Dobarro

Universidad de Oviedo

Resumen

El objetivo del presente estudio se ha centrado en analizar la percepción de utilidad del alumnado con respecto a diferentes formas de evaluación continua o formativa, así como la posible existencia de diferencias entre ellas, contando con los Sistemas de Respuesta Personal o “clickers” en el proceso. Para ello 243 estudiantes universitarios, utilizando estos dispositivos, participaron en diferentes sesiones y modalidades de evaluación en las que se combinaron cuatro variables (presentación de las preguntas individual o en grupo; evaluación continua o sumativa; evaluación solo formativa; diferentes incentivos por la participación). Los resultados muestran como los estudiantes, en general, perciben esta metodología útil para favorecer el estudio continuo y autónomo de las asignaturas. Asimismo, los resultados muestran como un mayor incentivo en la nota final y conocer el momento en que se va a realizar la sesión de evaluación fomenta una mayor percepción de utilidad entre los estudiantes.

Palabras clave: clickers, estudiantes universitarios, evaluación continua, evaluación formativa, percepción de utilidad y sistemas de respuesta personal.

Abstract

The purpose of the study is the analysis of students' perceived utility according to different methods of continuous assessment. Besides, there are of interest the differences between those methods, carried out through the use of an Personal Response System o “clickers”. In this study participated 243 university students. There were developed different sessions and assessment methods (questions showed individually or in group; continuous assessment; training evaluation; different incentives for participation). Results showed that students, in general, perceive these systems as useful to their daily and autonomous study. In the same way, results showed higher students' perceived utility when their participation is considered in their grade and when they know the moment when the assessment session will be done.

Keywords: clickers, university students, continuous assessment, training evaluation, perceived utility and active responses systems.

La necesidad de centrar el proceso de enseñanza-aprendizaje en el alumno implica cambios importantes en el diseño de las actividades académicas y en los roles de docentes y discentes. Se trata de lograr una mayor comprensión y profundización en los contenidos por parte del estudiante, fomentando así un aprendizaje significativo de los mismos. Desde este nuevo enfoque, ha comenzado a considerarse la evaluación no solo como una mera herramienta acreditadora, “*evaluación del aprendizaje*”, sino como un instrumento anclado en la instrucción, que puede servir para motivar, guiar e incentivar el trabajo autónomo y autorregulado del alumnado, “*evaluación para el aprendizaje*” (Crisp, 2012; Hernández, 2012; López, 2012). Aunque no parece existir un acuerdo generalizado a la hora de definir y clasificar los diferentes tipos de evaluación, en general, y además de la evaluación final tradicional, podrían considerarse al menos otros tres tipos de evaluación. Por un lado, se contaría con la evaluación sumativa y continua, que tiene lugar en diferentes momentos, pero cuya función es principalmente juzgar el grado de dominio de los estudiantes sobre la materia. Esta ha sido relacionada con mejores resultados académicos debido a que favorece el trabajo continuado de los alumnos durante el curso (Arribas, 2012; Delgado y Cuello, 2009). En segundo lugar, estaría

la evaluación formativa, que también tiene lugar en diferentes momentos a lo largo del curso, pero con el fin de ofrecer feedback a los estudiantes para que puedan dirigir y supervisar sus tareas, ajustarse al contexto y a las exigencias académicas, analizar su nivel de comprensión de las materias y valorar en qué medida los contenidos de la asignatura están siendo adquiridos. Este proceso de monitorización y análisis del aprendizaje que se potencia con la evaluación formativa fomenta un aprendizaje más profundo de los contenidos, repercutiendo, en ocasiones, en mejores resultados académicos (Carrillo de la Peña y Pérez, 2012; Nicol y MacFarlane-Dick, 2006). Y, por último, podría tenerse en cuenta una tercera opción, resultado de la combinación de las dos anteriores, es decir, evaluación continua (o sumativa) y formativa, en la que las diferentes sesiones de evaluación, además de cumplir una función acreditadora, sirven para proporcionar feedback al estudiante, fomentando así la revisión y control del alumno sobre su aprendizaje. Este tipo de evaluación se entiende que podría resultar estimulante, ya que el feedback ofrecido es una potente herramienta que ayuda a los alumnos a autorregular su proceso de estudio y, además, puede favorecer el logro de mejores resultados académicos (Coll, Rochera, Mayordomo y Naranjo, 2007). De esta manera, se esta-

ría implicando a todos los alumnos, no sólo a aquellos cuyas metas académicas son intrínsecas y están orientadas al aprendizaje sino también a aquellos cuyas metas son principalmente extrínsecas y están orientadas al logro (Valle, Cabanach, Rodríguez, Núñez y González-Pienda, 2006; Weurlander, Söderberg, Scheja y Wernerson, 2011).

Parece que introducir un sistema de evaluación en el aula, cuando va acompañada de feedback y cuando además éste es inmediato (Ghosh y Renna, 2010; Trees y Jackson, 2007), puede ofrecer información útil tanto para el alumnado, como para el profesorado. Estas estrategias, orientadas a mantener al estudiante activo durante su proceso de aprendizaje, pueden facilitar que los alumnos seleccionen correctamente la información relevante de la asignatura, la organicen para poder integrarla con los conocimientos previos sobre la misma y, una vez recibido el feedback sobre sus respuestas, para que pongan en marcha estrategias metacognitivas con el fin de mejorar su aprendizaje y preparar las posibles preguntas de un examen (Anthis, 2011; Mayer, 2008).

Los Sistemas de Respuesta Personal (*Personal Response Systems*) o “clickers”, muy similares a otros instrumentos como los *Classroom Response Systems* (Fies y Marshall, 2006), *Audience Response Systems* (Vana, Silva,

Muzyka e Hiriani, 2011), *Electronic Response Systems* (Ghosh y Renna, 2009), *Interactive Classroom Communication Systems* (Kay, 2009), *Student Response Systems* (Gok, 2011), etc., son herramientas que permiten generar una serie de preguntas a través de un software y recoger las respuestas a las mismas de manera automática, utilizando dispositivos muy similares a un teléfono móvil. Con ellos los profesores pueden evaluar el nivel de comprensión del grupo de una manera sencilla, rápida y eficaz y ofrecer un feedback inmediato al alumnado, tanto antes, como durante y después de las sesiones de aprendizaje (MacArthur y Jones, 2008). De esta forma se insta a los alumnos a reflexionar sobre su proceso de aprendizaje, estimulando el uso de estrategias metacognitivas que les ayuden a estructurar y a tomar conciencia acerca de su conocimiento sobre la materia. Además, los clickers han sido utilizados con la intención de promover la asistencia al aula, favorecer una interacción y comunicación bidireccional profesor-alumno, aun cuando hay un elevado número de alumnos (Fies y Marshall, 2008; Gauci, Dantas, Williams y Kemm, 2009; Kay y LeSage, 2009; Tlhoaele, Hoffman, Naidoo y Winnips, 2014), mantener un procesamiento activo cognitivo durante las clases (Kay y LeSage, 2009; Mayer et al., 2009) y, por tanto, directa o indirectamente, también han

sido herramientas utilizadas para potenciar una mejora en el rendimiento académico (Kay y LeSage, 2009; King y Joshi, 2008; Mayer et al., 2009).

No obstante, no es el uso de los dispositivos en sí mismo lo que favorece estos resultados, sino el tipo de diseño metodológico que subyace a su aplicación (Caldwell, 2007; Fies y Marshall, 2008; Morling, McAuliffe, Cohen y DiLorenzo, 2008; White, Syncox y Alters, 2011). Los investigadores y profesores han utilizado distintas modalidades de evaluación. Algunas difieren de otras en el tipo de preguntas, ya sea utilizando gráficos, imágenes o texto en su presentación (Gray, Owens, Liang y Steer, 2012). En función del momento de la evaluación: antes, durante o después de las clases. Otras, sin embargo, varían en el peso de las evaluaciones sobre la nota final o en los incentivos que los estudiantes pueden recibir con su participación (Oswald y Rothen, 2014; Yourstone, Kraye y Albaum, 2008).

A pesar de la rápida incorporación de los clickers a las aulas, sobre todo, en el contexto anglosajón (Hoon, 2014), no se utilizan habitualmente como sistema de evaluación continua-sumativa y formativa, dando un peso total a las respuestas de los estudiantes en la nota final y, al mismo tiempo, ofreciendo un feedback que fomente la reflexión del estudiante sobre su propio proceso de aprendizaje (Cald-

well, 2007); y, por tanto, tampoco se ha valorado la percepción de los estudiantes en esta modalidad resultando interesante su investigación (Keller et al., 2007). Tal vez porque no se han tenido en cuenta alternativas a la clásica presentación de las preguntas de autoevaluación (proyectadas para todo el aula) para evitar ciertos problemas como, por ejemplo, que los alumnos puedan ver las respuestas de sus compañeros o que se escuden en fallos del programa a la hora de recoger la información enviada (Fies y Mashall, 2008; White et al., 2011). La mayor parte de los profesores han optado por un sistema de incentivos con un pequeño impacto en la nota final (entre el 10% y 15% de la nota total) en función de las respuestas correctas dadas por los estudiantes (evaluación formativa con incentivos) (White et al., 2011).

Hay controversia en cuanto a la valoración de la herramienta cuando se tiene en cuenta el peso de las respuestas en las notas finales. Las investigaciones sostienen que los estudiantes podrían responder de una manera más comprensiva cuando se incentiva la exactitud de sus respuestas (White et al., 2011). Sin embargo, parece que es posible que este sistema no sea el diseño mejor valorado por los alumnos ya que, con esta modalidad, podrían percibir su participación como obligatoria y más exigente, pudiendo suponer una traba para mantener un ambiente de

aprendizaje positivo (Kay y LeSage, 2009) y, por tanto, profundo y significativo. La valoración de los alumnos sobre la utilidad con los clickers podría variar en función de la modalidad de evaluación en la que están participando, pero también de las variables que se estén midiendo.

Con el fin de encontrar estrategias docentes útiles para fomentar un aprendizaje profundo y significativo, se ha llevado a cabo una experiencia con estudiantes universitarios, contando con los *Sistemas de Respuesta Personal* en la iniciativa. Por un lado, se han integrado los dispositivos en varias asignaturas para poder ofrecer feedback inmediato y en grupo al alumnado. Por otro, se han utilizado diferentes modalidades de evaluación (presentación de las preguntas individual o en grupo; evaluación continua y sumativa; evaluación sólo formativa; diferentes incentivos por la participación) con el fin de determinar qué combinación, de todas las posibilidades existentes, puede resultar más útil, según la percepción del alumnado, de acuerdo con diferentes variables implicadas en el proceso de aprendizaje y que pueden verse influenciadas por los distintos métodos de presentación: gestión diaria del proceso de aprendizaje, interacción en el aula, autonomía en la integración de la información y rendimiento. Se asume que los alumnos que participan en modalidades con

mayor peso en la nota final serán los que valoren con mejor puntuación la utilidad de los clickers con respecto a estas variables; es decir: los alumnos de aquellas asignaturas en las que la participación tiene mayor repercusión en la nota final serán los que a) considerarán que la evaluación con los clickers favorece en mayor medida la gestión diaria del proceso de estudio; b) considerarán que la metodología usada favorece una mayor integración de la información; c) que los clickers favorecen la interacción en el aula por encima del resto de modalidades; y d) que puede repercutir con más intensidad en el rendimiento académico. Seguidos de los alumnos que reciben incentivos en función de la exactitud de sus repuestas y de los alumnos que participan en modalidades de evaluación formativa (percepción de utilidad mayor cuando están informados de la fecha de la sesión).

Método

Participantes

En la investigación han participado 243 estudiantes universitarios, matriculados en la Universidad de Oviedo durante el curso académico 2011-2012 en diferentes asignaturas de distintas titulaciones, aunque todas ellas afines al ámbito de la Psicología y de la Edu-

cación. Noventa y un estudiantes de la asignatura de *Neurociencias* de primer curso del Grado en Psicología (*grupo G1*); 51 alumnas de la asignatura de *Dificultades de Aprendizaje* del segundo curso del Grado en Educación Infantil (*grupo G2*); 69 estudiantes de la asignatura de *Desarrollo, Comunicación y Adquisición del Lenguaje* del primer curso de Grado en Psicología (*grupo G3*); y 32 alumnos del total de alumnos matriculados en la asignatura de *Trastornos y Dificultades del Aprendizaje* de tercer curso de Licenciatura en Psicología (*grupo G4*).

Variables e instrumentos

El objetivo del estudio es valorar la utilidad percibida por parte de los alumnos de diferentes metodologías de enseñanza-aprendizaje, a través de evaluaciones en el aula cuyo eje central es el uso de los “clickers”, así como comparar si existen diferencias en función de la modalidad de evaluación empleada. Para ello se elaboró y aplicó un cuestionario *ad hoc*. El cuestionario consta de dieciocho ítems presentados en un formato tipo Likert (1 nada; 2 poco; 3 bastante; 4 mucho; 5 totalmente). El análisis factorial de componentes principales, utilizando el método de rotación Oblimin (se seleccionó esta rotación por considerar que los factores resultantes podrían estar relacionados) reveló cuatro fac-

tores, contemplados en la elaboración del cuestionario, relacionados con el proceso de aprendizaje del alumnado según el método de trabajo en el aula (evaluaciones) y herramientas utilizadas (clickers). El primer factor “*Gestión*”, formado por siete ítems (e.g., *El uso de los clickers ha favorecido que trabajas habitualmente en la lectura de los apuntes*), hace referencia al uso de estrategias de apoyo, en concreto, de organización y planificación del estudio de manera constante (a diario). El segundo factor, “*Interacción en el aula*”, cuenta con tres ítems (p.e., *Te ha resultado más fácil pedir ayuda al/ los profesor/es para resolver dudas*) relacionados con el fomento de la comunicación y feedback directo e inmediato entre el profesor y el alumno. El tercer factor, “*Integración de la información*” (p.e., *Te ha animado a buscar materiales complementarios o a acudir a otras fuentes de información*) está constituido por cuatro ítems relacionados con la autonomía del alumno para poner en marcha estrategias de búsqueda e integración de la información (análisis de los conocimientos previos y búsqueda de información complementaria e integración con la previa). Por último, el cuarto factor, “*Rendimiento*”, está relacionado con la percepción de utilidad de la metodología y herramientas empleadas para conocer previamente la estructura del examen final y lograr un mejor

rendimiento académico en la asignatura (e.g., *Te ha servido para “perder el miedo” al examen final*). El Alpha de Cronbach total del cuestionario es de .80. Además, se habilitó un espacio para que los alumnos pudieran recoger, cualitativamente, aspectos positivos y negativos del uso de los dispositivos.

Procedimiento

Aunque en líneas generales el procedimiento seguido en la presente investigación ha sido similar en las cuatro aulas ya que se han utilizado los clickers en todas ellas, hay matices y diferencias sustanciales en la metodología utilizada entre las mismas. A continuación se describe el procedimiento utilizado de manera genérica con todos los alumnos y, posteriormente, se detalla el seguido con cada uno de los grupos.

A través de los Sistemas de Respuesta Personal (*Personal Response Systems*) o “clickers”, los profesores responsables de las asignaturas diseñaban diferentes sesiones de evaluación para el aula través de preguntas de elección múltiple (entre 10 y 20 cuestiones) sobre los contenidos teóricos y aplicados de su asignatura. Se ha optado en todos los casos por preguntas tipo test con múltiples opciones, generalmente, textuales o una combinación de preguntas gráficas, visuales y textuales en función de la materia, ofreciendo así diferentes sistemas de

estimulación, visual y verbal, que favorecen la integración de la información y acceso a las respuestas (Gray et al., 2012; Mayer et al., 2009). En función del diseño metodológico que subyace al procedimiento seguido con cada aula, el profesor utilizaba un sistema de presentación de las preguntas grupal o individual. Con el sistema grupal, el docente proyectaba las preguntas y, a través de un USB conectado al ordenador, iba recogiendo las respuestas de los estudiantes y visualizando los resultados una a una. Con el sistema individual, cada estudiante disponía, en papel, de las preguntas y respondía en el mismo a la vez que iba registrando las respuestas en el dispositivo. De esta forma se evitaba que los alumnos pudieran copiarse o presentar quejas posteriores relacionadas con fallos en el programa (Fies y Mashall, 2008; White et al., 2011). Una vez que los alumnos habían guardado sus respuestas, se recogían a través del USB todas a la vez, pudiendo ver los resultados de cada una al finalizar la sesión. Independientemente de la modalidad de presentación de las preguntas, las respuestas quedaban grabadas en una base de datos, generando un portafolio de calificaciones individualizado (cada alumno tenía asignado un clicker específico vinculado a su nombre). Como el software permite editar una serie de informes automáticamente, con gráficos de los resultados, se reali-

zaron las correcciones a todas las preguntas al finalizar las sesiones. Estos gráficos, que no reflejan la respuesta de cada usuario sino el cómputo global del aula, facilitaron la corrección colectiva de las preguntas, fomentando la participación y cooperación de toda la clase en su resolución (Fitch, 2004) (ver Figura 1).

En la asignatura de Neurociencias (G1) se siguió un sistema de evaluación continua-sumativa y formativa, es decir, durante el tiempo que duró la asignatura (primer semestre del curso escolar) se realizaron siete sesiones de evaluación en el aula (aunque la primera fue de toma de contacto con los clickers y no se tuvo en cuenta para la nota final). Los alumnos estaban avisados del día en que se iba a realizar la evaluación. Los estudiantes podían superar la asignatura cuando la nota media de dichas evaluaciones era suficiente para obtener un aprobado o más. No obstante, en todos los casos, podían presentarse a un examen final, tanto para superar la asignatura como para subir su nota. Dado que el peso de los resultados en esta modalidad es realmente relevante, se utilizó el sistema de respuesta individual, aunque el análisis de las respuestas a cada pregunta y el feedback se realizaron de manera grupal en el aula.

En la asignatura de Dificultades de Aprendizaje en Educación Infantil (G2) se siguió un sistema de eva-

luación formativa, con tres sesiones, previo aviso de los alumnos. Los estudiantes podían obtener un incentivo para la nota final en función las respuestas correctas emitidas en cada una de las sesiones (15% de la nota final). En este caso, puesto que el peso de la participación en la nota final era menor, se utilizó el sistema de presentación de las preguntas grupal. El examen final de la asignatura se realizó a través de los clickers para poder realizar la sesión de revisión grupal y ofrecer feedback a los alumnos (en el examen el sistema de presentación de las preguntas fue individual).

En la asignatura de Desarrollo, Comunicación y Lenguaje (G3), se siguió un sistema de evaluación formativa, tres sesiones, previo aviso de los alumnos. Los alumnos no obtenían ningún tipo de incentivo por la participación, por lo que el sistema de presentación de las preguntas fue grupal, al igual que la revisión y discusión de las respuestas sobre las mismas.

En Trastornos y Dificultades del Aprendizaje de Psicología (G4), se siguió también un sistema de evaluación formativa, tres sesiones, pero los alumnos no estaban informados de la fecha en la que iban a tener lugar cada una de ellas. También en este caso se utilizó un sistema de presentación, revisión y discusión de las preguntas grupal.

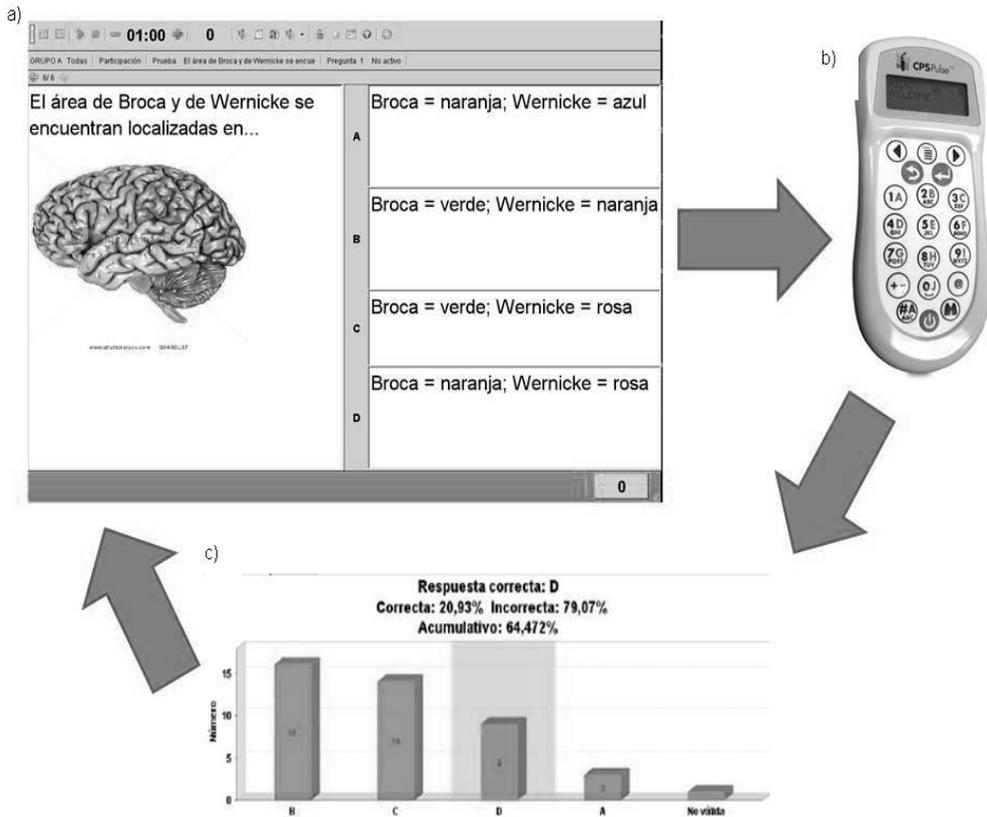


Figura 1. Uso de los “clickers” en el aula: a) proyección de la pregunta; b) recogida de las respuestas de los alumnos a través de los “clickers”; c) presentación del informe de aciertos y errores.

En la asignatura de Neurociencia (G1) se siguió un sistema de evaluación continua-sumativa y formativa, es decir, durante el tiempo que duró la asignatura (primer semestre del curso escolar) se realizaron siete sesiones de evaluación en el aula (aunque la primera fue de toma de contacto con los clickers y no se tuvo en cuenta para la nota final). Los alumnos estaban avisados del día en que se iba a realizar la evaluación. Los estudiantes podían superar la asignatura cuando la nota

media de dichas evaluaciones era suficiente para obtener un aprobado o más. No obstante, en todos los casos, podían presentarse a un examen final, tanto para superar la asignatura como para subir su nota. Dado que el peso de los resultados en esta modalidad es realmente relevante, se utilizó el sistema de respuesta individual, aunque el análisis de las respuestas a cada pregunta y el feedback se realizaron de manera grupal en el aula.

En la asignatura de Dificultades

de Aprendizaje en Educación Infantil (G2) se siguió un sistema de evaluación formativa, con tres sesiones, previo aviso de los alumnos. Los estudiantes podían obtener un incentivo para la nota final en función las respuestas correctas emitidas en cada una de las sesiones (15% de la nota final). En este caso, puesto que el peso de la participación en la nota final era menor, se utilizó el sistema de presentación de las preguntas grupal. El examen final de la asignatura se realizó a través de los clickers para poder realizar la sesión de revisión grupal y ofrecer feedback a los alumnos (en el examen el sistema de presentación de las preguntas fue individual).

En la asignatura de Desarrollo, Comunicación y Lenguaje (G3), se siguió un sistema de evaluación formativa, tres sesiones, previo aviso de los alumnos. Los alumnos no obtenían ningún tipo de incentivo por la participación, por lo que el sistema de presentación de las preguntas fue grupal, al igual que la revisión y discusión de las respuestas sobre las mismas.

En Trastornos y Dificultades del Aprendizaje de Psicología (G4), se siguió también un sistema de evaluación formativa, tres sesiones, pero los alumnos no estaban informados de la fecha en la que iban a tener lugar cada una de ellas. También en este caso se utilizó un sistema de presentación, revisión y discusión de las preguntas

grupal.

Análisis de datos

A través de un análisis descriptivo de los datos obtenidos en el cuestionario se examinó la percepción de utilidad de los sistemas de evaluación propuestos. Para comprobar si existían diferencias estadísticamente significativas entre los distintos sistemas metodológicos de evaluación seguidos con cada aula, y determinar cuál de ellos podría suponer la mejor opción para utilizar en el futuro en la universidad, se realizaron análisis multivariados de la varianza (MANOVA) y valoración del tamaño del efecto (Frías, Pascual, y García, 2000). Se tuvo en cuenta como variable dependiente el sistema de evaluación seguido en cada aula (grupo-clase) y como variables independientes los cuatro factores relacionados con el proceso de enseñanza-aprendizaje medidos a través del cuestionario (Gestión, Interacción en el aula, Integración de la información y Rendimiento). Puesto que la variable dependiente cuenta con más de dos grupos, se utilizó como prueba de comparaciones múltiples post-hoc la prueba de Scheffé.

Resultados

En la Tabla 1 se muestran los resultados de los análisis descriptivos (me-

días, desviaciones típicas, asimetría, curtosis) correspondientes a las variables dependientes incluidas en el MANOVA, y las medias y desviaciones típicas teniendo en cuenta la modalidad de evaluación. Se observan valores de asimetría y curtosis razonables dentro del intervalo 1 y -1 indicando que no hay más casos de los aconsejables en los valores extremos de la escala de medida (mínimo-1; máximo-5). Los resultados muestran como, en general y según la opinión de los alumnos, utilizar sistemas de evaluación continua o formativa a través de clickers resulta bastante útil para la gestión e integración de la información y el rendimiento académico. Además, se observa como aquellas modalidades de evaluación que tienen algún peso en la nota final tienden a ser mejor valoradas en cuanto a la gestión e integración de la información. Sin embargo, esta tendencia no se sostiene cuando las variables analizadas tienen que ver con la interacción profesor-alumno y el rendimiento académico.

Para el estudio de las diferencias entre cada uno de los sistemas de evaluación se procedió al análisis de las diferencias inter-grupo mediante el uso de MANOVA. Los resultados obtenidos indican que, en general, existen diferencias estadísticamente significativas entre los cuatro grupos de estudiantes [Λ de Wilks = .545; $F(12.625) = 13.41$; $p < .001$; $\eta^2 =$

.183], lo cual indica que la modalidad de implementación tiene resultados diferenciales respecto de los cuatro factores del cuestionario aplicado.

Sin embargo, analizando de modo individual los resultados para cada uno de los cuatro factores, los resultados aportados por el MANOVA indican que las diferencias son estadísticamente significativas únicamente para tres de los cuatro factores medidos: gestión [$F(3.239) = 15.83$; $p < .001$; $\eta^2 = .166$], integración de la información [$F(3.239) = 9.73$; $p < .001$; $\eta^2 = .109$] y rendimiento [$F(3.239) = 3.67$; $p = .013$; $\eta^2 = .044$]; pero no son significativas para la interacción profesor-alumno en el aula [$F(3.239) = 1.19$; $p = .313$; $\eta^2 = .015$].

A través de la prueba Scheffé de comparaciones múltiples se pudo observar entre qué grupos, en función del sistema de evaluación implementado, se daban esas diferencias (ver Tabla 2). Los resultados revelaron que para el factor gestión las diferencias se encontraban entre el grupo G1, que llevó a cabo una evaluación continua y formativa con total peso en la nota final, y el grupo G3 y G4, que no recibió recompensa alguna en la nota final; mientras que entre el grupo G1 y el G2 (que llevó a cabo una evaluación formativa con un pequeño peso en la nota final) no se observaron diferencias estadísticamente significativas. También se observaron diferencias estadísti-

camente significativas entre el grupo G4 (evaluación formativa sin previo aviso) y el resto de sistemas de evaluación implementados para el factor gestión, siendo la percepción de utilidad para la gestión menor en el grupo G4 que en el resto de los grupos.

En cuanto a la variable integración de la información se encontraron diferencias estadísticamente significativas entre el grupo G1 y el G2 y G1 y G4, pero no entre el G1 y el G3, siendo la percepción de utilidad mayor en el grupo G1 que en el resto de los grupos. Asimismo, se observaron diferencias estadísticamente significativas entre el grupo G2 (evaluación formativa con un pequeño peso en la nota final) y G3 (sin recompensa alguna en la nota final), a favor de este último.

Por último, con respecto a variable rendimiento, únicamente se observaron diferencias entre el grupo G1 y G2, autoinformando los alumnos del segundo grupo de una mayor percepción de utilidad con respecto del primero para el factor rendimiento. En cuanto a la interacción profesor-alumno en el aula, tal como indicaban los resultados del MANOVA, no se observaron diferencias estadísticamente significativas entre ninguno de los grupos.

Para valorar los aspectos positivos y negativos planteados por los alumnos se realizó una categorización de las respuestas dadas por los mismos (ver Tabla 3), resultando tanto los aspectos

positivos como los negativos similares en todos los casos, es decir, independientemente del sistema de evaluación empleado. Los aspectos positivos que los alumnos observan en mayor medida son que favorece llevar la materia al día y les ayuda a obtener un feedback de su grado de comprensión de la misma. Los aspectos negativos mayormente observados tienen que ver con el miedo a equivocarse a la hora de enviar las respuestas, incluso cuando el sistema de recogida es individual y en caso de error es posible rectificarlo con las respuestas registradas a través de papel y lápiz, y con el tiempo o tipo de presentación de las preguntas, destacando, además, un porcentaje pequeño de alumnos que considera como negativo el hecho de que “obligan”, de alguna forma, a asistir a clase y llevar la materia al día.

Discusión y conclusiones

El objetivo del presente estudio se ha centrado en analizar la percepción de utilidad del alumnado con respecto a diferentes formas de evaluación continua o formativa, así como la posible existencia de diferencias entre ellas, contando con los Sistemas de Respuesta Personal en el proceso. Las formas de evaluación propuestas han combinado tres variables en función de las cuales han resultado las cuatro

Tabla 1.

Media, desviación típica, asimetría y curtosis de las variables de aprendizaje evaluadas en general y media y desviación de cada grupo de evaluación en estas variables

	<i>Media</i>				<i>DT</i>				<i>Asim</i>	<i>Curt</i>
Gestión	3.60				.97				-.49	-.41
	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>		
	3.92	3.80	3.42	2.73	.87	.80	.99	.87		
Interacción	2.84				.95				-.059	-.39
	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>		
	2.71	2.98	2.84	2.98	1.06	.72	.90	1.06		
Integración	3.16				.85				.053	-.40
	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>		
	3.45	2.78	3.21	2.81	.90	.60	.76	.88		
Rendimiento	3.57				.87				-.57	.22
	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>		
	3.41	3.85	3.65	3.37	.94	.75	.74	.98		

DT = desviación típica; Asim = asimetría; Curt =curtosis.

Tabla 2.

Diferencia de medias entre los grupos en función de la metodología de evaluación implementada para cada de las variables dependientes

		Grupos					
		G1-G2	G1-G3	G1-G4	G2-G3	G2-G4	G3-G4
Gestión	Dif. M	.120	.501	1.189	.380	1.069	.688
	Sig.	.897	.007	.000	.151	.000	.005
Interacción	Dif. M	-.270	-.130	-.279	.140	-.009	-.149
	Sig.	.445	.866	.567	.889	1.000	.911
Integración	Dif. M	.666	.235	.633	-.431	-.033	.398
	Sig.	.000	.342	.003	.040	.998	.150
Rendimiento	Dif. M	-.441	-.240	-.037	.201	.448	.277
	Sig.	.036	.379	.998	.656	.108	.514

Tabla 3.

Comentarios de los alumnos acerca de los aspectos positivos o negativos de los sistemas de evaluación propuestos (%)

Aspectos positivos	%
Ayuda a repasar y llevar la materia al día.	47
Ayuda a autoevaluarte, resolver dudas y te ofrece feedback sobre tu grado de aprendizaje de los contenidos.	35
Es innovador y motivador.	9
Es sencillo y rápido de usar.	9
Aspectos negativos	%
Es fácil equivocarse al introducir la respuesta y es fácil que el sistema falle o se registren las respuestas de manera equivocada.	68
Cambiaría el tiempo y forma de presentación de las preguntas.	20
Te obliga a asistir regularmente a clase y a llevar la materia al día.	9
Resulta estresante.	3

modalidades de evaluación: forma de presentación de las preguntas (individual o grupal); relevancia de las respuestas sobre la nota final en la asignatura; y posibilidad de anticipación de las sesiones de evaluación.

En general, y al igual que en investigaciones anteriores (Fies y Marshall, 2009; Kay y LeSage, 2009), los alumnos perciben la incorporación de sesiones de evaluación continua o formativa, a través de dispositivos de respuesta interactiva, como bastante útiles para organizar la asignatura y favorecer un estudio al día de la misma; así como para favorecer la autonomía del alumno para poner en marcha estrategias de búsqueda e integración de la información relevante y, también, como sistema a través el cual poder conocer previamente la estructura del

examen final y lograr un mejor rendimiento académico en la asignatura. Sin embargo, los estudiantes que han participado en el presente estudio, no parecen valorar tan positivamente la utilidad de los sistemas de evaluación incorporados para interactuar con el profesorado, como ocurría en estudios como los mencionados. Tal vez, porque los grupos-aula con los que se ha trabajado en la presente investigación no son especialmente numerosos, entendiéndose que el alumnado, normalmente, tiene la posibilidad de interactuar con el profesorado, algo que es más complicado en aquellos casos en los que las aulas están formadas por un número más elevado de alumnos. Además, la revisión de las respuestas se realizaba de manera grupal en el aula, dando la posibilidad a los estudiantes

de discutir entre ellos sobre cuáles eran las respuestas acertadas y, seguramente, favoreciendo más la interacción entre pares que entre alumnos y profesores. Sería interesante tener en cuenta la percepción sobre la interacción con los compañeros en futuras investigaciones, tal como han estudiado otros autores como Caldwell (2007) y Trees y Jackson (2007).

Con respecto a la percepción de los alumnos, en función de la modalidad de evaluación, los resultados confirman, en parte, las hipótesis de partida, siendo dicha percepción diferente en función del factor analizado: gestión, interacción en el aula, integración de la información o rendimiento. En cuanto a la gestión del trabajo por parte de los estudiantes, los resultados muestran cómo el peso en la nota final y conocer el momento en que se va a realizar la sesión de evaluación fomenta una mayor percepción de utilidad de los sistemas de evaluación propuestos. Asimismo, se observa que, si bien el peso en la nota final estimula en los alumnos la necesidad de llevar la asignatura al día, no hay diferencias entre un mayor o menor peso en las calificaciones. Es decir, los estudiantes consideran que el uso de evaluaciones continuas y formativas a través de los clickers les ayudan a gestionar la asignatura diariamente, independientemente de si a través de estas evaluaciones obtienen total o parcialmente la calificación fi-

nal de la asignatura. Teniendo en cuenta la integración de la información, se observan diferencias estadísticamente significativas en los sistemas de evaluación cuando éstos están planteados como sistemas de evaluación continua y formativa (con peso total en la nota final) o sólo formativa (sin peso en la calificación final). Es decir, según los resultados observados, parece que el sistema de evaluación en esta variable es mejor valorado cuando los estudiantes ven una utilidad a nivel de calificación o una utilidad a nivel formativo, pero no cuando comparte ambos criterios. Es importante tener en cuenta que estos resultados han de ser contrastados en investigaciones futuras, ya que pueden estar interfiriendo variables no controladas como el tipo de asignatura impartida o las características de los docentes de las mismas. Teniendo en cuenta el rendimiento, no se observan apenas diferencias estadísticamente significativas entre los grupos. No obstante, en general, los estudiantes consideran que cualquiera de los sistemas de evaluación empleados son bastante útiles a la hora de ayudarles a mejorar su rendimiento. Además, cabe destacar que, aunque autores como Oswald y Rothen (2014) plantearon que los incentivos a los alumnos pueden favorecer su rendimiento, los alumnos valoran más positivamente su utilidad cuando su participación no tiene peso en la nota final, o tiene

poco peso, siempre que estén informados del momento en que se realizará la evaluación, tal como recogen autores como Kay y LeSage (2009). Es posible que esto se deba a que, por un lado, estas evaluaciones fomentan en el estudiante un aprendizaje estratégico y más activo de las materias pero, por otro, el nivel de estrés o la preocupación ante los posibles fallos es menor, por lo que su percepción del uso de estas metodologías es más favorable con menos presión, independientemente de que eso repercuta realmente en el rendimiento o no (White et al., 2011). Sería interesante contrastar en el futuro el rendimiento real del alumnado que utiliza clickers con su percepción sobre los efectos en ese rendimiento, así como comparar los resultados con

alumnado que no utiliza estos sistemas de evaluación.

Es necesario tener en cuenta las limitaciones del presente trabajo, así como ampliar la investigación antes de extraer conclusiones que puedan generalizarse. Por un lado, existe un posible sesgo con respecto a las aulas, a los profesores, contenidos de las asignaturas y curso en el que están matriculados los estudiantes que no ha sido estudiado y controlado. Por otro, tanto los alumnos participantes como los profesores mantuvieron una experiencia limitada con los clickers, por lo que sería interesante realizar estudios longitudinales donde fuera posible controlar el efecto de la novedad en el uso de los dispositivos.

Referencias

- Anthis, K. (2011). Is it the clicker, or is it the question? Untangling the effects of student response system use. *Teaching of Psychology*, 38, 189-193.
- Arribas, J. M. (2012). El rendimiento académico en función del sistema de evaluación empleado. *Revista Electrónica de Investigación y Evaluación Educativa*, 18, 1-15.
- Delgado, A., y Cuello, O. (2009). Interacción entre la evaluación continua y la autoevaluación formativa: la potenciación del aprendizaje autónomo. *Red-U. Revista de Docencia Universitaria*, 4, 1-13.
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *Life Science Education*, 6, 19-20.
- Carrillo de la Peña, M. T., y Pérez, J. (2012). Continuous assessment improved academic achievement and satisfaction of psychology students in Spain. *Teaching*

- of Psychology*, 39, 45-47.
- Coll, C., Rochera, M. J., Mayordomo, M. R., y Naranjo, M. (2007). Evaluación continua y ayuda al aprendizaje. Análisis de una experiencia de innovación en educación superior con apoyo de las TIC. *Revista Electrónica de Investigación Psicoeducativa*, 13, 783-804.
- Crisp, B. (2012). Integrative assessment: Reframing assessment practice for current and future learning. *Assessment and Evaluation in Higher Education*, 37, 33-43.
- Fies, C., y Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, 15, 101-109.
- Fies, C., y Marshall, J. (2008). The C3 framework: evaluating classroom response system interactions in university classrooms. *Journal of Science Education and Technology*, 17, 483-499.
- Fitch, J. L. (2004). Student feedback in the college classroom: A technology solution. *Educational Technology Research and Development*, 52, 71-81.
- Frías, M. D., Pascual, J. y García, F. (2000). Tamaño del efecto del tratamiento y significación estadística. *Psicothema*, 12, 236-240.
- Gauci, S. A., Dantas, A. M., Williams, D. A., y Kemm, R. E. (2009). Promoting student-centered active learning in lectures with a personal response system. *Advances in Physiology Education*, 33, 60-71.
- Gok, T. (2011). An evaluation of student response systems from the viewpoint of instructors and students. *The Turkish online Journal of Educational Technology*, 10, 67-83.
- Gosh, S., y Renna, F. (2009). Using electronic response systems in economics classes. *Journal of Economic Education*, 40, 354-365.
- Gray, K., Owens, K., Liang, X., y Steer, D. (2012). Assessing multimedia influences on student responses using a personal response system. *Journal of Science Education and Technology*, 21, 392-402.
- Hernández, R. (2012). Does continuous assessment in higher education support student learning? *Higher Education: The International Journal of Higher Education and Educational Planning*, 64, 489-502.
- Hoon, J. (2014). Closing the missing links and opening the relationships among the factors: A literature review on the use of clicker technology using the 3P model.

- Educational Technology & Society*, 17, 150-168.
- Kay, R. H. (2009). Examining gender differences in attitudes toward interactive classroom communications systems (ICCS). *Computers & Education*, 52, 730-740.
- Kay, R. H., y LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education* 53, 819-827
- Keller, C., Finkelstein, N., Perkins, K., Pollock, S., Turpen, C., y Dubson, M. (2007). Research-based practices for effective clicker use. In Hsu, L. et al. (Eds.), *Physics education research conference* (pp. 128-131). Melville, NY: American Institute of Physics.
- King, D. B., y Joshi, S. (2008). Gender differences in the use and effectiveness of personal response devices. *Journal of Science Education and Technology*, 17, 544-552.
- López, V. M. (2012). Evaluación formativa y compartida en la universidad: clarificación de conceptos y propuestas de intervención desde la red interuniversitaria de evaluación formativa. *Psychology, Society, & Education*, 4, 117-130.
- MacArthur, J. R., y Jones, L. L. (2008). A review of literature reports of clickers applicable to college chemistry classrooms. *Chemistry Education Research and Practice*, 9, 189-195.
- Mayer, R. E. (2008). *Learning and instruction*. New York: Pearson Merrill Prentice Hall.
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, A., Bimber, B., Chun, D., (...), y Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34, 51-57.
- Morling, B., McAuliffe, M., Cohen, L., y DiLorenzo, T. (2008). Efficacy of personal response systems ("Clickers") in large, introductory psychology classes. *Teaching of Psychology*, 35, 45-50.
- Nicol, D. J., y Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31, 199-218.
- Oswald, K. M., y Rothen, S. E. (2014). Improving classroom clicker practices: Effects of incentives and feedback on retention. *North American Journal of Psychology*, 16, 79-88.

- Tlhoaele, M., Hofman, A., Naidoo, A., y Winnips, K. (2014). Using clickers to facilitate interactive engagement activities in a lecture room for improved performance by students. *Innovations in Education and Teaching International*, 51, 497-509.
- Trees, A. R., y Jackson, M. H. (2007). The learning environment in clicker classrooms: student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, 32, 21-40.
- Valle, A., Cabanach, R. G., Rodríguez, S., Núñez, J. C., y González-Pienda, J. A. (2006). Metas académicas, estrategias cognitivas y estrategias de autorregulación en el estudio. *Psicothema*, 18, 165-170.
- Vana, K., Silva, E., Muzyka., D., y Hirani, L. (2011). Effectiveness of an audience response system in teaching pharmacology to baccalaureate nursing students. *CIN -Computers Informatics Nursing*, 29, 326-334.
- Weurlander, M., Söderberg, M., Scheja, M., Hult, H., y Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, 37, 747-760
- White, P., Syncox, D., y Alters, B. (2011). Clicking for grades? Really? Investigating the use of clickers for awarding grade-points in postsecondary education. *Interactive Learning Environments*, 19, 551-561.
- Yourstone, S. A., Krave, H. S., y Albaum, G. (2008). Classroom questioning with immediate electronic response: Do clickers improve learning? *Decision Sciences Journal of Innovative Education*, 6, 75-88.

Estrella Fernández Alba. Profesora Ayudante Doctora. Departamento de Psicología. Área de Psicología Evolutiva y de la Educación. Doctora en Psicología Evolutiva y de la Educación. Sus principales líneas de investigación se centran en la autorregulación del aprendizaje, dificultades de aprendizaje, nuevas tecnologías y variables relacionadas con el rendimiento académico. Entre sus publicaciones cabe destacar la publicación en 2016, junto con otros investigadores, de un artículo titulado “*Comparison of Personal, Social and Academic Variables Related to University Drop-out and Persistence*” en la Revista *Frontiers in Psychology* (IF: 2,463).

Rebeca Cerezo Menéndez. Profesora Contratada Doctora. Departamento de Psicología. Área de Psicología Evolutiva y de la Educación. Doctora en Psicología Evolutiva y de la Educación. Su línea de investigación se centra en la autorregulación del aprendizaje y metacognición, en especial, en contextos virtuales de

aprendizaje. Ha publicado en revistas como *Metacognition and Learning*, pudiendo destacar una de sus últimas aportaciones titulada “*Students’ LMS interaction patterns and their relationship with achievement: A case study in higher education*” publicada en la Revista *Computers and Education* (IF: 2.881).

Marta Méndez. Profesora Contratada Doctora. Departamento de Psicología. Área de Psicobiología. Doctora con Mención Europea en Psicología por la Universidad de Oviedo. Su experiencia docente e investigadora se circunscribe al ámbito de la Psicobiología y la Neurociencia Comportamental. Su investigación se centra en el estudio de las bases biológicas del aprendizaje y la memoria, con especial interés en su evolución a lo largo del desarrollo y su afectación en diversas patologías neurológicas.

Natalia Suárez Fernández. Profesora Ayudante Doctora. Su línea de investigación principal son los deberes escolares y el rendimiento académico (implicación de los padres, feedback de los profesores, motivación del alumnado, autorregulación del aprendizaje, etc). La más reciente ha sido publicada en la prestigiosa revista *Frontiers in Psychology* (IF: 2.463) con el título “*Homework Involvement and Academic Achievement of Native and Immigrant Students*”.

Alejandra Dobarro. Titular de Universidad Interina. Departamento de Psicología. Área de Psicología Evolutiva y de la Educación. ha colaborado en varios proyectos de investigación, entre ellos “Análisis de la convivencia escolar en los centros educativos del Principado de Asturias” y “Factores de riesgo asociados a cibervictimización en la adolescencia.” Coautora de publicaciones relacionadas con la convivencia escolar, como el cuestionario CUVE3 y diversos artículos en revistas como *Anales de Psicología*, *Educación XXI*, *Infancia y Aprendizaje* o *Revista de Psicodidáctica*.

Correspondencia. Estrella Fernández Alba. Universidad de Oviedo. Facultad de Psicología. Departamento de Psicología. Área de Psicología evolutiva y de la Educación. Plaza Feijoo s/n. 33003, Oviedo, Asturias, España. Teléfonos: 985.10.32.80. E-mail: fernandezestrella@uniovi.es