

Rivalizar o no: análisis del modo competición de Wooclap basado en rendimiento y procesamiento de audio

Óscar Cánovas Reverte, Pilar González Férez
Departamento de Ingeniería y Tecnología de Computadores
Universidad de Murcia
30100 Murcia
ocanovas@um.es, pilargf@um.es

Resumen

Este artículo presenta los resultados de aplicar, en una asignatura del Grado en Ingeniería Informática, distintos modos de uso para un sistema interactivo de respuesta. En el estudio se ha usado la plataforma Wooclap, la cual permite lanzar preguntas a los estudiantes para que las contesten en el aula mediante sus dispositivos móviles. Se ha diseñado una experiencia en la que se analizan dos usos de la plataforma: sin competición y con competición basada en puntos. Para ello, se han empleado dos grupos disjuntos de estudiantes que han respondido, en pruebas distintas guiadas por el mismo docente, a las mismas preguntas pero empleando modos de uso distintos en cada caso. Esta investigación cuasi-experimental toma como fuentes de datos el rendimiento de los estudiantes en dichas pruebas (preguntas acertadas correctamente) y el nivel de interacción general en el aula durante el desarrollo de las mismas. Hemos empleado un sistema de análisis de las grabaciones de audio efectuadas, basado en el uso de inteligencia artificial, para caracterizar la participación de los distintos actores. Los resultados obtenidos muestran que, dependiendo de la modalidad, hay patrones distintos en el nivel de interacción de los estudiantes durante las pruebas. Sin embargo, los resultados también indican que la modalidad usada no influye en el rendimiento obtenido por los estudiantes.

Abstract

This paper presents the results of applying, in a subject of the Degree in Computer Science, different modes of use for an interactive response system. The study used the Wooclap platform, which allows the teacher to prompt questions to the students to be answered in class using their mobile devices. An experience has been designed in which two different uses of the platform were analyzed: without competition and with point-based competition. Two separate groups of stu-

dents answered, in different tests guided by the same teacher, the same questions but using different modes of use in each case. This quasi-experimental research takes as data sources the performance of the students in these tests (questions correctly answered) and the general level of interaction in the classroom. We have used an AI-based audio recording analysis system to characterize the participation of different actors. The results obtained show that, depending on the mode, there are different patterns in the level of interaction of the students during the tests. However, the results also indicate that the mode used does not influence the performance obtained by the students.

Palabras clave

Sistema interactivo de respuesta, Wooclap, competición, investigación cuasi-experimental, análisis de grabaciones de audio, inteligencia artificial.

1. Introducción

Los sistemas interactivos de respuesta (también denominados de respuesta de audiencia o, simplemente, cuestionarios interactivos) son herramientas de uso habitual en las aulas en todos los niveles educativos, incluyendo la educación superior [4]. Además, estas herramientas pueden emplearse de formas muy diversas en función del enfoque pedagógico seguido [2].

Entre las distintas modalidades de uso, algunas plataformas proporcionan una experiencia de usuario muy cercana al concurso. Se pone especial énfasis en aspectos de gamificación como las puntuaciones, tablas de clasificaciones o una música y estética acordes. La percepción subjetiva de los estudiantes difiere en función de si la experiencia es más competitiva o menos [14], lo que quizá pueda conllevar a un rendimiento distinto por parte de los estudiantes.

Este artículo presenta una experiencia docente basada en el uso de Wooclap en una asignatura del Grado

en Ingeniería Informática. En él se analiza la influencia que puede tener usar o no el modo competición en el rendimiento de los estudiantes durante las pruebas realizadas en el aula. Además se indaga acerca del nivel de interacción que genera cada modo durante la dinámica. Teniendo todo ello en consideración, estas son las dos preguntas de investigación planteadas:

Pregunta 1 ¿Existen diferencias significativas en las puntuaciones obtenidas al utilizar el modo competición en Wooclap frente a no usarlo?

Pregunta 2 ¿Existen diferencias en los patrones de interacción de los estudiantes en el aula dependiendo del modo de uso de Wooclap?

En relación con la dinámica generada en el aula, muchos resultados de investigaciones educativas han demostrado que la interacción exitosa de los estudiantes en un aula, es decir, compartir ideas con el profesor o participar en discusiones, se correlaciona con un aprendizaje de alta calidad [11]. Las interacciones entre profesores y alumnos puede medirse a través de métodos de observación estandarizados que proporcionen a los profesores datos sobre características relevantes de las interacciones en el aula. En nuestro caso analizamos los audios de las grabaciones realizadas en el aula durante el desarrollo de las pruebas que son objeto de estudio, utilizando para ello un sistema de análisis basado en técnicas de aprendizaje computacional.

La principal contribución de este trabajo es analizar si el rendimiento de los estudiantes al usar un sistema interactivo está influenciado por el uso o no del modo competición de la herramienta, midiendo para ello los resultados obtenidos por los alumnos al realizar las pruebas. Otra contribución importante es la medición del nivel de interacción de los estudiantes durante las pruebas en función del modo de uso.

Los resultados obtenidos en esta experiencia docente con estudiantes de Ingeniería en Informática parecen indicar que el modo competición no produce puntuaciones significativamente distintas respecto al modo sin competición, pero sí genera mayor interacción en el aula durante el desarrollo de las pruebas.

El resto del artículo se estructura de la siguiente forma. En la Sección 2 se proporcionan los detalles acerca de las herramientas empleadas en la experiencia, para la realización de las pruebas y el análisis del audio. La Sección 3 presenta detalles del contexto de la experiencia, como la materia, titulación, tipo de pruebas y métodos de obtención de datos. Los resultados de las pruebas se analizan y se visualizan en la Sección 4, donde se contesta a las preguntas de investigación formuladas a la luz de los mismos. Finalmente el artículo concluye en la Sección 5 con la discusión y las conclusiones obtenidas de la experiencia.

2. Herramientas utilizadas

En esta sección presentamos los detalles de las dos tecnologías implicadas en la experiencia docente. Primero, describimos brevemente Wooclap, la herramienta que ha sido objeto de estudio. A continuación, exponemos las principales características del sistema de análisis de audio empleado para medir el nivel de interacción en el aula durante el desarrollo de las pruebas.

2.1. Wooclap

Wooclap pertenece a la categoría de los denominados sistemas interactivos de respuesta. No son una tecnología reciente, dado que hay experimentos documentados en la década de 1960 [5], pero es cierto que hoy en día son más fáciles de usar, más completos y sólo se necesita un teléfono inteligente convencional para responder, por lo que se han popularizado mucho en la última década.

En un uso típico de esta tecnología, a los estudiantes se les presentan preguntas de verdadero/falso o de opción múltiple que tienen que responder utilizando la aplicación o el hardware asociado. A continuación, el docente muestra las respuestas y proporciona retroalimentación, que puede implicar una discusión colectiva posterior. Varios estudios, por ejemplo [9, 14], resumen algunos usos comunes de este tipo de tecnologías documentados en la literatura, que no deben reducirse únicamente a plantear cuestiones de repaso.

Existen muchas plataformas, sobradamente conocidas, como Kahoot!, Socrative, Wooclap, Quizizz, Quizlet o Poll Everywhere, por nombrar algunas. Wooclap es una herramienta muy similar a Socrative, tanto en lo funcional como en el aspecto visual, alejada de la estética de concurso de Kahoot!. Sin embargo, Wooclap incluye gran parte de la versatilidad de Kahoot!, como es la gran diversidad de tipos de preguntas a plantear y la posibilidad de insertar las preguntas entre las diapositivas que forman parte de una presentación. También dispone de un modo competitivo, aunque con menos elementos de gamificación que Kahoot!. Este modo competitivo se diferencia del modo sin competición en los siguientes aspectos:

- Los estudiantes deben identificarse siempre, bien con un seudónimo o con el nombre real.
- Los estudiantes obtienen puntuaciones por las respuestas correctas, en función del orden en el que llega su respuesta respecto a la del resto de estudiantes, siendo mayor cuanto antes se envía la respuesta.
- El docente puede mostrar en cualquier momento la clasificación en pantalla, pero solo si pulsa el botón correspondiente, ya que por defecto no se muestra la clasificación.

Podemos encontrar en la literatura algunos trabajos de investigación sobre Wooclap que analizan la implicación del alumnado en función del uso de la funcionalidad de mensajes [7], pero este trabajo es el primero en analizar las consecuencias del modo competición en el rendimiento y en la interacción.

2.2. Análisis de audio

Con el fin de medir el nivel de interacción de los estudiantes durante el desarrollo de las pruebas, empleamos técnicas de análisis de audio basadas en el aprendizaje automático. Este sistema de análisis no verbal se ha utilizado anteriormente [3] para analizar los audios de diferentes actividades en el aula, como las clases magistrales, la resolución de problemas, el aula invertida o el uso de cuestionarios interactivos. Se trata, por tanto, de una herramienta idónea para estudiar las diferencias entre el uso del modo competición y la ausencia de la misma con Wooclap.

El procesamiento del audio realizado por dicho sistema conlleva varias etapas. En primer lugar, se utiliza un método de umbralización para separar el silencio de los segmentos de voz y, utilizando pequeños segmentos de audio, se calculan los coeficientes cepstrales de frecuencia de Mel (MFCC), el tono y la energía media de cada segmento de voz. Estas características se utilizan ampliamente en la literatura científica [12] como entrada para la diarización del hablante.

La diarización es la segunda etapa y consiste en determinar “quién habló cuándo”, es decir, etiquetar segmentos de la grabación con el correspondiente hablante. En este artículo no distinguimos qué estudiante concreto está hablando, y todos los estudiantes que intervienen se etiquetan simplemente como Estudiantes. En concreto, distinguimos entre Docente, Estudiantes y Silencio. Este proceso de diarización se basa en técnicas de aprendizaje automático y es un proceso semi-supervisado, ya que hay que entrenar el sistema con algunos fragmentos de audio que se etiquetan como Docente o Estudiantes. Este método proporciona una precisión de clasificación en torno al 95 % con otros audios de prueba. En consecuencia, obtenemos una secuencia de etiquetas ordenada en el tiempo que se utilizará en la siguiente etapa, pero que también puede visualizarse mediante líneas de tiempo. En este trabajo no ponemos en el foco en qué se está diciendo en cada momento, sino simplemente en quién interviene, como hacen otros trabajos previos de análisis no verbal del clima en el aula [8].

El tercer paso es la obtención de las características del discurso no verbal a partir de la secuencia de etiquetas. De todas las características que el sistema proporciona, basadas en los trabajos previos [1, 10], se han empleado las que nos resultan más útiles para este trabajo. Específicamente, empleamos dos características

de discurso por rol (profesor, estudiantes):

- Ratio de participación del hablante (PSR). Mide la relación de participación de cada rol durante la grabación.
- Duración promedio de la intervención del hablante (APSUD). Medida en segundos, es la duración promedio de las intervenciones de cada tipo de participante.

Además, también se emplean, entre otras, las siguientes características globales de discurso:

- Duración promedio del lapso (ALD). Mide cuán largo es el intervalo de silencio promedio entre intervenciones.
- Conteo de turnos (TTC). Representa cuántos cambios de turno ocurrieron en el diálogo entre estudiantes y docente.

Finalmente, el sistema nos permite la visualización de los datos, utilizando diferentes técnicas de representación para mostrar las características calculadas y también los resultados del proceso de diarización. Como vemos más adelante, estos gráficos proporcionan información valiosa para analizar las actividades del aula, como los niveles de interacción.

3. Contexto de la experiencia

3.1. Materia y participantes

Este trabajo presenta la experiencia docente desarrollada durante el curso académico 2022/23 en la asignatura Fundamentos de Computadores, de primer curso del Grado en Ingeniería Informática de la Facultad de Informática de la Universidad de Murcia. Se trata de una asignatura de 6 créditos ECTS, 3 de teoría y 3 de prácticas. En el grupo que se ha realizado el estudio hay un total de 68 estudiantes matriculados, de los cuales han asistido a clase regularmente alrededor de 55. Para las clases prácticas los alumnos se dividen en tres grupos, cada grupo está compuesto por, aproximadamente, 22 alumnos. El docente de todos los grupos es siempre la misma persona.

En esta asignatura la teoría y las prácticas están estrechamente relacionadas, de forma que las clases de prácticas se usan para afianzar conceptos de teoría a través de ejercicios teórico-prácticos. Estos ejercicios algunas veces siguen el modelo tradicional y se desarrollan en pizarra, mientras que otras veces son ejercicios que se apoyan en el ordenador, usando alguna herramienta informática. Del total de clases prácticas, más del 75 % de las mismas usan herramientas informáticas, mientras que en menos del 25 % de las mismas no se usa el ordenador y se realizan ejercicios en pizarra siguiendo el modelo tradicional.

La metodología usada en esta asignatura abarca un gran abanico de opciones. Así, se hace uso de clases magistrales para introducir los conceptos de la misma. En las clases de prácticas, que se llevan a cabo en laboratorio con grupos reducidos, se resuelven problemas en ordenador/pizarra para afianzar dichos conceptos, en este caso se fomenta en todo momento la participación del alumno, tanto al resolver el docente el ejercicio, como proponiendo que los estudiantes resuelvan individualmente parte de los ejercicios (pero con apoyo del profesor cuando es necesario). Además, en dos de las sesiones teóricas del curso se hace uso de la metodología de aula invertida. Por último, al final de cada tema, se utiliza la herramienta Wooclap con cuestiones específicas del tema correspondiente, lo que permite resolver dudas y comprobar cómo los alumnos han afianzado los conceptos del mismo. El único cambio sustancial respecto a años anteriores ha sido que dos de las pruebas de Wooclap han sido movidas de las sesiones de teoría a las sesiones de laboratorio, para poder realizar nuestra investigación.

3.2. Diseño de actividades

Hemos usado la herramienta Wooclap como sistema interactivo de respuesta para hacer nuestro estudio. Las dos pruebas realizadas con Wooclap en las sesiones de prácticas se han configurado de la misma manera: están compuestas por 10 preguntas, con un tiempo máximo de respuesta de 60 segundos por pregunta.

La primera prueba está compuesta de nueve preguntas de “Buscar en una imagen” y una pregunta de respuesta múltiple (con tres posibles respuestas y sólo una de ellas válida). En las preguntas de “Buscar en una imagen” se presenta la imagen de una placa base completa, se le pregunta por un componente de la misma, y el alumno tiene que marcar qué zona de la imagen se corresponde con dicho componente.

En la segunda prueba las diez preguntas giran alrededor de los distintos lenguajes del computador. Son de respuesta múltiple, teniendo cada pregunta tres posibles respuestas, de las cuales sólo una es válida.

El estudio se ha llevado a cabo durante las sesiones de prácticas y sólo dos de los tres grupos de prácticas han participado en la investigación aquí presentada. En uno de los grupos prácticas se han tenido que hacer adaptaciones pedagógicas específicas debido al alumnado presente, por lo que se sigue una metodología algo diferente y no se incluye en esta investigación.

Por tanto, en la investigación participan los grupos 2.1 y 2.2, que en adelante llamaremos G. 2.1 y G. 2.2, respectivamente. En ambos grupos, y en ambas sesiones, se realizan las mismas pruebas, salvo en el aspecto de la competición. En concreto, para el G. 2.1 la primera prueba se configura sin competición, mientras

que para el G. 2.2 se configura con competición autenticada. Para la segunda prueba, esta configuración se invierte, de forma que el G. 2.1 realiza la segunda prueba con competición autenticada y el G. 2.2 la hace sin competición. El horario de los grupos de prácticas es consecutivo dentro de la misma mañana.

3.3. Procedimiento de obtención de datos

Con el fin de responder a las preguntas de investigación, se han recopilado varios elementos de información a partir de diversas fuentes de datos. En primer lugar, se dispone de los resultados obtenidos a través de Wooclap. Se trata de las puntuaciones numéricas de los estudiantes para cada prueba realizada.

En segundo lugar, respecto a la grabación de los audios, todos los estudiantes fueron informados acerca del procedimiento (aprobado por la institución en el marco de un proyecto de innovación docente), la finalidad del estudio y el compromiso de no desvelar por ningún medio la identidad de los participantes intervinientes. Para efectuar las grabaciones se situó una grabadora digital de mano (TASCAM DR-07X) a una distancia mínima de 1,5 metros del profesor y de los alumnos de la primera fila. Posteriormente se procesaron los audios para seleccionar únicamente los fragmentos que abarcan desde el instante en el que se lanzaba el enlace para conectar hasta el momento en el que se da por concluida la prueba. El resultado son cuatro grabaciones con duraciones de entre 13 y 15 minutos. Posteriormente dichas grabaciones fueron procesadas por el sistema de análisis de audios [3] para generar la diarización de las mismas y la extracción de características, datos que se analizan en detalle en la próxima sección.

4. Resultados

En esta sección se analizan los datos recogidos para responder a las dos preguntas de investigación formuladas en la sección introductoria.

Para analizar el rendimiento obtenido en ambas pruebas se han analizado las distribuciones de las puntuaciones obtenidas. Los Cuadros 1 y 2 muestran el rendimiento alcanzado en las pruebas 1 y 2, respectivamente, junto con el número de estudiantes que participaron. En concreto, en dichas tablas usamos la siguiente notación: N para el número de participantes, Media para la nota media, D.E. para la desviación estándar y la p de Shapiro-Wilk para comprobar el criterio de normalidad en las puntuaciones. “Comp.” indica que la prueba se configuró con competición, mientras que “Sin Comp.” indica que la prueba se configuró sin competición. Para completar nuestro análisis descriptivo la Figura 1 presenta las distribuciones de rendimiento obtenidas por ambos grupos en ambas pruebas.

| | Grupo | Valor |
|--------------------------|--------------------|-------|
| N | Comp. (G. 2.2) | 18 |
| | Sin Comp. (G. 2.1) | 20 |
| Media | Comp. (G. 2.2) | 7.28 |
| | Sin Comp. (G. 2.1) | 6.75 |
| D.E. | Comp. (G. 2.2) | 1.74 |
| | Sin Comp. (G. 2.1) | 1.33 |
| p de Saphiro-Wilk | Comp. (G. 2.2) | 0.004 |
| | Sin Comp. (G. 2.1) | 0.34 |

Cuadro 1: Descriptivas del rendimiento de la prueba 1.

| | Grupo | Valor |
|--------------------------|--------------------|-------|
| N | Comp. (G. 2.1) | 19 |
| | Sin Comp. (G. 2.2) | 17 |
| Media | Comp. (G. 2.1) | 6.84 |
| | Sin Comp. (G. 2.2) | 8.06 |
| D.E. | Comp. (G. 2.1) | 2.22 |
| | Sin Comp. (G. 2.2) | 1.43 |
| p de Saphiro-Wilk | Comp. (G. 2.1) | 0.168 |
| | Sin Comp. (G. 2.2) | 0.271 |

Cuadro 2: Descriptivas del rendimiento de la prueba 2.

4.1. Prueba 1

A partir de los datos del Cuadro 1 podemos comprobar que, dado un número similar de participantes, los resultados del grupo con competición (G. 2.2) fueron superiores en media a los del grupo sin competición (G. 2.1). No obstante, la Figura 1 parece indicar que las diferencias no son estadísticamente significativas. Para corroborarlo, dado que no se cumple el criterio de normalidad en las puntuaciones del grupo 2.2, se ha empleado el test no paramétrico de Kruskal-Wallis, con $\chi^2 = 2.32$ y $p = 0.127$, lo que confirma que las diferencias no son significativas ($p > 0.05$).

Sin embargo, si analizamos las grabaciones de audio de la Prueba 1 podemos observar diferencias en el patrón de interacción. Las Figuras 2a y 2b muestran líneas temporales en las que se resaltan los instantes en los cuales intervienen los distintos tipos de participantes. Un análisis comparativo de ambas indica un mayor porcentaje de intervención de los estudiantes en el modo con competición. De hecho, el grado de participación casi se triplica, como se aprecia en las características de PSR (ratio de participación de hablante) y TTC (conteo de turnos) que se ilustran en las Figuras 2c y 2d.

4.2. Prueba 2

Para la segunda prueba, volvemos a disponer de un número similar de participantes en cada grupo. Observamos que la nota media del grupo sin competición

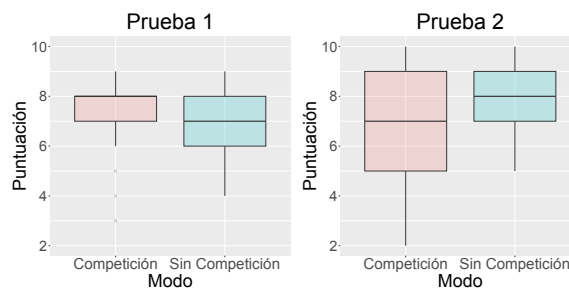


Figura 1: Distribución de puntuaciones para las pruebas 1 y 2 en base al modo competición en Wooclap.

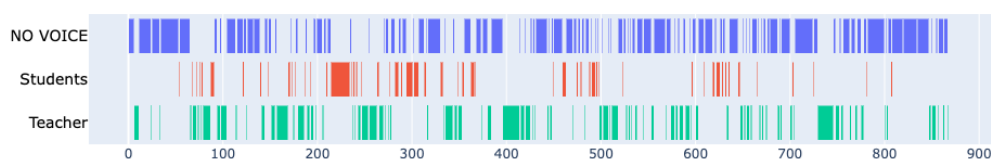
(G. 2.2) es superior a la del grupo con competición. En la Figura 1 apreciamos una diferencia en la distribución de puntuaciones, por lo que tenemos que corroborar si la diferencia es estadísticamente significativa. En esta ocasión podemos emplear un test paramétrico dado que se cumple la condición de normalidad en las distribuciones. El resultado con ANOVA de Welch es $F = 3.90$ y $p = 0.057$, por lo que de nuevo se rechaza la hipótesis de que las puntuaciones difieran estadísticamente ($p > 0.05$).

Por tanto, se aprecia que en las pruebas realizadas, el grupo 2.2 siempre saca una nota media más alta, independientemente del modo. Esto queda también reflejado en la Figura 1, ya que para el G. 2.2 la distribución de puntuaciones ocupa siempre un menor rango teniendo el valor mínimo más alto.

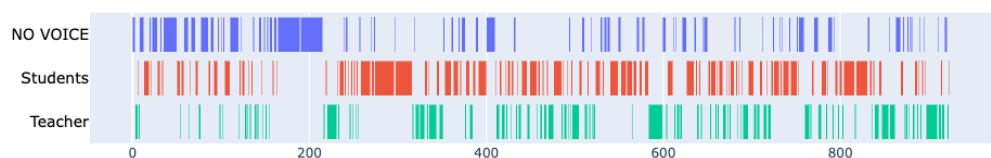
Al analizar las grabaciones de audio de la Prueba 2 encontramos diferencias aún más acusadas en los patrones de interacción. Las Figuras 3a y 3b evidencian que el grado de interacción en el modo sin competición es prácticamente inexistente, muy inferior al del grupo 2.1. El silencio es mayoritario durante el 70 % del tiempo, como indica la Figura 3d, con una duración de las pausas del docente también notablemente superior. Sin embargo, podemos apreciar que el grupo en modo competición tiene un patrón bastante regular de intervenciones, que en este caso se relaciona con los murmullos y las conversaciones que se producen entre los estudiantes mientras que esperan el resultado de las preguntas o cuando interrogan al docente.

4.3. Respuesta a la pregunta 1

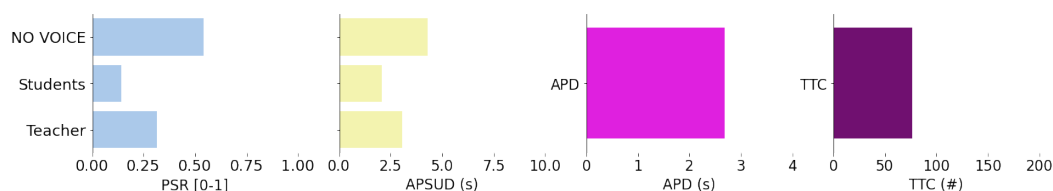
La primera pregunta planteaba si existen diferencias significativas en las puntuaciones obtenidas al utilizar el modo competición en Wooclap frente a no usarlo. Según nuestro análisis la respuesta es no. Los datos indican que estadísticamente no hay diferencias significativas. En este caso, el grupo G. 2.2 obtiene siempre mejores puntuaciones, sin importar si la prueba fue configurada en modo competición o no. Quizá sea importante apuntar que también se analizaron los resulta-



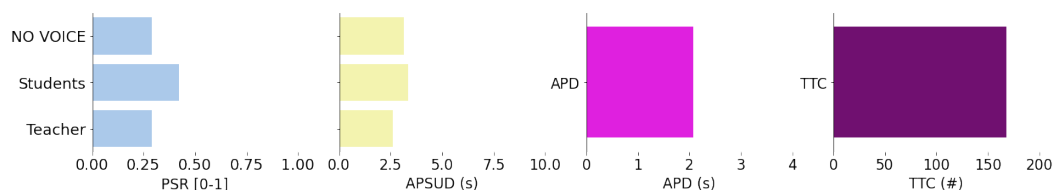
(a) Línea temporal de intervenciones. Sin competición (G. 2.1).



(b) Línea temporal de intervenciones. Con competición (G. 2.2).



(c) Descriptores de los audios. Sin competición (G. 2.1).



(d) Descriptores de los audios. Con competición (G. 2.2).

Figura 2: Resultados del procesamiento de los audios de la prueba 1.

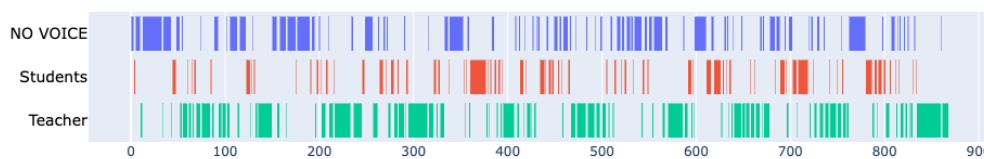
dos del examen final de teoría (igual para todos los grupos) para comprobar si efectivamente los alumnos del grupo G. 2.2 también obtenían mejores calificaciones en media, pero fueron similares en ambos grupos. En concreto, calificación obtenida en el examen final por los alumnos del G. 2.1 fue de ($\mu = 5.75, \sigma = 1.51$), mientras que los alumnos del G. 2.2 obtuvieron ($\mu = 5.22, \sigma = 1.95$), sin diferencias estadísticamente significativas entre ambos grupos.

Por tanto, basándonos en los resultados presentados, podemos concluir que usar el modo de competición no influye en el resultado obtenido por los alumnos en nuestra experiencia.

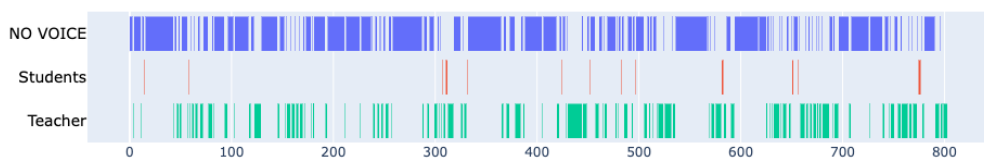
4.4. Respuesta a la pregunta 2

Respecto a la segunda pregunta, acerca de la diferencia en los patrones de interacción de los estudiantes,

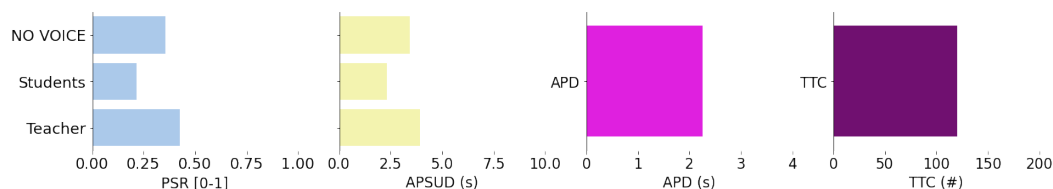
los datos confirman que sí hay diferencias importantes en las líneas temporales y en las características extraídas de los audios. Las pruebas basadas en competición muestran un nivel de interacción mayor, de forma más uniforme a lo largo del desarrollo de las mismas. La participación de los estudiantes se triplica, al menos, respecto al modo sin competición. También el docente interviene de forma más continua, con pausas más cortas y por tanto de forma más dinámica. Por ejemplo, cuando hay competición el docente tiene que poner los elementos de la gamificación en relieve, lo que incrementa su número de intervenciones. Además, hay más alternancia en las intervenciones entre docente y estudiantes, como indican los valores más elevados de TTC (conteo de turnos).



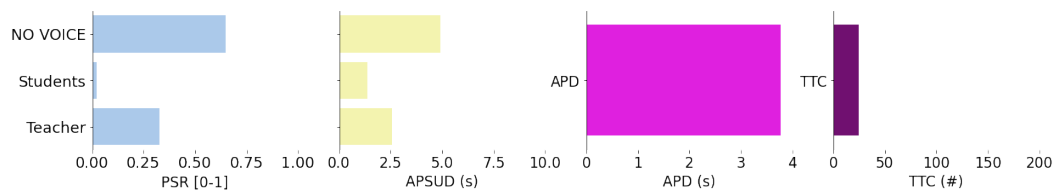
(a) Línea temporal de intervenciones. Con competición (G. 2.1).



(b) Línea temporal de intervenciones. Sin competición (G. 2.2).



(c) Descriptores de los audios. Con competición (G. 2.1).



(d) Descriptores de los audios. Sin competición (G. 2.2).

Figura 3: Resultados del procesamiento de los audios de la prueba 2.

5. Discusión y conclusiones

Son varios los estudios que reflejan que el uso de sistemas interactivos de respuesta en clase beneficia el clima generado, favoreciendo la participación activa del alumnado [6] y contribuyendo, en cierta medida, a mejorar los resultados del proceso de aprendizaje.

En este trabajo hemos indagado acerca de si el uso del modo competición o no en estas herramientas influye en el rendimiento de los estudiantes. Para ello hemos basado nuestro análisis en las evidencias recogidas a través del rendimiento obtenido en unas pruebas con Wooclap configuradas con y sin competición, y en los patrones de interacción de los alumnos durante dichas pruebas. Hasta ahora, la mayoría de los trabajos previos se habían limitado a la percepción subjetiva de los estudiantes, por ejemplo respecto a usar el modo competición de forma reiterada [13].

Los resultados presentados reflejan dos aspectos interesantes. Por un lado, el rendimiento obtenido por el estudiante no depende de la configuración de la prueba. Sin embargo, esta configuración sí influye en el grado de interacción del alumnado en clase, ya que las pruebas basadas en competición muestran un nivel de interacción mayor que las realizadas sin competición.

Nuestra experiencia docente a lo largo de las pruebas realizadas indica que el clima generado en clase con las pruebas con competición es más lúdico y ameno, el alumnado está más integrado en las mismas, y parece que muestra más interés aunque sólo sea por el hecho de rivalizar entre compañeros. Con el modo competición se genera más interacción en todos los sentidos, es decir docente-estudiante y estudiante-estudiante. Desde el punto de vista del docente, el modo competición introduce nuevos elementos en sus intervenciones, dado que tiene que hacer referencia explícita a las puntua-

ciones, clasificaciones y los nombres o los seudónimos de los estudiantes, lo que le sitúa en otro rol distinto al habitual por momentos. Por otro lado, los estudiantes también participan más, comentan los resultados y la clasificación. Da la sensación subjetiva de que los estudiantes están más motivados y atentos.

Destacar que este ambiente también puede llevar a un mayor nivel de distracción de los alumnos y un mayor nivel de ruido en el aula entre cada pregunta de la prueba, lo que puede conllevar un mayor esfuerzo por parte del docente para continuar con la prueba.

Cuando la prueba se lleva a cabo sin competición, el clima es más pasivo y menos lúdico. El estudiante no es tan participativo, y se queda a la espera de recibir la respuesta correcta, normalmente sin interactuar mucho, en un clima similar al de las clases magistrales, donde el profesor diserta y el alumno escucha. En este modo los estudiantes parecen prestar más atención a las explicaciones del docente tras cada pregunta formulada, pero es una apreciación subjetiva.

La conclusión final es que debe ser el docente el que elija el modo con competición o sin competición según el clima que quiera crear en clase, dado que no parece afectar a las puntuaciones obtenidas. Así, si el docente quiere tener un clima más interactivo, con alumnos más participativos, aunque ello pueda conllevar una posible pérdida de atención a la realimentación que el docente realiza tras las preguntas, debe configurar la prueba en modo competición. Por el contrario, si el docente quiere un ambiente más relajado, en el que los alumnos puedan centrar más la atención en sus explicaciones, aunque serán muy probablemente menos participativos, es una buena alternativa configurar la prueba sin competición.

Como vía de trabajo futuro, queremos extender el estudio al uso de distintas plataformas de respuesta interactiva, dado que hay diferencias de diseño importantes en el sistema de gamificación que cada una incorpora. Esto nos permitirá analizar si el rendimiento y la interacción también están condicionados por el tipo de plataforma.

Referencias

- [1] Indrani Bhattacharya, Michael Foley, Ni Zhang, Tongtao Zhang, Christine Ku, Cameron Mine, Heng Ji, Christoph Riedl, Brooke Foucault Welles y Richard J Radke. A multimodal-sensor-enabled room for unobtrusive group meeting analysis. En *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, pp. 347–355, 2018.
- [2] Óscar Cánovas. Sistemas interactivos de respuesta: hay vida más allá de los concursos. *Actas de las XXVII JENUI*, pp. 243–250, 2021.
- [3] Óscar Cánovas y Félix J. García. Analysis of classroom interaction using speaker diarization and discourse features from audio recordings. En *Learning in the Age of Digital and Green Transition*, pp. 67–74, 2023.
- [4] José I. Castillo-Manzano, Mercedes Castro-Nuño, Lourdes López-Valpuesta, María Teresa Sanz-Díaz y Rocío Yñiguez. Measuring the effect of ars on academic performance: A global meta-analysis. *Computers & Education*, 96:109–121, 2016.
- [5] Emily Dill. Do clickers improve library instruction? Lock in your answers now. *The Journal of Academic Librarianship*, 34(6):527–529, 2008.
- [6] Debra Filer. Everyone's answering: Using technology to increase classroom participation. *Nursing education perspectives*, 31(4):247–250, 2010.
- [7] Jérôme Hutain y Nicolas Michinov. Improving student engagement during in-person classes by using functionalities of a digital learning environment. *Computers & Education*, 183:104496, 2022.
- [8] Anusha James, Yi Han Victoria Chua, Tomasz Maszczyk, Ana Moreno Núñez, Rebecca Bull, Kerry Lee y Justin Dauwels. Automated classification of classroom climate by audio analysis. En *9th International Workshop on Spoken Dialogue System Technology*, pp. 41–49. Springer, 2019.
- [9] Robin H Kay y Ann LeSage. Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3):819–827, 2009.
- [10] Catherine Lai, Jean Carletta y Steve Renals. Modelling participant affect in meetings with turn-taking features. En *Proc. Workshop of Affective Social Speech Signals*, 2013.
- [11] Tuan Dinh Nguyen, Marisa Cannata y Jason Miller. Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *The Journal of Educational Research*, 111(2):163–174, 2018.
- [12] Tae Jin Park, Naoyuki Kanda, Dimitrios Dimitriadis, Kyu J Han, Shinji Watanabe y Shrikanth Narayanan. A review of speaker diarization: Recent advances with deep learning. *Computer Speech & Language*, 72:101317, 2022.
- [13] Alf Inge Wang. The wear out effect of a game-based student response system. *Computers & Education*, 82:217–227, 2015.
- [14] Weiwen Wang, Sun Ran, Linda Huang y Valerie Swigart. Student perceptions of classic and game-based online student response systems. *Nurse educator*, 44(4):6–9, 2019.