



Universitat d'Alacant  
Universidad de Alicante

**Memorias del Programa  
de Redes-I3CE de calidad,  
innovación e investigación  
en docencia universitaria**

Convocatoria  
**2020-21**

**Memòries del Programa  
de Xarxes-I3CE de qualitat,  
innovació i investigació  
en docència universitària**

Convocatòria  
**2020-21**



Satorre Cuerda, Rosana (Coordinación)  
Menargues Marcilla, María Asunción; Díez Ros, Rocío; Pellín Buades, Neus (Eds.)

**UA**

UNIVERSITAT D'ALACANT  
UNIVERSIDAD DE ALICANTE

Vicerectorat de Transformació Digital  
Vicerrectorado de Transformación Digital  
Institut de Ciències de l'Educació  
Instituto de Ciencias de la Educación

*Memorias del Programa de Redes-I3CE de calidad, innovación e investigación en docencia universitaria. Convocatoria 2020-21 / Memòries del Programa de Xarxes-I3CE de qualitat, innovació i investigació en docència universitària. Convocatòria 2020-21*

Organització: Institut de Ciències de l'Educació (Vicerectorat de Transformació Digital) de la Universitat d'Alacant/ *Organización: Instituto de Ciencias de la Educación (Vicerrectorado de Transformación Digital) de la Universidad de Alicante*

Edició / *Edición*: Rosana Satorre Cuerda (Coord.), Asunción Menargues Marcillas, Rocío Díez Ros, Neus Pellin Buades

Revisió i maquetació: ICE de la Universitat d'Alacant/ *Revisión y maquetación: ICE de la Universidad de Alicante*

Primera edició / *Primera edición*: desembre 2021/ diciembre 2021

© De l'edició/ *De la edición*: Rosana Satorre Cuerda, Asunción Menargues Marcillas, Rocío Díez Ros & Neus Pellin Buades

© Del text: les autores i autors / *Del texto: las autoras y autores*

© D'aquesta edició: Universitat d'Alacant / *De esta edición: Universidad de Alicante*

ice@ua.es

Memorias del Programa de Redes-I3CE de calidad, innovación e investigación en docencia universitaria. Convocatoria 2020-21 / Memòries del Programa de Xarxes-I3CE de qualitat, innovació i investigació en docència universitària. Convocatòria 2020-21 © 2021 by Universitat d'Alacant / Universidad de Alicante is licensed under [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) 

ISBN: 978-84-09-34941-8

Qualsevol forma de reproducció, distribució, comunicació pública o transformació d'aquesta obra només pot ser realitzada amb l'autorització dels seus titulars, llevat de les excepcions previstes per la llei. Adreceu-vos a CEDRO (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) si necessiteu fotocopiar o escanejar algun fragment d'aquesta obra. / *Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) si necesita fotocopiar o escanear algún fragmento de esta obra.*

Producció: Institut de Ciències de l'Educació (ICE) de la Universitat d'Alacant / *Producción: Instituto de Ciencias de la Educación (ICE) de la Universidad de Alicante*

Aquesta publicació s'ha fet seguint les directrius d'accessibilitat UNE-EN 301549:2020 / Esta publicación se ha hecho siguiendo las directrices de accesibilidad UNE-EN 301549:2020.

EDITORIAL: Les opinions i continguts dels treballs publicats en aquesta obra són de responsabilitat exclusiva de les autores i dels autors. / *Las opiniones y contenidos de los trabajos publicados en esta obra son de responsabilidad exclusiva de las autoras y de los autores.*

## 43. Impacto del factor humano en el éxito de actividades de evaluación entre iguales

C. Cachero Castro\*; J.R. Rico Juan\*; H. Macià Soler\*\*\*; S. Meliá Beigbeder\*; S. Luján Mora\*; P. Barra Hernández\*\*; M. de Juan Vigaray\*\*\*\*

ccachero@dlsi.ua.es (Cristina Cachero Castro), juanramonrico@ua.es (Juan Ramón Rico-Juan), hermenegilda.macia@uclm.es (Mere Macià Soler), santi@ua.es (Santiago Meliá Beigbeder), slujan@dlsi.ua.es (Sergio Luján Mora), mpbarra@ucam.edu (Pilar Barra Hernández); mayo@ua.es (Mayo de Juan Vigaray)

\* Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, Alicante, 03690, Spain

\*\* Departamento de Económicas. Universidad Católica San Antonio de Murcia, Murcia, 30107, Spain

\*\*\* Departamento de Matemáticas. Universidad Castilla-La Mancha. Albacete, 02071, Spain.

\*\*\*\* Departamento de Marketing. Universidad de Alicante. Alicante, 03690, Spain.

## Resumen (Abstract)

En este trabajo se analiza la relación entre personalidad, intención de adopción y las distintas modalidades de evaluación entre pares. El propósito de este análisis es discernir si la personalidad influye en la preferencia por una evaluación entre pares frente a una evaluación tradicional (realizada por un experto), y si existen diferencias significativas en las preferencias con respecto a la modalidad de evaluación (corrección individual, en parejas o en tríos). Para ello, se presenta un quasi-experimento realizado en la Universidad de Castilla-La Mancha y coordinado con la Universidad de Alicante con una muestra de 85 estudiantes de primer curso del grado en Ingeniería Informática. Los estudiantes realizaron tres entregas, y en cada una evaluaron a sus compañeros con una modalidad distinta, de manera que al final todos los estudiantes habían experimentado los tres modos de evaluación. Los resultados muestran que las personas más amables y/o menos neuróticas presentan más reticencias a evaluar a los compañeros. Además, los resultados indican que las modalidades de evaluación entre pares no afectan de manera significativa la intención de adopción de esta técnica.

**Palabras clave:** Evaluación colaborativa, estudio empírico, quasi-experimento, factor humano, peer assessment

## 1. Introducció

El proceso de la evaluación entre iguales, también conocido como evaluación por o entre pares (*peer assessment*) es un método de evaluación cada vez más nombrado como alternativa a la evaluación tradicional en el contexto universitario, debido a que ofrece diferentes ventajas tanto al profesor -como instructor- como al alumno -como evaluador-. Por parte del instructor, permite realizar una evaluación formativa en la que los alumnos reciben una retroalimentación rápida sobre los trabajos realizados por parte de sus compañeros. El papel del instructor en este proceso consiste en la supervisión de todo el proceso. Esta supervisión, sobre todo en asignaturas con muchos alumnos, es menos costosa y permite una retroalimentación más rápida que la revisión del instructor de todos los trabajos. Por su parte, el alumno adquiere un rol de evaluador sobre un trabajo cuyo enunciado ya conoce, dado que ha tenido que realizarlo previamente. En este papel, el alumno se ve obligado a ser crítico con respecto a diferentes aspectos tanto de su propio trabajo (autoevaluación) como del trabajo de sus compañeros para determinar el nivel de acierto en cada uno de ellos (Rico-Juan et. al, 2019). Por ello, el alumno desarrolla un hábito de reflexión y de espíritu crítico constructivo ante soluciones tanto propias como propuestas por compañeros. Además, este proceso ayuda a interiorizar los criterios de corrección que el instructor hace explícitos en las instrucciones de corrección. Aún más, si la evaluación se hace por grupos, es de esperar que las discusiones entre iguales sobre las soluciones ayuden a entender mejor los conceptos estudiados que si los explicara el profesor (Mazur, 1997). Todas estas ventajas se han comprobado de manera empírica en algunos meta-análisis recientes que se han enfocado en el efecto de la evaluación entre pares sobre el aprendizaje en distintos contextos de aprendizaje (Li et al., 2020, Zheng et al., 2020).

Sin embargo, con el fin de maximizar los beneficios sobre el aprendizaje, es importante tener en cuenta los retos inherentes a la implementación exitosa de este tipo de evaluación. En este sentido, Adachi et al. (2018) y Liu et al. (2016) identificaron cuatro posibles obstáculos a la hora de conseguir este éxito: restricciones de tiempo y recursos, relaciones de poder, la pericia percibida y la

fiabilidad y exactitud de las habilidades de juicio de los estudiantes. Estos inhibidores parecen estar en la base de la baja adopción en la práctica de la autoevaluación y la evaluación entre pares en el contexto de la educación superior (Panadero et al. 2019). Sin embargo, existen otros posibles inhibidores, como pueden ser las diferencias individuales de los estudiantes, que no han sido suficientemente investigados (Chang et al., 2021). Dado que, dentro del campo de la psicología educativa, es un hecho reconocido el impacto que tienen estas diferencias individuales sobre el aprendizaje y el logro (An & Carr, 2017), este trabajo pretende contribuir a clarificar este aspecto.

## 1.1 Problema o cuestión específica del objeto de estudio

Dentro de la relación entre Factor Humano (FH) y efectos y preferencias de aprendizaje, en esta red nos hemos centrado en el estudio de la personalidad de los alumnos y de su percepción subjetiva (utilidad percibida, facilidad de uso, compatibilidad y, en último término, intención de adopción) del proceso de evaluación entre pares. Según el APA (APA, 2021), la personalidad se define como *las diferencias individuales en patrones característicos de pensar, sentir y comportarse*. Para la operacionalización de este constructo hemos adoptado el modelo de personalidad de los cinco factores, también conocido como Big Five (BF, John et al., 2008), debido a sus beneficios en términos de completitud y fiabilidad de sus medidas (Balijepally et al. 2006). Este modelo define cinco factores como fuentes fundamentales de las diferencias individuales: Extraversión (E), Apertura a la experiencia (O), Amabilidad (A), Escrupulosidad (*Conscientiousness*, C) y Neuroticismo (N). A continuación, definimos cada una de estas cinco dimensiones (Ter Laak, 1996):

- Extraversión: grado en el que una persona se abre al resto. Se relaciona con la imaginación, la creatividad, la curiosidad, la sensibilidad artística, una vida interior rica, el gusto por la variedad, una tendencia a buscar la compañía de otros e independencia de juicio. Niveles altos de E también se relacionan con personas asertivas, habladoras, optimistas y en búsqueda constante de emociones nuevas.

- Neuroticismo (N): grado en el que una persona es inestable emocionalmente. Las personas con un nivel alto de N tienden a mostrar signos de ansiedad, preocupación y una percepción sesgada hacia situaciones negativas. Tienden también a mostrar una menor tolerancia al estrés.
- Amabilidad (A): grado en el que una persona es respetuosa, tolerante y tranquila. Se refiere a tendencias interpersonales. Valores altos de A se relacionan con personas altruistas, consideradas, generosas y ayudadoras.
- Escrupulosidad (*Conscientiousness*, C): también conocida como ‘deseo de triunfar’. Es el grado en que una persona está centrada en conseguir sus metas. Se relaciona con el autocontrol, la planificación, la organización y la ejecución de tareas. Valores altos de C implican personas que son muy responsables y en las que se puede confiar a la hora de asignar tareas. Valores bajos se relacionan con personas más relajadas, informales y hedonistas.
- Apertura a la Experiencia (*Openness to experience*, O): grado en que a una persona le gusta buscar nuevas experiencias de manera creativa.

Por su parte, la intención de adopción de distintos tipos de innovaciones ha sido estudiada ampliamente en la literatura científica, donde se han propuesto distintos modelos para su operacionalización (Diéguez et al., 2012). La mayoría de ellos se derivan del Modelo de Aceptación de la Tecnología (*Technology Acceptance Model*, TAM) (Davis, 1989), y han sido aplicados a distintos dominios, incluido el dominio de educación (Rakoczy et al., 2019). En este trabajo hemos adoptado el modelo de Intención de Adopción (IA) presentado en (Martínez et al., 2013), que propone Utilidad (U), Facilidad de Uso (*Ease of Use*, EoU) y Compatibilidad (Comp) como los tres componentes principales que conforman la Intención de Adopción (IA) de nuevos procesos de trabajo, como es el caso del nuevo proceso que implica la evaluación entre pares. A continuación, definimos cada uno de ellos:

- Utilidad (U): Grado en que el sujeto cree que usar un método particular mejorará su rendimiento.

- Facilidad de uso (EoU): Grado en que el sujeto cree que usar el método o proceso no va a costarle trabajo.
- Compatibilidad (Comp): Grado en que un nuevo método o proceso se percibe como consistente con los valores, principios, prácticas y experiencia pasada del sujeto.
- Intención de Adopción (IA): Intención global de adoptar el nuevo método o proceso en el futuro si el sujeto tiene oportunidad.

Por último, es importante destacar cómo la evaluación entre pares puede adoptar dos modalidades distintas: individual (cada evaluador revisa de manera individual el trabajo de sus compañeros), o colaborativa (los evaluadores se agrupan en grupos para evaluar el trabajo de sus compañeros). Estos grupos pueden estar formados además por distintos números de componentes: dos (parejas), tres (tríos), etc. En base a la teoría de desarrollo social de Vygotsky (Vygotsky, 1980), que enfatiza el rol tan importante que juega la interacción social en el aprendizaje, es plausible argumentar que una evaluación colaborativa debería aumentar el impacto del proceso en el aprendizaje. Sin embargo, no está claro si ocurre lo mismo con la intención de adopción de esta modalidad de evaluación.

Esta investigación es de utilidad para cualquier instructor que desee maximizar el impacto de introducir procesos de evaluación entre iguales en su contexto educativo.

## 1.2 Revisión de la literatura

La personalidad es un factor humano que tiene un impacto sustancial tanto en la productividad y en las actitudes de los sujetos (Curtis et al., 1998) como en la intención de adoptar nuevos procesos (O'Connor & Yilmaz, 2015). A pesar de este hecho reconocido, las innovaciones educativas rara vez incluyen la personalidad como variable mediadora de su impacto. Este constructo sí se ha estudiado sin embargo en otros contextos. En Devaraj et al. (2008) se propone un modelo conceptual que incluye los rasgos de personalidad incluidos en el BF como predictores de la IA. El trabajo concluye que tanto N como A influyen la U, mientras que E, C y A moderan las relaciones entre U e IA. En Jarillo-Nieto et al.

(2015) los autores también usan el modelo BF para mostrar cómo niveles más altos de neuroticismo se relacionan con un menor apego a los procesos estructurados, mientras que altos niveles de A, O y C se relacionan con una mayor satisfacción con la adopción de nuevos procesos. Otro trabajo que muestra cómo tolerancia a la ambigüedad y O se asocian con una mayor habilidad para adaptarse a innovaciones tecnológicas es el realizado por Gallivan (2004).

En relación con el trabajo en grupo, en Feldt et al. (2010) los autores muestran cómo altos niveles de E se relacionan con una preferencia a trabajar en grupos, mientras que niveles altos de O se relacionan con la preferencia hacia tomar la responsabilidad de todo un proyecto y no de partes individuales.

### **1.3 Propósitos u objetivos**

El objetivo general de este trabajo es el estudio de cómo las distintas dimensiones de la personalidad según el modelo BF se relacionan con la preferencia por un proceso de evaluación entre pares frente a una evaluación tradicional. Además, se pretende estudiar cómo las percepciones subjetivas de U, EoU y Comp varían en función de la modalidad concreta de evaluación entre pares.

## **2. Método**

### **2.1. Descripción del contexto y de los participantes**

En este estudio participaron 85 estudiantes (8 mujeres) del curso Cálculo y Métodos Numéricos, ubicado en el primer semestre del grado de Ingeniería Informática de la UCLM. La edad media de los estudiantes fue de 18 años. Los estudiantes estaban divididos en dos grupos aleatorios (A y B). En ambos grupos, los estudiantes realizaron 3 actividades de evaluación entre pares, cada una con una modalidad distinta (individual, pares, tríos). La tabla 1 muestra el diseño final del estudio. Todos los participantes realizaron las tres actividades individuales a la vez en condiciones de examen. La evaluación entre pares se

efectuó dos días después de haber realizado la actividad en una sesión de clase. La duración de todas las sesiones de evaluación entre pares se estableció en 1.5 h. Una vez acabada la evaluación, y antes de abandonar el aula, todos los alumnos cumplimentaron un cuestionario donde se recogía su intención de adopción de la modalidad de evaluación entre pares que habían aplicado.

Tabla 1. Diseño del estudio experimental

Actividad	Semana	# Total Previstos	# Total Presentados	Grupo	Modalidad	# Previstos	# Presentados
A1	4	85	70	A	Individual	44	36
				B	Pares	41	34
A2	9	85	64	A	Pares	44	30
				B	Tríos	41	34
A3	14	85	59	A	Tríos	44	31
				B	Individual	41	28

## 2.2. Instrumento utilizado para realizar la investigación

Tal y como hemos comentado en el trabajo relacionado, el instrumento utilizado para capturar el perfil de personalidad de cada alumno ha sido el cuestionario BF.

Para capturar la intención de adopción, el instrumento utilizado ha sido una simplificación del UMAM-Q (Diéguez et al., 2012).

Además, se añadió una pregunta donde el alumnado debía indicar si prefería realizar una evaluación entre pares o seguir con el sistema tradicional de ser evaluado por un experto.

## 2.3. Procedimiento

El proceso experimental puede dividirse en cuatro fases:

- Complimentación del perfil de personalidad. Cada estudiante debía cumplimentar el cuestionario BF, disponible online a través de la plataforma Qualtrics.
- Ejecución de la actividad propiamente dicha. Cada una de las actividades se realizó de manera individual en condiciones de examen, con lápiz y papel. Se asignó un tiempo de 1.5 h para realizar cada actividad. Para considerarse entregada, la actividad tenía que ser subida a un taller de Moodle.
- Ejecución del proceso de evaluación entre pares. Con anterioridad a la sesión de evaluación, el experto debía realizar en cada grupo la asignación de correctores a trabajos, así como publicar una propuesta de solución y una guía de evaluación con criterios cuantitativos claramente especificados. Una vez en el aula con los alumnos, el proceso se definió como sigue:
  - El experto explica la solución de la actividad y los criterios de evaluación (20 min)
  - Los estudiantes se agrupan según modalidad (parejas, tríos) (5 min)
  - Los estudiantes evalúan los trabajos asignados (55 min)
  - Los estudiantes rellenan el instrumento UMAM-Q (10 min)
- Tras la realización de las tres actividades (con las tres modalidades de evaluación entre pares), el alumno contesta un cuestionario final donde se le pregunta qué tipo de evaluación prefiere.

## 3. Resultados

### 3.1. Perfil de Personalidad de las distintas preferencias

La Tabla 2 muestra los estadísticos descriptivos para las variables de personalidad (Extraversión, Amabilidad, Escrupulosidad, Neuroticismo y

Apertura a la Experiencia) en función de si los sujetos prefieren una evaluación tradicional (realizada por un experto) o una evaluación entre pares. Según estos datos, los alumnos que prefieren la evaluación entre pares parecen ser ligeramente más extrvertidos, menos amables, más escrupulosos, más neuróticos y menos abiertos a la experiencia.

Tabla 2. Estadísticos descriptivos BF

Variable	Preferencia	Media	SD	Mediana
BF-E	Experto	23.45	5.23	23
	Entre pares	23.76	4.51	25
BF-A	Experto	35.10	4	36
	Entre pares	33.89	2.99	33
BF-C	Experto	29.98	5.72	30
	Entre pares	31.50	5.19	32
BF-N	Experto	19.90	5.12	19
	Entre pares	21.65	5.88	22
BF-O	Experto	34.27	3.82	34
	Entre pares	33.80	4.85	32.50

Para comprobar si estas diferencias son significativas, se ha intentado aplicar un t-test a cada uno de los componentes de la personalidad.

Las primeras tres asunciones de este test tienen que ver con el diseño del estudio: existencia de una variable continua dependiente (en este caso los cinco componentes de la personalidad), una variable independiente categórica con dos grupos (experto/pares) e independencia de observaciones.

Las otras tres asunciones se relacionan con la naturaleza de los datos, y se detallan a continuación:

- No debería haber valores extremos en los dos grupos de la variable independiente en términos de la variable dependiente. En las Figuras 1, 2, 3, 4 y 5 puede verse cómo sí existen valores extremos. Sin embargo, una revisión de los datos mostró que eran valores genuinamente extremos, por lo que se dejaron en el análisis.
- La variable dependiente debería estar distribuida de manera aproximadamente normal para cada grupo de la variable independiente. Para comprobar esta asunción, se ha aplicado la prueba de Shapiro-Wilk, que ha arrojado valores de  $p < 0.05$  para todos los grupos excepto la amabilidad de los que prefieren una evaluación entre pares. Dado que existía un sesgo moderado de los datos hacia la derecha, se ha aplicado una transformación de raíz cuadrada, que no ha conseguido subsanar el problema. Es por ello que se ha optado por aplicar un test de Mann-Whitney en lugar del T-Test, que requiere que las distribuciones de los dos grupos tengan una forma similar. La distribución de las cinco variables para experto y pares no son similares, tal y como puede verse con una inspección visual de las Figuras 6, 7, 8, 9 y 10, por lo que en lugar de testar las medianas testamos los rangos promedio.

El resultado de aplicar el test muestra cómo los rangos promedio son significativamente distintos para N, con experto=64.17 y pares=83.43 ( $U=2757$ ,  $z=2.648$ ,  $p=0.008$ ) y A, con experto=76.13 y pares=59 ( $U=1633$ ,  $z=-2.36$ ,  $p=0.018$ ), mientras que para E ( $U=2231$ ,  $z=0.307$ ,  $p=0.759$ ), O ( $U=1874$ ,  $z=-1.28$ ,  $p=0.20$ ) y C ( $U=2466$ ,  $z=1.352$ ,  $p=0.176$ ) los rangos promedio no muestran diferencias significativas.

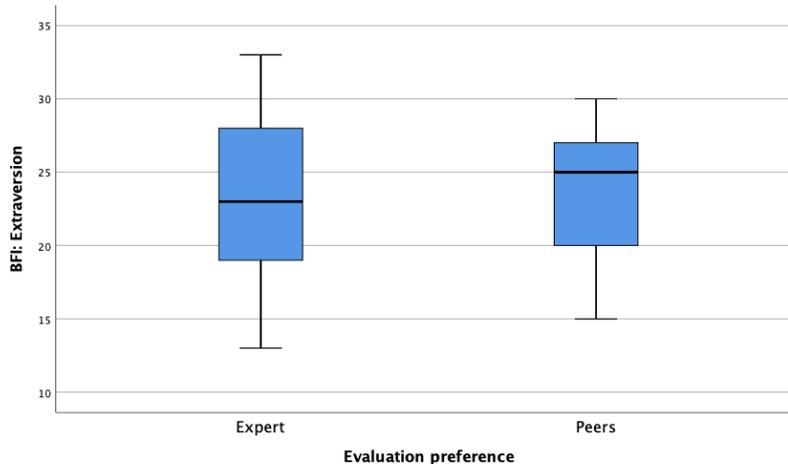


Figura 4: Diagrama de Caja de la variable Extraversión en función de la preferencia de evaluación del alumno

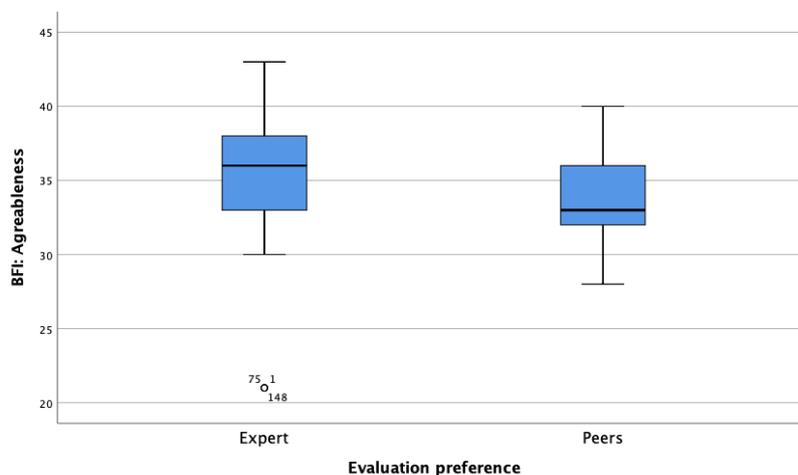


Figura 5: Diagrama de Caja de la variable Amabilidad en función de la preferencia de evaluación del alumno

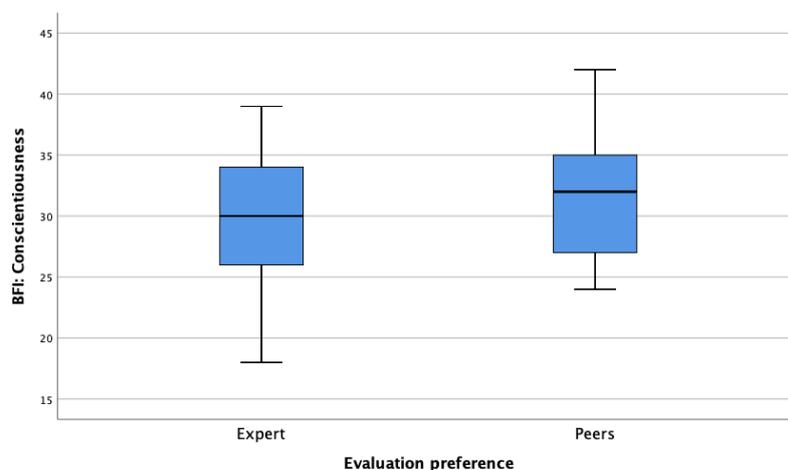


Figura 6: Diagrama de Caja de la variable Escrupulosidad en función de la preferencia de evaluación del alumno

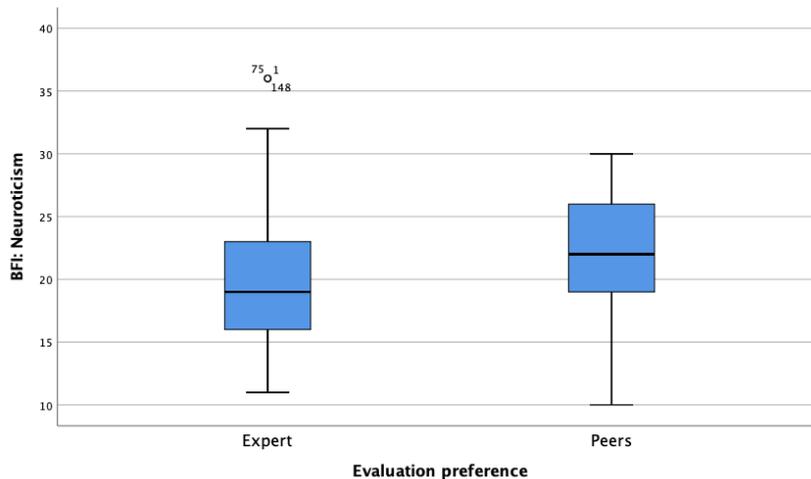


Figura 7: Diagrama de Caja de la variable Neuroticismo en función de la preferencia de evaluación del alumno

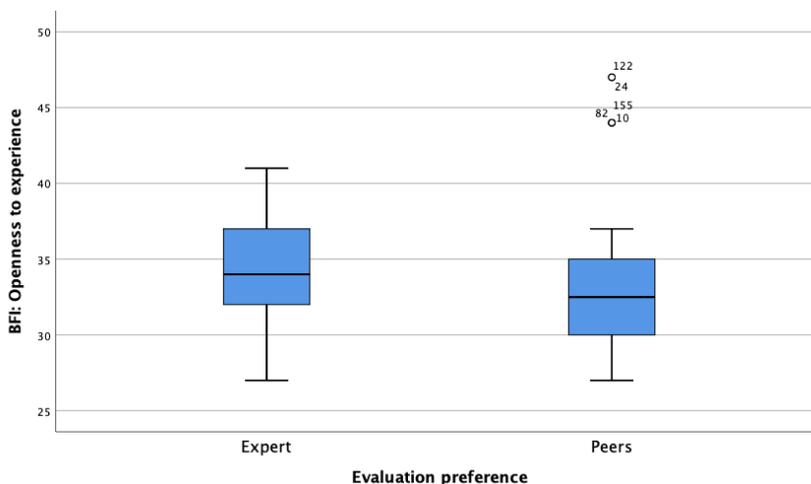


Figura 8: Diagrama de Caja de la variable Apertura a la Experiencia en función de la preferencia de evaluación del alumno

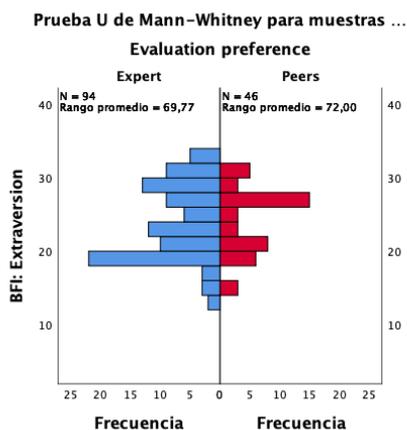


Figura 9: Comparativa distribuciones y rango promedio de la variable Extraversión

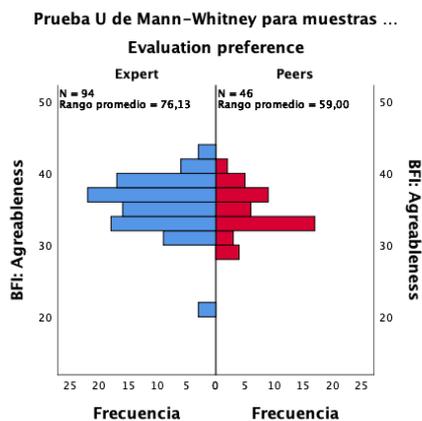


Figura 10: Comparativa distribuciones y rango promedio de la variable Amabilidad

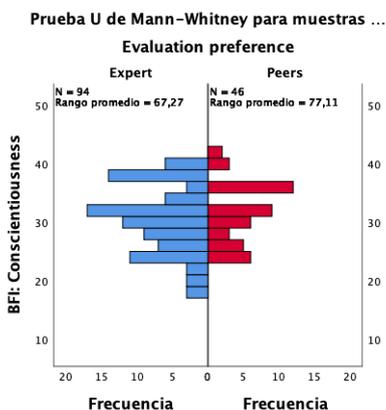


Figura 11: Comparativa distribuciones y rango promedio de la variable Escrupulosidad

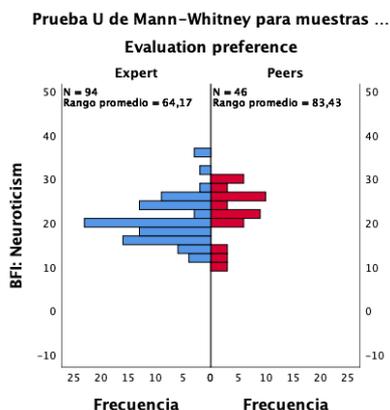


Figura 12: Comparativa distribuciones y rango promedio de la variable Neuroticismo

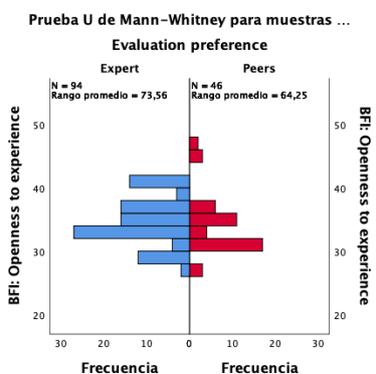


Figura 13: Comparativa distribuciones y rango promedio de la variable Apertura a la Experiencia

## 3.2. Intención de Adopción de las distintas modalidades

La Tabla 3 muestra los estadísticos descriptivos para las variables de intención de adopción (U, EoU, Comp e IA global). La utilidad percibida ha sido ligeramente mayor para la modalidad de parejas, así como la facilidad percibida. En cuanto a la compatibilidad, los alumnos parecen sentirse ligeramente mejor evaluando con compañeros. En conjunto, la IA global también favorece ligeramente la modalidad de parejas.

Tabla 3. Estadísticos descriptivos IA

Variable	Modalidad	Media	SD	Mediana
UMAMQ-U	1	34.87	9.03	34.50
	2	37.24	7.51	38.50
	3	35.34	8.46	35
UMAMQ-EoU	1	32.56	6.33	33
	2	33.74	5.72	34
	3	32.96	7.01	33
UMAMQ-Comp	1	32.19	8.92	31.50
	2	33.12	7.88	33
	3	33.58	8.01	33
UMAMQ-IA	1	31.37	10.59	31
	2	32.50	8.76	32
	3	32.53	9.66	31

Para comprobar si existen diferencias significativas entre las distintas modalidades de evaluación entre pares se ha aplicado un one-way ANOVA. Este test tiene tres asunciones que se refieren al diseño del estudio: (a) una variable continua dependiente (las distintas dimensiones de la IA), (b) una variable dependiente categórica con dos o más valores (en este caso la modalidad, con tres valores: individual, parejas, tríos) y (c) existe independencia de observaciones. Las otras tres asunciones se refieren a cómo los datos encajan en el modelo one-way ANOVA, y son:

- No debería haber valores extremos en los grupos de la variable independiente en términos de la variable dependiente. En las Figuras 11, 12, 13 y 14, puede verse cómo sí existen valores extremos. Sin embargo, una revisión de los datos mostró que eran valores genuinamente extremos, por lo que se dejaron en el análisis
- La variable dependiente debería estar distribuida de manera aproximadamente normal para cada grupo de la variable independiente. Para comprobar esta asunción, se ha aplicado la prueba de Shapiro-Wilk, que ha arrojado valores de  $p > 0.05$  para todos los grupos.
- La varianza debería ser igual en cada grupo de la variable independiente. Para comprobar esta asunción se ha aplicado el test de igualdad de varianzas de Levene, que nuevamente ha arrojado valores de  $p > 0.05$  para todos los grupos.

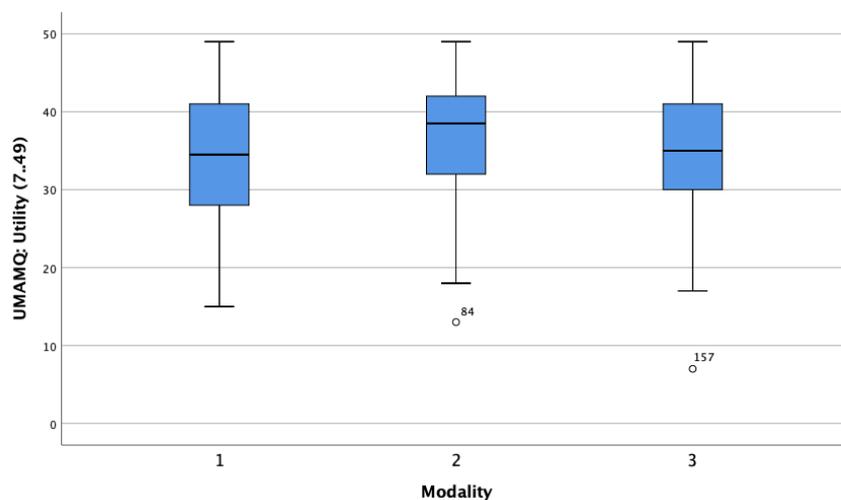


Figura 14: Gráfico de cajas de la variable Utilidad en función de la modalidad (individual, parejas, tríos)

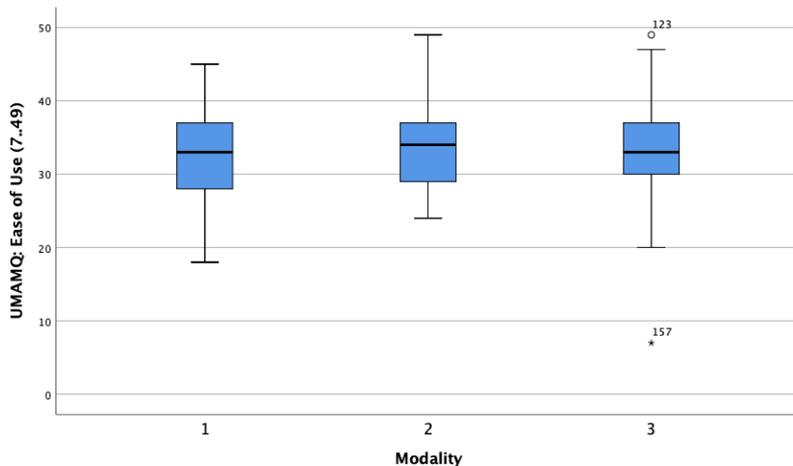


Figura 15: Gráfico de cajas de la variable Facilidad de Uso en función de la modalidad (individual, parejas, tríos)

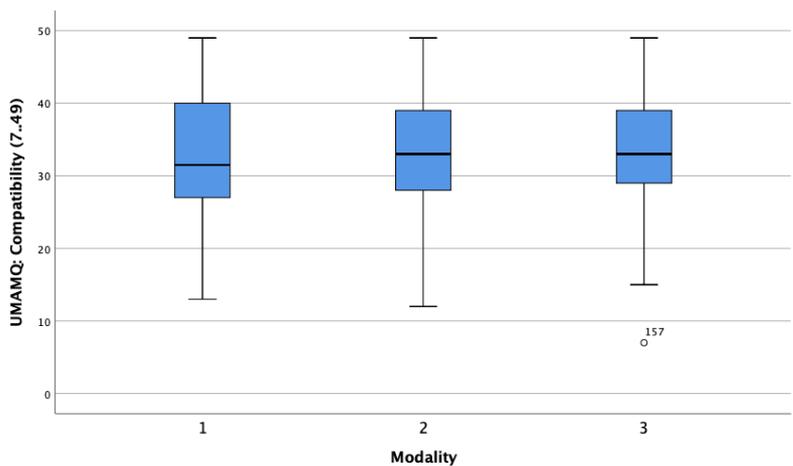


Figura 16: Gráfico de cajas de la variable Compatibilidad en función de la modalidad (individual, parejas, tríos)

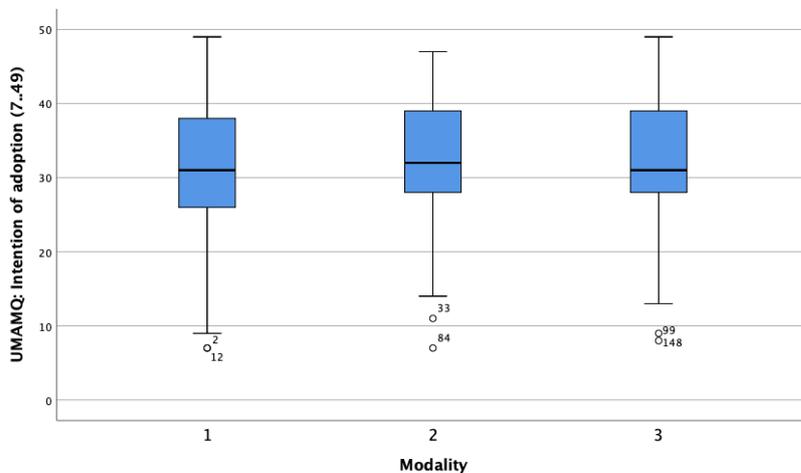


Figura 17: Gráfico de cajas de la variable Intención de Adopción global en función de la modalidad (individual, parejas, tríos)

Un vez comprobadas las asunciones, los resultados de aplicar el test muestran que, a pesar de que los alumnos muestran una ligera inclinación a preferir la evaluación en parejas, no existen diferencias significativas entre la intención de adopción de las tres modalidades,  $F_U(2,162)=1.28$ ,  $p=0.281$ ,  $F_{EoU}(2,162)=0.506$ ,  $p=0.604$ ,  $F_{Comp}(2,162)=0.399$ ,  $p=0.672$ ,  $F_{IA}(2,162)=0.254$ ,  $p=0.776$ .

## 4. Conclusiones

Los análisis realizados muestran cómo las personas más amables muestran más reparos a evaluar a sus compañeros. Esto tiene sentido, ya que estas personas pueden percibir la evaluación como una crítica al trabajo de otros. Por otro lado, las personas más neuróticas prefieren la evaluación entre pares, lo que se puede explicar si pensamos que estas personas, más desconfiadas por naturaleza, pueden sentirse más cómodas si son evaluadas por sus compañeros en lugar de serlo por una persona que perciben como con más autoridad. La extraversión, la escrupulosidad y la apertura a la experiencia no muestran diferencias significativas.

Por lo que respecta a la intención de adopción, los resultados muestran que no existen diferencias significativas a la hora de elegir una u otra modalidad de evaluación entre pares.

Como trabajo futuro se pretende analizar si la preferencia entre evaluación individual/entre pares se ve influenciada por la personalidad. También se pretende comprobar si la escrupulosidad está relacionada con la precisión de la evaluación entre pares en la modalidad individual.

## 5. Tareas desarrolladas en la red

Tabla 4. Tareas desarrolladas en la red

Participante de la red	Tareas que desarrolla
Cristina Cachero Castro	Coordinación de la red, estado del arte, configuración de los instrumentos en Qualtrics, análisis y redacción
Juan Ramón Rico Juan	Limpieza de datos, análisis y redacción
Hermenegilda Macià Soler	Recopilación de datos, limpieza de datos, análisis y redacción
Santiago Meliá Beigbeder	Estado del arte y redacción
Sergio Luján Mora	Estado del arte y redacción
Pilar Barra Hernández	Estado del arte y redacción
Mayo de Juan Vigaray	Estado del arte y redacción

## 6. Referencias bibliográficas

Adachi, C., Hong-Meng Tai, J., & Dawson, P. (2018). Academics' perceptions of the benefits and challenges of self and peer assessment in higher education. *Assessment and Evaluation in Higher Education*, 43 , 294–306.

An, D., & Carr, M. (2017). Learning styles theory fails to explain learning and achievement: Recommendations for alternative approaches. *Personality and Individual Differences*, 116 , 410–416.

(APA), A. P. A. (2017). Personality. [http://www.apa.org/topics/ 530 personality/](http://www.apa.org/topics/530/personality/). [Online; last accessed March 2021].

Balijepally, V., Mahapatra, R., & Nerur, S. P. (2006). Assessing personality profiles of software developers in agile development teams. *Communications of the Association for Information Systems*, 18 , 4.

Chang, C.-Y., Lee, D.-C., Tang, K.-Y., & Hwang, G.-J. (2021). Effect sizes and research directions of peer assessments: From an integrated perspective of meta-analysis and co-citation network. *Computers & Education*, 164 , 104123.

Curtis, B., Krasner, H., & Iscoe, N. (1988). A field study of the software design process for large systems. *Communications of the ACM* , 31 , 1268–1287.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, (pp. 319–340).

Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note: how does personality matter? relating the five-factor model to technology acceptance 575 and use. *Information Systems Research*, 19 , 93–105.

Diéguez, M., Sepúlveda, S., & Cachero, C. (2012). Umam-q: An instrument to assess the intention to use software development methodologies. In *Information Systems and Technologies (CISTI), 7th Iberian Conference on* (pp. 1–6). IEEE.

Feldt, R., Angelis, L., Torkar, R., & Samuelsson, M. (2010). Links between the personalities, views and attitudes of software engineers. *Information and Software Technology*, 52 , 611–624.

Gallivan, M. J. (2004). Examining it professionals' adaptation to technological change: the influence of gender and personal attributes. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* , 35 , 28–49.

Jarillo-Nieto, P. I., Enríquez-Ramírez, C., & Sánchez-Herrera, R. A. (2015). Identificación del factor humano en el seguimiento de procesos de software en un medio ambiente universitario. *Computación y Sistemas*, 19, 577–588.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, 3, 114–158.

Li, H., Xiong, Y., Hunter, C. V., Guo, X., & Tywoniw, R. (2020). Does peer assessment promote student learning? a meta-analysis. *Assessment & Evaluation in Higher Education*, 45, 193–211.

Liu, N.-F., & Carless, D. (2016). Peer feedback: the learning element of peer assessment. *Teaching in Higher education*, 11, 279–290.

Martínez, Y., Cachero, C., & Meliá, S. (2013). Mdd vs. traditional software development: A practitioner's subjective perspective. *Information and Software Technology*, 55, 189–200.

Mazur, E. (1997). *Peer instruction: A user's manual*. Series in Educational Innovation. Prentice Hall.

O'Connor, R. V., & Yilmaz, M. (2015). Exploring the belief systems of software development professionals. *Cybernetics and Systems*, 46, 528–542.

Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment and Evaluation in Higher Education*, 44, 379–397.

Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*, 60, 154–165.

Rico-Juan, J. R., Gallego, A.-J., & Calvo-Zaragoza, J. (2019). Automatic detection of inconsistencies between numerical scores and textual feedback in peer-assessment processes with machine learning. *Computers & Education*, 140, 103609.

Ter Laak, J. J. (1996). Las cinco grandes dimensiones de la personalidad. *Revista de psicología*, 14, 129–181.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Zheng, L., Zhang, X., & Cui, P. (2020). The role of technology-facilitated peer assessment and supporting strategies: a meta-analysis. *Assessment and Evaluation in Higher Education*, 45, 372–386.