

Explorando la Eficacia de las Prácticas con Wiris-Quizzes en Asignaturas de Matemáticas de Estudios de Ingeniería en Línea

Josep Figuerola-Cañas y Teresa Sancho-Vinuesa

Title—Exploring the efficacy of practices with Wiris-Quizzes in online engineering mathematics .

Abstract— The use of online self-assessed questionnaires Wiris-Quizzes as formative assessment tool is an increasingly widespread practice in a variety of subjects in higher education. Some previous studies have examined the impact of such questionnaires on learning outcomes in classroom environments or hybrid teaching. Our work presents an exploratory study that proves that in a course of Mathematical Analysis in a completely online environment, practices with Wiris-Quizzes improve learning outcomes. We have conducted an ex post facto research from the results of the Practice Tests and Continuous Assessment Tests carrying out an analysis using the trimmed means Yuen's test.

Index Terms— Formative assessment, Online Engineering Mathematics , Learning, Wiris-Quizzes

I. INTRODUCCIÓN

LOS estudiantes de matemáticas de las ingenierías de la Universitat Oberta de Catalunya (UOC), añaden a la falta de motivación y de conocimientos básicos del área características personales específicas de una universidad en línea: responsabilidades familiares y/o profesionales que se traducen en un tiempo reducido de dedicación semanal al estudio, y las dificultades propias del aprendizaje de las matemáticas en línea [1], [2].

En dos asignaturas de esta área se ha implantado una metodología docente basada en la actividad continua del estudiante a través de una herramienta de evaluación y feedback automático, con la finalidad de mejorar la calidad docente [1]: el primer semestre del curso 2010-11, en la asignatura optativa de Iniciación a las Matemáticas del grado de ingeniería informática, y en el primer semestre del curso 2011-12, en la asignatura obligatoria de Análisis Matemático de los grados de ingeniería informática y de ingeniería en telecomunicaciones. En esta última, la implantación se ha llevado a cabo de manera progresiva hasta ser completada en el primer semestre del curso 2015-16, periodo en el cual se ha realizado este trabajo. Más de 10 años de experiencia en este contexto nos condujo a definir una metodología que ayudase al estudiante a realizar una actividad regular durante el curso obteniendo

retroalimentación inmediata. El recurso de aprendizaje que dio respuesta a nuestras necesidades fueron los Wiris-quizzes. Son cuestionarios diseñados en el entorno Moodle, que cuentan con el soporte del programa de cálculo simbólico Wiris y su editor de fórmulas matemáticas. Cuatro de sus características principales justificaron su elección en la fase inicial de proyecto: a) son autoevaluables y proporcionan respuesta automática; b) la respuesta y el feedback son inmediatos; c) sus enunciados tienen parámetros; y d) existe la posibilidad de introducir números y expresiones matemáticas que serán interpretados por el sistema. De esta manera cada vez que se abre un cuestionario, el enunciado es diferente en tanto que los valores de dichos parámetros son distintos, y por lo tanto hay un elevado número de variantes [2]. La posibilidad de introducir expresiones matemáticas de manera ágil reduce el tiempo empleado por los estudiantes en la redacción de las respuestas y favorece que el tipo de preguntas planteadas tengan un grado de complejidad superior. En cuanto a la respuesta del sistema, el estudiante recibe un feedback correctivo [3] mediante la provisión de la respuesta correcta con la resolución detallada del ejercicio, así como una referencia al contenido de los materiales didácticos de la asignatura. Además, se le proporciona la calificación obtenida.

En el presente artículo exploramos las características del aprendizaje de la asignatura de Análisis Matemático del grado de ingeniería informática a través de la realización de cuestionarios con feedback automático. Concretamente, el objetivo principal de este trabajo es determinar, en primera aproximación dado que se trata de un estudio exploratorio, si resulta eficaz la realización de Cuestionarios de Prácticas para superar la evaluación continua de la asignatura Análisis Matemático del grado en ingeniería informática de la UOC. Estudiamos dicha eficacia a través del análisis de los resultados de los estudiantes en los Cuestionarios de Evaluación Continua. Concretamente, pretendemos comprobar la existencia o no de diferencias de resultados entre los estudiantes que optan por la entrega de Cuestionarios de Prácticas, estudiantes *activos*, respecto de aquellos que optan por no hacerlo, estudiantes *no-activos*, tomando como referencia la clasificación de estudiantes *activos* de [5].

Hasta donde llega nuestro conocimiento, el contexto de nuestro estudio, una asignatura de matemáticas en estudios de ingeniería en un entorno completamente en línea, así como el análisis de datos utilizado, basado en contraste de medias truncadas, no han sido empleados en otros trabajos publicados.

Josep Figuerola-Cañas es estudiante de la Universitat Oberta de Catalunya, España. (e-mail: jfiguerola@uoc.edu, ORCID ID: 0000-0002-6790-9142)

Teresa Sancho-Vinuesa es profesora de los Estudios de Informática, Multimedia y telecomunicaciones de la Universitat Oberta de Catalunya, España (e-mail: tsancho@uoc.edu, ORCID ID: 0000-0002-0642-2912)

II. ANTECEDENTES

El uso de cuestionarios en línea con finalidad formativa ha sido empleado en la enseñanza superior en línea en ámbitos como el de matemáticas en ingeniería [2] o en ciencias sociales [6]; pero, también, en enseñanza superior presencial, en metodologías llamadas híbridas, en áreas de conocimiento como matemáticas en estudios de economía [5], [7], [8], matemáticas en ingenierías [9], biología [10] o economía [11].

Diversos son los trabajos de investigación dedicados a determinar la incidencia de los cuestionarios de prácticas en línea sobre el aprendizaje. Estos trabajos coinciden en afirmar que la realización de dichos cuestionarios mejora el aprendizaje. Existen diferencias entre ellos, tanto en la forma de elegir la variable encargada de cuantificar el aprendizaje, como en la metodología empleada para llevar a cabo la investigación, principalmente en cómo efectúan el análisis de datos. A continuación presentaremos sus principales características.

Por lo que se refiere a la variable que mide el aprendizaje, la calificación del examen final es empleada como variable dependiente en [5], [7], [8], [9], [10]. El trabajo de [11], en cambio, utiliza la calificación de exámenes parciales correspondientes a unidades didácticas y el de [2], la nota de la asignatura en su conjunto.

En cuanto a la metodología y resultados, la referencia [11] implementa un test de diferencias entre medias en las calificaciones de exámenes de unidades didácticas entre dos grupos de estudiantes, según si han realizado cuestionarios de prácticas o no. La pertenencia a los grupos es decidida por los estudiantes, tratándose por tanto de un diseño no-experimental. Establece que la realización de los cuestionarios de prácticas mejora el aprendizaje, en tanto que la media de los cuestionarios de evaluación de los estudiantes que realizan cuestionarios de prácticas es superior, con una significación estadística inferior al 5%, a la de los estudiantes que optan por no realizarlos. La referencia [6] también utiliza un test de diferencias entre medias, en su caso en las calificaciones del examen final, entre dos grupos de estudiantes. En este caso, se trata de un diseño cuasi-experimental con un *grupo de control* formado por estudiantes sin la posibilidad de realizar cuestionarios de prácticas, y un segundo grupo, *grupo de experimentación*, formado por estudiantes obligados a realizar cuestionarios de prácticas. El estudio concluye la mejora del aprendizaje ante la realización de cuestionarios de prácticas, teniendo en cuenta que la media del examen final de la asignatura es significativamente superior en los estudiantes de un grupo obligados a realizar cuestionarios de prácticas, grupo de *experimentación*, respecto a la del grupo que no disponía de cuestionarios, grupo de *control*. En la referencia [10], el análisis de datos consiste en comparar la media de los exámenes finales entre promociones que no disponían de cuestionarios de prácticas y aquellas que sí lo hacían. La conclusión para [10] difiere según la voluntariedad en la realización de cuestionarios de prácticas. Para una promoción con voluntariedad absoluta, no se aprecia mejora significativa en el aprendizaje, mientras que en otra promoción en que los cuestionarios de prácticas representan el 20% de la calificación total de la asignatura, se observa que la media del examen final de los estudiantes es superior a la de promociones sin el uso de cuestionarios. Un diseño *expost facto*, junto a una prueba de análisis de varianzas (ANOVA) es utilizada por [9], quien llega a la misma

conclusión que [10], en el sentido que únicamente al realizar cuestionarios con repercusión directa sobre la calificación final se aprecia mejora de aprendizaje, no así en el caso que los cuestionarios no contribuyen a la calificación final. La referencia [9] demuestra que los estudiantes que entregan mayor número de cuestionarios obtienen mejores calificaciones en el examen final. Mediante un estudio correlacional y una regresión logística, la referencia [7] prueba la mejora del aprendizaje, estableciendo, que existe asociación entre realizar cuestionarios de prácticas y la nota en el examen final. Otro estudio correlacional, en este caso mediante una regresión de cresta [8] concluye la eficacia de los cuestionarios, al probar la existencia de asociación positiva entre el número de cuestionarios realizados y la nota del examen final. Un diseño *expost facto* y un test de diferencia de medias [5] concluye la mejora del aprendizaje en los estudiantes que realizan cuestionarios de prácticas, estudiantes *activos*, observando que las medias de los exámenes finales son superiores respecto a sus compañeros *no-activos*

III. CONTEXTO EDUCACIONAL Y MÉTODO PEDAGÓGICO

La Universitat Oberta de Catalunya, UOC, establece de forma obligatoria la realización de un examen final presencial en todas las asignaturas de matemáticas del grado de ingeniería informática, así como la posibilidad, recomendada por el conjunto de profesores a los estudiantes, de participar, en la Evaluación Continua. Cada asignatura determina cómo concreta la Evaluación Continua, cuya finalidad es tanto formativa como sumativa. El sistema de calificación de la asignatura bajo estudio integra de forma ponderada la calificación de la Evaluación Continua, EC, y la calificación del Examen Final presencial, EF. Así, la nota final de la asignatura, NF, se obtiene según (1)

$$NF = \max(EF; 0,65 \cdot EF + 0,35 \cdot EC) \quad (1)$$

Este sistema garantiza el carácter de voluntariedad en la Evaluación Continua para aquellos estudiantes que, o bien optan por ser evaluados exclusivamente mediante el Examen Final (EC=0), o bien, para aquellos cuya calificación de la Evaluación Continua, obtenida a lo largo del proceso de aprendizaje, supondría empeorar la calificación del Examen Final, mediante su incorporación en la media ponderada con un peso del 35%. En definitiva, la Evaluación Continua únicamente constituirá un elemento de la evaluación sumativa, al contribuir cuantitativamente a la calificación final [4], siempre que favorezca al estudiante, mientras que en todo momento mantiene su carácter formativo, dado que la información que se le facilita al estudiante pretende contribuir a su aprendizaje.

La Evaluación Continua consta de siete Cuestionarios de Evaluación Continua (CEC), en donde los estudiantes disponen de un único intento para responder. Los cinco primeros, planificados para ser realizados semanalmente, se centran en aspectos básicos de Análisis Matemático, buena parte de los cuales ya han sido estudiados con anterioridad en enseñanzas medias o en la asignatura Iniciación a las Matemáticas para la Ingeniería del propio grado. Se trata de cuestionarios Wiris-quizzes de seis preguntas, donde las cinco primeras son del tipo elección múltiple o bien, respuesta corta. La corrección se efectúa automáticamente y de forma instantánea, ofreciendo al estudiante feedback

inmediato. La sexta pregunta pertenece al tipo de respuesta abierta en que el estudiante debe responder de manera razonada. La corrección corresponde al profesor y el tiempo de respuesta puede llegar a ser de hasta una semana. Los dos últimos Cuestionarios de Evaluación Continua, programados para ser realizados mensualmente, se centran en aspectos avanzados de Análisis Matemático, no trabajados anteriormente en otras asignaturas.

Previamente al inicio de cada Cuestionario de Evaluación Continua, los estudiantes disponen de la posibilidad de responder uno o más Cuestionarios de Prácticas (CP), formados por 10 preguntas parametrizadas del tipo elección múltiple o respuesta corta, véase como ejemplo la mostrada en la Fig.1. Una vez enviadas las respuestas, como en el caso de la Fig.2 mediante la introducción de expresiones matemáticas en el programa Wiris, el sistema proporciona automática e inmediatamente un feedback correctivo informando de la respuesta correcta, ofreciendo una ayuda para su comprensión (Fig.3), valorando numérica la respuesta (Fig.1), así como el conjunto de las diez. La finalidad de estos cuestionarios es familiarizar al estudiante en este tipo de actividad y permitirles total flexibilidad en su realización. El número de modelos de Cuestionarios de Prácticas difiere según el Cuestionario de Evaluación Continua.

IV MÉTODO

La aproximación metodológica de este trabajo es cuantitativa. Se trata de un diseño *ex post facto* retrospectivo, en tanto que la entrega de Cuestionarios de Prácticas es una decisión voluntaria de cada estudiante, sobre la que los autores de la investigación no han tenido ninguna capacidad de control. La fuente de datos se encuentra en el registro de actividades del Moodle de la asignatura ya finalizada que contiene, principalmente, las calificaciones de Cuestionarios de Prácticas y de Cuestionarios de Evaluación Continua. En cuanto al tipo de análisis, se ha efectuado un análisis mediante contraste de medias truncadas.

A. Participantes

Participan en este estudio los 131 estudiantes de la asignatura de análisis matemático del primer semestre del curso 2015-16 del grado en ingeniería informática de la UOC. No se ha contemplado ningún diseño muestral, dado que el estudio incluye toda la población de estudiantes que entregan Cuestionarios de Evaluación Continua.

B. Instrumentos

Las calificaciones de los Cuestionarios de Prácticas y de los Cuestionarios de Evaluación Continua constituyen la principal fuente de datos del presente estudio. El registro de actividades del Moodle de la asignatura de análisis

Pregunta 5	Determina la recta tangente a la función $f(x) = -x^3 - 3x^2 - 3x - 1$ en $x = 0$.
Correcta	
Puntua 1,00	Escribe la respuesta de la forma $m \cdot x + n$ donde m es la pendiente de la recta y n su ordenada al origen. Es decir, de manera que la ecuación explícita de la recta sea $y = m \cdot x + n$
sobre 1,00	

Fig. 1. Ejemplo de pregunta de respuesta corta en un Cuestionario de Prácticas y de valoración numérica

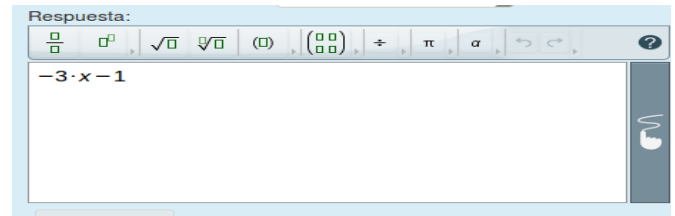


Fig. 2. Ejemplo de respuesta introducida mediante el programa Wiris

matemático permite disponer de las calificaciones globales y separadas por preguntas.

C. Procedimiento

Con el fin de mantener la confidencialidad de los datos, las referencias personales han sido codificadas convenientemente. Además, se ha efectuado un filtrado de los datos con el programa Calc de Libreoffice, de manera que de cada Cuestionario de Evaluación Continua se conserva el código de cada estudiante y sus calificaciones (por pregunta y total). La información que se almacena de los Cuestionarios de Prácticas es el código de estudiante y la calificación total de cada entrega. Los 17 archivos, correspondientes al total de cuestionarios, una vez efectuado el filtrado descrito anteriormente, son la base para tratamiento estadístico posterior.

D. Preprocesamiento de Datos

De acuerdo con [5], para cada Cuestionario de Evaluación continua separamos los estudiantes en dos grupos según su nivel de actividad: estudiantes *activos* y estudiantes *no-activos*. Consideramos un estudiante activo en un Cuestionario de Evaluación Continua si ha demostrado un comportamiento activo en la realización de los Cuestionarios de Prácticas asociados a dicho cuestionario, que se evidencia con su envío y posterior registro en la plataforma Moodle de la asignatura. La pertenencia a cada grupo es decidida por el propio estudiante y puede ser distinta entre diferentes Cuestionarios. En el caso de Cuestionarios de Evaluación Continua con un único modelo de Cuestionario de Prácticas asociado, entendemos por comportamiento activo el hecho de haber entregado al menos uno. En el caso del Cuestionarios de Evaluación con más de un modelo de Cuestionario de Prácticas asociado, entendemos comportamiento activo haber entregado al menos un Cuestionario de Prácticas de cada uno de los distintos modelos de Cuestionarios de Prácticas asociados.

En los cinco primeros Cuestionarios de Evaluación Continua, las variables dependientes estudiadas son tres: la calificación global, la calificación del total de preguntas de

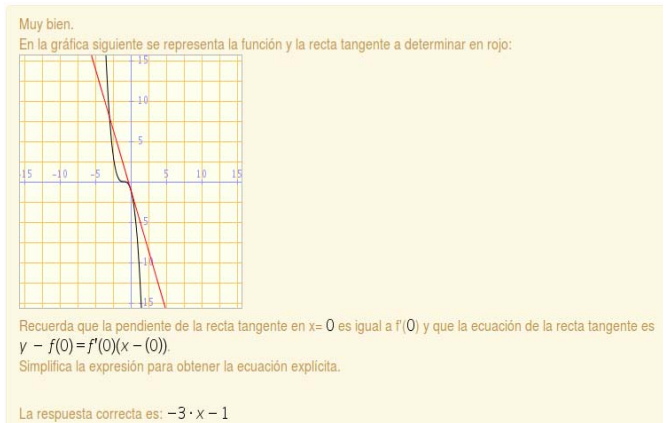


Fig.3 Ejemplo de respuesta correcta y ayuda para su comprensión.

corrección automática y la calificación de la pregunta de exposición razonada. En los dos últimos Cuestionarios de Evaluación Continua, la variable dependiente estudiada es, únicamente, la calificación global.

E. Análisis de datos

Los diagramas de caja y bigotes de las variables definidas para cada uno de los siete Cuestionarios de Evaluación Continua muestran asimetría. Sirva como ejemplo el mostrado en Fig.4. Aplicando el test de Brown-Forsythe no podemos suponer la homogeneidad de varianzas entre los dos grupos, por otro lado, hecho bastante frecuente en estudios del ámbito de la educación. El uso de medias truncadas para realizar comparaciones entre grupos, y en particular la prueba de Yuen, es sugerido por [12] para conseguir robustez ante la existencia de no-normalidad y heteroscedasticidad como en la situación en que nos encontramos. Diversos autores citados por [12] consideran distintos porcentajes de truncamiento que oscilan entre el 10% y el 25%, tanto para valores superiores como inferiores. En nuestro estudio consideraremos un 25% de truncamiento. De esta forma, estamos calculando la media de los datos que constituyen el rango intercuartílico, es decir, la media del 50% de las calificaciones centrales. Para la prueba de Yuen establecemos un nivel de confianza del 95%.

La hipótesis nula considera que las medias truncadas son iguales para el grupo de estudiantes activos que para el de no-activos, y la hipótesis alternativa, que las medias truncadas son diferentes para los dos grupos. Los cálculos se han realizado con el programa R mediante el uso de la función *yuen()*, que efectúa la prueba de Yuen para determinar la existencia de diferencias significativas entre dos grupos; y de la función *yuen.effect.ci()*, que proporciona la medida explicatoria del tamaño del efecto, ξ [13]. Si bien, dado que no se trata de un diseño ni experimental ni cuasiexperimental, por lo que no se puede establecer ninguna relación *causa-efecto*, nos parece oportuno mostrar el tamaño del efecto como información complementaria.

V. RESULTADOS

En general, para cada Cuestionario de Evaluación Continua, los estudiantes activos obtienen, mejores resultados, en la calificación global y en la calificación de las preguntas de corrección automática, pero no en las preguntas de exposición razonada.

A. Los Estudiantes Activos Obtienen Calificaciones Globales Superiores

La media truncada de las calificaciones globales de los estudiantes activos es significativamente superior, según los resultados mostrados en la Tabla I, en cuatro de los siete Cuestionarios de Evaluación Continua, dado que el p-valor es inferior a 0,05. De acuerdo con ello, podemos afirmar que las calificaciones globales de los estudiantes activos son superiores, en más del 50% (4 sobre 7) de los Cuestionarios de Evaluación Continua. En tres de los Cuestionarios en que se observa diferencia, ésta es superior a 1 punto sobre 10, y en el número 6 la diferencia supera los 2 puntos. El tamaño del efecto en todos los casos se mueve en valores próximos a 0,5, que es considerado como alto en una escala de bajo, medio, alto [13]. Se observa, además, que es en los Cuestionarios extremos (números 1, 2 y 7) donde no se

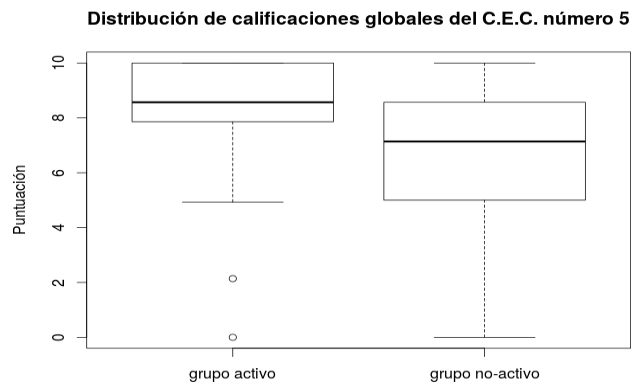


Fig. 4. Diagrama de caja y bigotes de la distribución de calificaciones globales del Cuestionario de Evaluación Continua número 5.

TABLA I
RESULTADOS DE LA PRUEBA DE YUEN (A)

CEC	Diferencia de Medias Truncadas al 25% (B)	p-valor	Tamaño del efecto
Núm. 1	+0,30	0,2721	-----
Núm. 2	+0,29	0,4282	-----
Núm. 3	+1,64 (0,92 ; 2,35)	0,0000*	0,53
Núm. 4	+0,84 (0,03 ; 1,64)	0,0430	0,44
Núm. 5	+1,95 (0,98 ; 2,92)	0,0003	0,56
Núm. 6	+2,08 (0,63 ; 3,53)	0,0057	0,53
Núm. 7	+0,58	0,1690	-----

* inferior a 0,0001

(A) Sobre las **calificaciones globales** de los Cuestionarios de Evaluación Continua (CEC)

(B) Diferencia entre estudiantes activos respecto no-activos. Entre paréntesis, el intervalo de confianza de la diferencia de medias truncadas al 95% de nivel de confianza.

Hipótesis alternativa: la media truncada es diferente en el grupo de estudiantes activos

aprecian diferencias significativas en las calificaciones globales.

B. Los Estudiantes Activos Obtienen Mejores Resultados en las Preguntas de Corrección Automática.

Únicamente en aquellos Cuestionarios de Evaluación Continua en que la media truncada de la calificación global de los estudiantes activos es superior, la media truncada de la calificación de las preguntas de corrección automática de los estudiantes activos es superior. Ello continúa suponiendo que en más del 50% de los Cuestionarios (3 sobre 5) los resultados son mejores para los estudiantes activos. La Tabla II muestra que en los Cuestionarios 3, 4 y 5 existe diferencia en la calificación de las preguntas de corrección automática, tal como mostraba la Tabla I para la calificación global. En

TABLA II
RESULTADOS DE LA PRUEBA DE YUEN (A)

CEC	Diferencia de Medias Truncadas al 25% (B)	p-valor	Tamaño del efecto
Núm. 1	+0,21	0,5124	-----
Núm. 2	+0,17	0,6482	-----
Núm. 3	+2,11 (1,14 ; 3,07)	0,0000*	0,66
Núm. 4	+1,30 (0,55 ; 2,05)	0,0011	0,49
Núm. 5	+2,72 (1,53 ; 3,92)	0,0000*	0,61

* inferior a 0,0001

(A) Sobre las **calificaciones de las preguntas de corrección automática** de los Cuestionarios de Evaluación Continua (CEC)

(B) Diferencia entre estudiantes activos respecto no-activos. Entre paréntesis, el intervalo de confianza de la diferencia de medias truncadas al 95% de nivel de confianza

Hipótesis alternativa: la media truncada es diferente en el grupo de estudiantes activos

los tres casos el incremento de la media en las preguntas de corrección automática es superior al mostrado en el global. El más destacado corresponde al número 5 con una diferencia de 2,72 puntos sobre 10 (Tabla II) en las preguntas de corrección automática, que también mostraba una diferencia máxima de 2,08 en el global (Tabla I). Por lo que se refiere al tamaño del efecto, los valores son algo superiores a los obtenidos en las calificaciones globales, por lo que continúan siendo, igualmente altos.

C. Los Estudiantes Activos no Obtienen Mejores Resultados en las Preguntas de Exposición Razonada.

En ninguno de los cinco Cuestionarios de Evaluación Continua con calificación de preguntas de exposición razonada se observa diferencia entre los estudiantes activos y los no-activos (Tabla III). En tres Cuestionarios la diferencia entre las medias truncadas no supera los 0,15 puntos sobre 10, e incluso en el número 4 la diferencia es 0. Esta unanimidad entre los Cuestionarios contrasta con lo expuesto en los apartados anteriores.

D. El Número de Estudiantes Activos se Reduce Notablemente en la Fase Final del Curso.

En la Tabla IV se observa que el número de estudiantes que entregan todos los modelos de Cuestionarios de Prácticas, estudiantes activos, es superior a los no-activos en los Cuestionarios de Evaluación Continua números 2, 4 y 5. Aquellos con un único modelo y de frecuencia semanal. En los números 1 y 3, con más de un modelo y, también, de frecuencia semanal, se invierte la dominancia, pasando a ser el número de estudiantes activos inferior al de no-activos. En los dos últimos Cuestionarios de Evaluación Continua la proporción de estudiantes no-activos respecto a activos resulta cercana a 3:1.

VI. DISCUSIÓN

La participación de los estudiantes, calculada sobre el total de estudiantes matriculados, en la entrega de Cuestionarios de Prácticas, correspondientes a los cinco primeros Cuestionarios de Evaluación Continua, presenta unos valores entre el 49% y el 54%, en los de modelo único, y entre 34% y 41%, en los de modelo múltiple. Recordemos que estos últimos valores corresponden a estudiantes que han entregado, como mínimo, un Cuestionario de Práctica de cada modelo. Todas estas cifras se encuentran entre el 33% de participación del estudio presentado por [10] y el 68% del trabajo de [11]. En cambio, en los dos últimos CEC, el porcentaje cae hasta el 20%.

Se pueden apuntar diversas posibles explicaciones para la alta participación en los cinco primeros CEC. En primer

TABLA III
RESULTADOS DE LA PRUEBA DE YUEN (A)

CEC	Diferencia de Medias Truncadas al 25% (B)	p-valor	Tamaño del efecto
Núm. 1	+0,14	0,3096	-----
Núm. 2	+0,75	0,3017	-----
Núm. 3	+0,31	0,4656	-----
Núm. 4	0	-----	-----
Núm. 5	+0,07	0,8815	-----

* inferior a 0,0001

(A) Sobre las calificaciones de las preguntas de respuesta razonada de los Cuestionarios de Evaluación Continua (CEC)

(B) Diferencia entre estudiantes activos respecto no-activos.

Hipótesis alternativa: la media truncada es diferente en el grupo de estudiantes activos

TABLA IV
NÚMERO DE ESTUDIANTES SEGÚN LA ACTIVIDAD

Cuestionarios de Evaluación Continua	Estudiantes activos	Estudiantes no-activos
Número 1	44	69
Número 2	71	43
Número 3	53	61
Número 4	64	42
Número 5	65	46
Número 6	27	74
Número 7	25	69

lugar, el nivel de conocimiento, básico tal como indica el propio plan docente, favorece que el estudiante se sienta capaz de responder a las preguntas que se le plantean. En segundo lugar, el hecho que la frecuencia de CEC sea semanal hace que ayude al estudiante a ser regular en su dedicación. Y en tercer lugar, dado el estudiante dispone de dos días para responder el CEC, prefiere practicar con los CP y obtener *feedback*, que le permita reducir la distancia entre su nivel de conocimiento y el exigido para superar el CEC, encarando así con más garantías de éxito el CEC, que en definitiva repercute de forma directa en su calificación final. La reducción en la participación disminuye en los casos de múltiples CP asociados a CEC (números 1 y 3). Ello puede ser debido al aumento de carga de trabajo semanal que les comporta entregar más de un Cuestionario de Práctica distinto.

El descenso de participación en dos últimos Cuestionarios de Evaluación Continua podría atender a diversas explicaciones. En primer lugar, el nivel avanzado de los contenidos y/o la dificultad de las preguntas. Se constata un aumento muy significativo del número de estudiantes que no entregan el CP, a pesar de haberlo descargado, lo que puede ser debido a la dificultad que encuentran los estudiantes a la hora de resolverlos. En segundo lugar, el aumento del tiempo permitido para la realización del CEC, que pasa a ser de hasta dos semanas, puede hacer que algunos estudiantes consideren que es tiempo suficiente para preparar y revisar el CEC con garantías, y prefieran concentrar su esfuerzo únicamente en la realización de ese tipo de cuestionarios, dado que son los únicos que tienen incidencia directa sobre la calificación final de la asignatura. En los trabajos de [5] y [9] a la entrega de cuestionarios en línea corresponde un porcentaje del 10% y 25%, respectivamente, en la calificación final de la asignatura. Finalmente, es posible, también, que los estudiantes no contrapesen suficientemente el beneficio del posible *feedback* que recibirían ante el coste que les supone entregar el CP.

Desde el punto de vista de los posibles beneficios que obtienen los estudiantes activos gracias a la entrega de CP podríamos indicar que el *feedback* que reciben aporta conocimiento de posibles preguntas similares en contenido y tipología. Ello podría incidir en la asociación observada entre calificaciones globales de CEC y la entrega CP asociados. La referencia [9] prueba que en las preguntas del examen final, el contenido del cual había sido cubierto previamente en los cuestionarios en línea, se obtenían calificaciones más elevadas. El hecho de que los resultados en cuanto a la asociación de calificaciones globales se extiendan a la de asociación de calificaciones de preguntas de corrección automática puede ser debido a la tipología de los CP, constituidos únicamente por ese tipo de preguntas. Esto último podría ser también el motivo por el cual no se aprecia asociación entre la realización de CP y las

calificaciones de las preguntas de respuesta razonada. Así, tal como están planteados los CP, su realización no parece que ayude a establecer diferencias entre estudiantes en cuanto a exposición razonada de contenidos matemáticos.

Otras justificaciones alternativas en relación a la realización de Cuestionarios de Prácticas podrían ser que los estudiantes activos provengan de un nivel de conocimiento previo superior y/o que la motivación sea un elemento diferencial respecto al resto de estudiantes. La referencia [5], sin embargo, descarta la motivación como factor explicativo. Dado que se aprecia diferencia según el tipo de preguntas (de corrección automática y de exposición razonada) pensamos que no se pueden atribuir exclusivamente al nivel de estudios previos las diferencias observadas en las calificaciones globales entre estudiantes activos y no-activos. De un modo u otro pues, la realización del CP influye en la existencia de esas diferencias.

En los CEC del inicio y del final del curso no se establecen las mismas asociaciones que en los CEC del medio. Podría ser que en el último, los estudiantes con buenos resultados previos ya piensen más en concentrar esfuerzos en la evaluación y renuncien a la práctica. Por lo que respecta al segundo cuestionario de evaluación podría ser que la similitud entre las preguntas de su CP sea inferior a la del resto, que las preguntas tengan un nivel de dificultad superior y/o que el feedback proporcionado no haya resultado de tanta utilidad como en otros CP. Finalmente, el CEC número 1 dispone de tres modelos de CP entregados en una sola semana, lo que hace más complejo las interacciones entre ellos y las calificaciones globales.

VII. CONCLUSIONES

De acuerdo con el objetivo del presente trabajo, podemos concluir, en primera aproximación, que el aprendizaje adquirido por los estudiantes que entregan Cuestionarios de Práctica tipo Wiris-quizzes es superior al de los estudiantes que optan por no entregarlos. Nuestra conclusión establece una asociación, estadísticamente significativa mediante la prueba de Yuen, entre calificaciones de Cuestionarios de Evaluación Continua como indicador de aprendizaje y la entrega de Cuestionarios de Prácticas. Esta misma relación coincide con la presentada por [11] en un entorno presencial y tan solo con un Cuestionario de Prácticas y uno de Evaluación continuada. En nuestro caso se ha comprobado en cuatro de los siete Cuestionarios de Evaluación Continua. Con calificaciones de examen final como indicador de aprendizaje, [5] y [9] en un contexto de enseñanza presencial complementada con la utilización de cuestionarios en línea también concluyen, igualmente, que existen diferencias, estadísticamente significativas, entre el aprendizaje de los estudiantes que entregan Cuestionarios de Prácticas y aquellos que no lo hacen. La mejora del aprendizaje se constata en las preguntas de respuesta corta o elección múltiple, a las que corresponde una corrección automática. En cambio, el análisis específico de la pregunta de exposición razonada de los Cuestionarios de Evaluación Continua nos indica que la entrega de Cuestionarios de Prácticas no se traduce en una diferencia significativa en la calificación de este tipo de preguntas.

La realización de Cuestionarios de Prácticas es recomendable porque permite mantener una constancia regular en el trabajo, y sobre todo porque permite disponer de feedback que facilita el proceso de autoregulación del estudiante, de gran importancia en estudios en línea. Por

contra, no resulta suficiente para mejorar la capacidad de expresar razonamientos matemáticos.

El presente trabajo contribuye a ampliar estudios previos en entornos presenciales a entornos completamente en línea en el ámbito de la enseñanza de matemáticas superiores. Además, utiliza la comparación entre grupos mediante el contraste de medias truncadas como estrategia analítica de los datos. El diseño *ex post facto* retrospectivo lleva consigo limitaciones en las conclusiones de relaciones causa-efecto, que en cierto modo pueden ser corregidas siempre que se introduzcan mecanismos que controlen las variables de confusión, como por ejemplo estableciendo subgrupos homogéneos [12]. En este sentido, y con el propósito de profundizar en la problemática planteada, se podrían introducir nuevas variables que recojan aspectos emocionales, como la motivación; o variables que informen sobre uso efectivo del feedback recibido, o que establezcan niveles de dificultades de los cuestionarios. Además, resultaría de interés ampliar la comparación de resultados a las calificaciones de los exámenes finales agrupando a los estudiantes según su comportamiento respecto los Cuestionarios de Prácticas y/o Cuestionarios de Evaluación Continua. También sería interesante avanzar hacia diseños correlacionales que permitan establecer asociaciones más estrechas entre variables como los planteados por [7] y hacia diseños cuasiexperimentales que permitan definir grupos equivalentes en cuanto a conocimientos de la materia con el objetivo que las diferencias encontradas sean debidas exclusivamente a la intervención: realización de Cuestionarios de Práctica. En nuestro caso, no podemos asegurar la equivalencia entre los dos grupos de estudiantes, dado que éstos se han autoseleccionado. Por último, debe destacarse la relevancia del análisis que aquí se presenta como prueba de la eficacia de la práctica continuada en el aprendizaje de matemáticas a través de una herramienta de evaluación automática. Éste podría ser un elemento clave en nuevas propuestas de metodologías docentes mixtas que incluyan la presencialidad y la virtualidad.

AGRADECIMIENTOS

Esta investigación ha sido financiada por el Gobierno de España a través del proyecto “ICT-FLAG” Enhancing ICT education through Formative assessment, Learning Analytics and Gamification (TIN2013-45303-P).

REFERENCIAS

- [1] T. Sancho-Vinuesa y N. Escudero Viladoms, “¿Por qué una propuesta de evaluación formativa con feedback automático en una asignatura de matemáticas en línea?,” *Revista Universidad y Sociedad del Conocimiento*, vol. 9, n.º 2, pp. 59–79, 2012.
- [2] R. Calm, R. Masià, C. Olivé, N. Parés, F. Pozo, J. Ripoll, y T. Sancho-Vinuesa, “Wiris Quizzes: un sistema de evaluación continua con feedback automático para el aprendizaje de matemáticas en línea,” *Teoría de la Educación; Educación y Cultura en la Sociedad de la Información*, vol. 14, n.º 14, pp. 452–472, 2013.
- [3] J. Hattie y H. Timperley, “The power of feedback,” *Review of Educational Research*, vol. 77, n.º 1, pp. 81–112, 2007.
- [4] M. Yorke, “Formative Assessment in Higher Education: Moves Towards Theory and the Enhancement of Pedagogic Practice,” *Higher Education*, vol. 45, n.º 4, pp. 477–501, 2003.
- [5] R. Huisman y H. E. Reedijk, “The Impact of Individual Online Tests in Addition to Group Assignments on Student Learning,” *Iciete 2012*, pp. 1–16, 2012.
- [6] B. M. Klecker, “The Impact of Formative Feedback on Student Learning in an Online Classroom,” *Journal of Instructional Psychology*, vol. 34, no. 3, pp. 161–165, 2007.
- [7] S. D. Angus y J. Watson, “Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a

- large data set,” *British Journal of Educational Technology*, vol. 40, n.º. 2, pp. 255–272, 2009.
- [8] E. Fitkov-Norris y B. Lees, “Online formative assessment: Does it add up to better performance in Quantitative modules?,” en *Proceedings of the 11th ECRM*, 2012, pp. 115–121
- [9] L. Chirwa, “A case study on the impact of automated assessment in engineering mathematics,” *Engineering Education.*, vol. 3, n.º. 1, pp. 13–20, 2008.
- [10] S. Voelkel, “Combining the formative with the summative: The development of a two stage online test to encourage engagement and provide personal feedback in large classes,” *Research in Learning Technology*, vol. 21, n.º. 1063519, pp. 1–18, 2013.
- [11] L. Sly, “Practice Tests as Formative Assessment Improve Student Performance on Computer Managed Learning Assessments,” *Assessment & Evaluation in Higher Education*, vol. 24, no. 3, pp. 339–343, 1999.
- [12] H.J. Keselman, A.R. Othman, R.R. Wilcox, y K. Fradette, “The new and improved two-sample T test”, *Psychological Science*, vol 15, no. 1, pp. 47-51, 2004.
- [13] R.R. Wilcox y T.S. Tian, “Measuring effect size: a robust heteroscedastic approach for two or more groups”, *Journal of Applied Statistics*, vol 38, no. 7, pp. 1359-1368, 2011.
- [14] Brewer, Khun, y N. J. Salkind, “Encyclopedia of Research Design Volume 1,” in *SAGE Publications*, 2010, pp. 124–130.

Josep Figuerola-Cañas, nacido en Manresa (España), es estudiante de doctorado de Educación y TIC de la Universitat Oberta de Catalunya y licenciado en ciencias físicas por la Universitat Autònoma de Barcelona (España, 1988).

Teresa Sancho-Vinuesa, nacida en Barcelona (España), es doctora en ingeniería electrónica por la universitat Ramon Llull (España, 1995) y licenciada en matemáticas por la Universitat de Barcelona (España, 1990).

Actualmente es profesora agregada de la Universitat Oberta de Catalunya (España) y dirige el grupo de investigación Learning Analytics for Innovation and Knowledge Application in Higher Education (LAIKA) y ha sido investigadora visitante en la Open UK. Durante el período 1990-1996, ha sido docente e investigadora en la Escuela de Ingeniería y Arquitectura, la Salle (España). Ha sido miembro del equipo pedagógico y editorial de TEXT de la Enciclopedia catalana (España). En la Universitat Oberta ha ocupado diversos cargos de responsabilidad: directora del programa de doctorado en Sociedad de la información y el conocimiento, Directora de investigación y vicerrectora de investigación e innovación. Aunque desde su ingreso en la Oberta de Catalunya ha trabajado en temas relacionados con la educación superior e Intenet, actualmente concentra su actividad investigadora en el uso de learning analytics para la mejora de la enseñanza y aprendizaje en línea, en particular, en los procesos de evaluación y feedback.

La profesora Sancho-Vinuesa ha participado en más de 10 comités técnicos y científicos y es revisora de distintas revistas académicas de reconocido prestigio en el ámbito educativo. Ha participado en las dos redes internacionales REMIC y JEM.