
El feedback y la experiencia evaluando como factores determinantes en la autorregulación de los estudiantes

Feedback and evaluative experience as decisive factors in student self-regulation

反馈和评估的经验作为学生自我调节的决定因素

Обратная связь и опыт оценивания как факторы, определяющие саморегуляцию студентов

Samuel Parra León

Universidad de Jaén

sparra@ujaen.es

<https://orcid.org/0000-0002-6980-2680>

Inmaculada García-Martínez

Universidad de Almería

imartin@ual.es

<https://orcid.org/0000-0003-2620-5779>

Fechas · Dates

Recibido: 2021/04/30

Aceptado: 2021/06/10

Publicado: 2021/06/30

Cómo citar este trabajo · How to Cite this Paper

Parra, S., & García-Martínez, I. (2021). El feedback y la experiencia evaluando como factores determinantes en la autorregulación de los estudiantes. *Publicaciones*, 51(1), 287–301. <https://doi.org/10.30827/publicaciones.v51i1.20738>

Resumen

La evaluación tiene un papel muy importante en los procesos de aprendizaje en contextos educativos. Así, las evaluaciones activas, donde el alumno tiene un papel protagonista, cada vez están tomando mayor presencia como complemento a las evaluaciones tradicionales debido a su probado valor formativo. El presente estudio trata de analizar el impacto que tiene sobre las evaluaciones hechas por los alumnos, la experiencia adquirida por dichos alumnos al someterlos a situaciones de evaluación y al recibir feedback por parte del profesor sobre dichas evaluaciones. Para ello, dos grupos naturales de clase realizaron evaluaciones del trabajo realizados por otros grupos y de su propio trabajo. La diferencia entre los grupos radicó en que uno de ellos recibió feedback del profesor tras la evaluación y el otro grupo no. Los resultados analizaron la precisión de los grupos evaluando a otros grupos de compañeros y evaluándose a sí mismo, con respecto a la evaluación del profesor. Los datos mostraron que ambos grupos mejoraron la precisión de las evaluaciones a lo largo de las sesiones, pero la mejora fue significativamente mayor en el caso del grupo que recibió feedback. Estos resultados nos permiten apuntar sobre la importancia de la evaluación dentro del proceso de aprendizaje y de la necesidad de proveer a los alumnos de feedback sobre la evaluación realizada.

Palabras clave: Autoevaluación, Feedback, autorregulación, universitarios.

Abstract

Assessment plays a key role in learning processes in educational environments. Thus, active evaluations, such as self-evaluation, are being used increasingly frequently as a supplement to traditional evaluations due to their proven formative value. The present study analyzes the impact of the evaluations made by the students on the experience acquired by these students when they are subjected to evaluation situations and when they receive feedback from the teacher on these evaluations. To this end, two natural class groups carried out evaluations of the work done by other groups and of their own work. The difference between the groups was that one group received feedback from the teacher after the assessment and the other group did not. The results were analyzed to determine the groups' accuracy in evaluating their peers and themselves with respect to the teacher's evaluation. The data showed that both groups improved the accuracy of the evaluations throughout the sessions, with a significantly greater improvement for the group that received feedback. These results highlight the importance of evaluation in the learning process and the need to provide students with feedback on their evaluation.

Keywords: Self-assessment, Feedback, self-regulation, university students.

概要

评估在教育里的学习过程中起着非常重要的作用。学生起主导作用的主动评估因其已证明的教育价值而越来越多地被作为传统评估的补充。本研究的目的是分析学生所作的评估,其置身于评价情境中获得的经验和从老师那边获得反馈的影响。为此,由两个自然班级小组对其他小组开展的工作和自己的工作进行了评估。两组之间的区别在于,其中一组在评估后收到了老师的反馈,而另一组则没有。结果分析了相对于教师的评价,各组评价其他同龄人组和评价自己的精度。数据显示,两组在整个过程中都提高了评估的准确性,但在收到反馈的情况下,改进幅度更大。这些结果使我们能够指出评估在学习过程中的重要性,以及向学生提供评估反馈的必要性。

关键词: 自我评价, 反馈, 自我调节, 大学生.

Аннотация

Оценивание играет очень важную роль в процессе обучения в образовательных контекстах. Таким образом, активное оценивание, в котором учащийся играет ведущую роль, все чаще используется в качестве дополнения к традиционному оцениванию благодаря своей доказанной формирующей ценности. Целью данного исследования является анализ влияния опыта, приобретенного учащимися, когда они подвергаются ситуациям оценивания и когда они получают обратную связь от учителя по этим оценкам, на оценки, сделанные учащимися. Для этого две естественные классные группы провели оценку работы, выполненной другими группами, и своей собственной работы. Разница между группами заключалась в том, что одна группа получила обратную связь от учителя после оценки, а другая - нет. В результатах анализировалась точность оценки группами других групп сверстников и оценки самих себя по отношению к оценке учителя. Данные показали, что обе группы улучшили точность оценок в ходе сеансов, но улучшение было значительно больше в группе, которая получила обратную связь. Эти результаты указывают на важность оценки в процессе обучения и необходимость предоставления студентам обратной связи по их оценке.

Ключевые слова: Самооценка, обратная связь, саморегуляция, студенты университета.

Introducción

La educación universitaria ha cambiado sustancialmente desde la entrada de los nuevos planes de Grado. Estos cambios no solo han afectado al diseño de los procesos de enseñanza y aprendizaje y a su instrucción, sino también al papel que desempeña el alumno y el anhelo por conseguir mayor autonomía en él (Vicario-Molina et al., 2020). La evaluación tradicional está siendo paulatinamente complementada por modalidades alternativas de evaluación, donde el alumnado asume un papel central como agente activo de su proceso de enseñanza y aprendizaje (Sáiz & Gómez, 2020).

En este panorama, las evaluaciones activas, donde el alumno tiene un papel protagonista en la evaluación, están tomando gran presencia. Dentro de estas evaluaciones activas la autoevaluación del alumnado (SSA) se está posicionando como una modalidad de evaluación complementaria a la del profesorado, que contribuye a la mejora del aprendizaje del alumnado (Panadero et al., 2016). Sin embargo, las reticencias de un amplio sector del profesorado para implementar esta modalidad de evaluación en sus clases, sumado al desconocimiento sobre cuáles son las condiciones más idóneas para su desarrollo hacen a que no se termine de consolidar esta propuesta educativa (Kambourova, 2020).

La autoevaluación del estudiante (SSA) puede definirse como ese proceso en el que el alumno valora la calidad de sus logros y ganancias obtenidas a lo largo de la construcción de su proceso de aprendizaje (Brown & Harris, 2014). La autoevaluación se ha identificado como una vía de mejora cuantitativa y cualitativamente el aprendizaje del alumnado, al ofrecer oportunidades para que el alumnado reflexione sobre lo que va aprendiendo, especialmente cuando existe una retroalimentación (Andrade, 2019). Aunque existen diferentes tipologías dentro de la autoevaluación del alumnado (Panadero et al., 2016), en función de su finalidad, el papel del profesor o los instrumentos utilizados, es aconsejable la implantación de las modalidades más formativas en detrimento de sumativas, debido, entre otras causas, a las posibilidades didácticas que

ofrecen para el aprendizaje (Siegesmund, 2017). La diferencia entre autoevaluación formativa o sumativa recae en el propósito con el que este tipo de evaluación plantea. Mientras que en la autoevaluación sumativa el propósito es conocer si el alumno es capaz de evaluar sus logros en una tarea de manera similar a como lo haría un profesor, en el caso de la autoevaluación formativa esta tiene un propósito orientado al aprendizaje, de manera que las estrategias que se usan en este proceso están relacionadas con la autorregulación del proceso de aprendizaje del alumnos, la interiorización de los aprendido, reconocimiento de los errores y procedimientos emergentes de reconstrucción cognitiva (Wanner & Palmer, 2018).

En esta línea, proporcionar al alumnado un instrumento donde se reflejen los criterios o estándares de evaluación con los que será evaluado facilita el ejercicio reflexivo que ha de hacer a la hora de atribuir valor a sus actuaciones y desempeño (Abella-García et al., 2020). De igual modo, este guion o instrumento con el que contrasta sus producciones con "modelos ideales" implica brindarle oportunidades para que haga los ajustes necesarios y despliegue las estrategias necesarias para mejorar su experiencia de aprendizaje (Andrade, 2019). Otras modalidades de evaluaciones activas estrechamente relacionadas con la autoevaluación son la evaluación por pares (Peer-assessment) o la evaluación colaborativa (Co-assesment).

El efecto positivo que tiene sobre el aprendizaje someter a los alumnos a una situación de evaluación o prueba es un efecto robusto ampliamente estudiado por la literatura de la psicología cognitiva del aprendizaje. Recientemente, Yang et al. (2021), han realizado un meta-análisis donde evalúan el efecto que suponen sobre el rendimiento académico someter a los alumnos a una prueba frente a otras condiciones en contextos académicos. Este estudio concluye que someter a los alumnos a una situación de evaluación aumenta el rendimiento académico de manera significativa ($g = .449$).

Además de los beneficios propios de la evaluación ya indicados, diferentes autores señalan que tanto la evaluación por pares como la evaluación cooperativa realizada durante el proceso de aprendizaje, contribuyen al desarrollo de las habilidades de autoevaluación de los estudiantes. Concretamente, Somervell (1993), Topping (1998) y Vickerman (2009), afirman que los resultados de evaluar a otros estudiantes pueden considerarse parte de la autoevaluación y pueden contribuir al desarrollo de las habilidades de autoevaluación.

En la búsqueda de un mayor aprendizaje autónomo en el contexto universitario, autoevaluación del estudiante y aprendizaje autorregulado (SRL) se unifican dentro del proceso de enseñanza y aprendizaje, evidenciándose una relación de interdependencia clara entre ellos (Panadero et al., 2017). En este prisma, el aprendizaje autorregulado se define como un proceso en el que el alumnado, fruto de la retroalimentación recibida, es consciente de su progreso de aprendizaje y reflexiona para continuar aprendiendo. En este proceso se identifican tres fases cíclicas. La primera se corresponde con la previsión inicial, donde se establecen los objetivos que se pretenden alcanzar en el desarrollo de las tareas a realizar en las sucesivas fases. A continuación, se sitúa la fase de ejecución, donde el alumnado ha de realizar las tareas propuestas, a través de la aplicación de estrategias de aprendizaje y, finalmente, la autorreflexión donde comparan sus producciones con los estándares establecidos y valoran el aprendizaje alcanzado. Así existe una fuerte relación entre autorregulación y autoevaluación (Panadero & Alonso-Tapia, 2013). La autoevaluación se presenta como un proceso clave asociado a la autorregulación (Panadero et al., 2017; Paris & Paris, 2001). En este proceso, el alumnado pone en práctica una serie de estrategias y recursos, donde dirige sus esfuerzos para optimizar mejor lo que aprende, al tiempo que considera sus

posibilidades y aspectos a mejorar (Larruzea-Urkixo & Ramírez, 2020). De esta manera, la transición desde un aprendizaje tradicional hacia un aprendizaje autorregulado implica habilidades metacognitivas como la autoconciencia y reflexión en el uso de estrategias dentro de contextos de aprendizaje personalizados, en los que la motivación y la socialización ostentan aspectos claves en el análisis del desempeño de los estudiantes (Daura, 2017). La toma paulatina de consciencia necesaria para desempeñar satisfactoriamente la autoevaluación se convierte en un componente decisivo en el logro de aprendizaje autorregulado por parte de los estudiantes (Panadero et al., 2016).

A su vez, el feedback se posiciona como un factor crucial en el análisis de la efectividad de SSA y SRL (Hawe & Dixon, 2017). Empoderar al alumnado y proporcionarle retroalimentación sobre cuál es su estado real de aprendizaje, favorece el desarrollo de estrategias reguladoras para consolidar aprendizajes anteriores y producir andamiajes en la adquisición de los nuevos (Panadero & Alonso-Tapia, 2013). No solo basta con que el alumnado sea capaz de calificar su trabajo sin un cuestionamiento lógico. Por el contrario, se requiere una orientación por parte de un experto o un instrumento que lo oriente y aporte solidez a los juicios que ha de emitir sobre su propio desempeño. Desde esta mirada, el feedback es concebido como un elemento central en el proceso de instrucción, más allá de ser una forma de contrastar si ha aprendido los conocimientos susceptibles de ser enseñados (Ibarra-Sáiz et al., 2020). Solo bajo esta consideración será posible concebir el aprendizaje autorregulado como esa aspiración a alcanzar por un alumnado maduro, activo y autónomo, donde su aproximación hacia el conocimiento se nutre directamente de las estrategias que pone en práctica para aproximarse y dominar el conocimiento (Panadero et al., 2017).

Sin embargo, estudios recientes señalan que la retroalimentación no siempre culmina en mejoras en el aprendizaje del alumnado (van der Kleij, 2017). Por ejemplo, el enfoque que adopte el profesorado a la hora de realizar esa retroalimentación va a ser decisivo. Sugerencias y comentarios más constructivos y estratégicos, tienden a ser más efectivos que los que se posicionan en una perspectiva más correctiva (van der Kleij et al., 2017). Del mismo modo, el rol que asuma el alumnado dentro de su proceso de aprendizaje va a determinar la eficacia de esa retroalimentación. De esta manera, para los alumnos pasivos que no centran sus esfuerzos en redirigir sus estrategias de aprendizaje en base a la retroalimentación aportada por el profesorado, el efecto de la retroalimentación será nulo o escaso, a diferencia de aquellos que adopten un papel más activo. Asimismo, la utilidad con la que perciba los comentarios recibidos por el profesor determinará que los incorpore o no a sus estrategias de aprendizaje (Jonsson et al., 2018). La familiaridad del alumnado con los criterios con los que se le evaluará desde el comienzo de la instrucción, contribuirá a reducir el impacto negativo de estos factores (Panadero & Alonso-Tapia, 2013).

Finalmente, el tipo de instrumento seleccionado para ofrecer retroalimentación a los alumnos y guiarlos en los procesos de autoevaluación cobran un importante papel (Panadero et al., 2017). En relación con esto último, la rúbrica surge como un instrumento plausible para garantizar la retroalimentación que el alumnado demanda en aras de avanzar en su proceso de aprendizaje de acuerdo a lo hallado en la literatura (Pui et al., 2020; Yan, 2020).

Teniendo en cuenta lo mencionado anteriormente, el objetivo del presente estudio es analizar como las evaluaciones grupales de los alumnos sobre el desempeño de los trabajos presentados en clase y sobre su propio desempeño mejora gracias a su experiencia previa de autoevaluación y el feedback recibido tras la evaluación. Para examinar esta cuestión, dos grupos serán sometidos a situaciones que les permitirán ad-

quirir experiencia evaluando. Sin embargo, solo uno de los grupos recibirá feedback del profesor experto tras las evaluaciones. Con el objetivo de favorecer las condiciones para la evaluación, se facilitó a los alumnos la misma rúbrica que usó el profesor para las evaluaciones, haciendo así que tanto alumnos como profesor conozcan los criterios de evaluación de antemano.

Método

Participantes

Un grupo natural de clase constituido por 71 estudiantes del 2º curso del grado de educación social participó en este estudio. La edad de los estudiantes estuvo comprendida entre 20 y 39 años ($M = 22.48$). Del total de estudiante 88.73% fueron mujeres y 11.26% fueron hombres. Estos porcentajes son proporcionales a la distribución de hombres y mujeres en la población total de estudiantes en España (Instituto Nacional de Estadística [INE], 2015). El grupo de clase estaba dividido en dos subgrupos para cada uno de los turnos de prácticas de la asignatura. El Grupo que recibirá feedback (grupo FB) con 35 alumnos y el Grupo que no recibirá feedback (grupo N-FB) con 36 alumnos.

Instrumentos

Para unificar criterios en la evaluación de los trabajos presentados, el profesor creó y facilitó a los alumnos una rúbrica. En esta rúbrica se establecían diferentes niveles de logros para cada uno de los apartados del informe de diagnóstico, así como el porcentaje máximo de la nota final que se podía obtener en cada apartado del informe. Las partes del informe de diagnóstico junto con los porcentajes sobre la nota final fueron: 1) Datos personales, escolares y familiares y 2) Antecedentes (punto 1 y 2 el 10% del total), 3) Motivos/justificación de la evaluación (20%), 4) Técnicas e instrumentos aplicados (20%), 5) Resultados obtenidos (10%), 6) Análisis y valoración de los resultados (20%), 7) Conclusiones, orientaciones y recomendaciones (20%). La rúbrica fue evaluada y usada por 5 docentes independientes y mostró un coeficiente de correlación intraclase de $r = .962$.

Procedimiento

Este estudio se realizó a lo largo de un cuatrimestre durante las prácticas de una asignatura del grado de educación social (Diagnóstico y Evaluación en Educación Social). En el estudio, se siguieron los estándares éticos que guían la investigación con personas, según la Declaración de Helsinki (WMA, 2009). Al comienzo del curso, el profesor de la asignatura dejó que los alumnos se inscribieran libremente a cada uno de los turnos de prácticas, con la única restricción de que los grupos no podían diferir en más de 4 alumnos. Esta inscripción se realizó por medio de la plataforma oficial de la universidad y los alumnos debían de inscribirse en uno de los dos subgrupos dentro del dominio de la asignatura. Esta división de la clase en dos subgrupos nos permitió crear nuestros dos grupos para el estudio. En la primera sesión de los seminarios prácticos, el profesor de la asignatura (el primer autor) explicó el procedimiento a seguir a lo largo de las prácticas. Se indicó a los alumnos que debían hacer grupos de no más de 4 alumnos, y que estos grupos trabajarían juntos a lo largo de todas las prácticas.

Tanto el Grupo FB como el Grupo N-FB contaba de 9 grupos de alumnos. A lo largo de las prácticas estos grupos debían de desarrollar un informe de diagnóstico basado en diferentes casos propuestos que tendrían que exponer a toda la clase al final del curso. A través de las diferentes sesiones el profesor ofrecía contenidos para que los alumnos pudieran aplicar y desarrollar los casos escogidos. En las semanas finales del curso, los diferentes grupos debían exponer a la clase el informe de diagnóstico desarrollado. Previa a la fase de exposición, los alumnos recibieron una sesión en la que el profesor indicó los criterios de evaluación que se usarían a la hora de evaluar los diferentes apartados del informe de diagnóstico, y los puntos que se podrían ganar con cada una de las partes del informe. Estos criterios y ponderaciones de las notas para cada apartado del informe se plasmaron en una rúbrica que el profesor facilitó a los alumnos mediante la plataforma de docencia virtual. Así todos conocían cuales serían los criterios y los pesos en la evaluación. Con el objetivo de motivar a los alumnos a ser lo más precisos posible en sus evaluaciones, se indicó que los 2 grupos más precisos recibirían 0.5 sobre la nota final de prácticas.

Durante las semanas de exposición, cada grupo tenía que hacer una evaluación de los informes de diagnóstico presentados por los diferentes grupos, y además deberían hacer una evaluación de su propio informe. Para hacer estas evaluaciones se pidió que usaran la rúbrica que se había facilitado en sesiones previas. Cada grupo contó con una sesión para la exposición, haciendo tres sesiones de exposición a la semana. El orden de exposición de los grupos se estableció al azar. Cada grupo de alumnos contó con 35 minutos para hacer la exposición del caso elegido y del informe diagnóstico realizado. Tras la exposición los grupos oyentes, que estaban evaluando, tuvieron 15 minutos para plantear preguntas al grupo que había hecho la exposición. Después, todos los grupos entregaron la rúbrica al profesor. El profesor al final de cada sesión hacía la evaluación del grupo que había realizado la exposición. Esta evaluación no se hizo pública hasta el final del curso.

La única diferencia que hubo entre el Grupo FB y el Grupo N-FB, consistió en el papel del profesor al final de la exposición de cada grupo. En el Grupo FB el profesor una vez recogidas las evaluaciones de los grupos, ofrecía feedback a los grupos sobre los puntos fuertes y los puntos débiles en cada uno de los apartados del informe de diagnóstico. Con este feedback el profesor de manera concreta indicaba a los alumnos no solamente las debilidades encontradas en el trabajo, sino qué información deberían haber incluido en cada uno de los apartados del trabajo para que hubiera alcanzado el nivel más alto de logro en la rúbrica, en base a los contenidos vistos en las diferentes sesiones de prácticas. La elección del grupo que recibió el feedback del profesor fue al azar.

Análisis de los datos

El α para todos los análisis realizados se fijó en .05. Todos los análisis se realizaron con el software jamovi (The jamovi Project.). Se recogieron todas las valoraciones que los grupos hicieron sobre las defensas del informe (GA, Group Assessment) y sobre su propia evaluación (GSA, Group Self-Assessment). Además, se recogió la evaluación que el profesor hizo de cada grupo (Exp, Expert). Para calcular la precisión tanto de las evaluaciones grupales como de SSA, se calculó restando estas evaluaciones (GA y GSA) con la evaluación emitida por el profesor (Exp). La habilidad de los grupos al evaluar el informe de diagnóstico de otros grupos se llamó Group Assessment-Skill (GA-Skill), y a la habilidad de los grupos de autoevaluar su propio informe se llamó Group

Self-Assessment skill (GSA-skill). Puntuaciones positivas en estas variables indicarán sobreestimación, y puntuaciones negativas indicarán subestimación, las puntuaciones cercanas a 0 serán las más precisas. Para analizar el desarrollo de las diferentes evaluaciones a lo largo de las sesiones de prácticas se realizó un ANOVA de medidas repetidas con las evaluaciones de todos los grupos para el grupo que exponía el informe de diagnóstico (GA-Skill) a lo largo de los 9 días de exposición. De manera adicional, se analizó la evaluación que los grupos realizaban sobre su propio informe (GSA-Skill) a lo largo de las 9 sesiones. Se presenta *generalized eta squared* (η^2_g) para medir el tamaño de los efectos.

Resultados

La Tabla 1 muestra los resultados descriptivos para las evaluaciones promedio de todos los grupos de cada clase a lo largo de las 9 sesiones de exposición.

Tabla 1

Descriptivos para la variable Group Assessment-Skill (GA-Skill) por sesión

		S1	S2	S3	S4	S5	S6	S7	S8	S9
FB	<i>Mean</i>	2.12	1.89	1.16	.87	.22	.12	-.42	-.51	-.09
	<i>SD</i>	.64	.67	.85	.80	.38	.93	.87	1.08	.97
	<i>Median</i>	2.50	1.90	1.00	1.30	.15	.20	-.30	-.30	-.30
N-FB	<i>Mean</i>	2.58	2.41	2.29	2.49	1.86	1.40	1.77	1.26	.91
	<i>SD</i>	1.25	.86	1.12	.86	1.06	1.08	.68	.84	1.30
	<i>Median</i>	2.65	2.60	2.30	2.60	1.90	1.15	1.60	1.05	1.10

Nota. FB=grupo FeedBack, N-FB grupo no FeedBack. S1 a S9 indica el número de sesión

Estos datos se representan en la Figura 1. Como se puede observar, las autoevaluaciones promedio de los grupos de ambas clases mostraron sobre estimación al comienzo de las sesiones (Group evaluation skill es superior a la evaluación del profesor), y estas evaluaciones fueron descendiendo, acercándose a 0 (volviéndose más precisas) conforme avanzaban las sesiones. No obstante, el grupo FB parecía mostrar un ajuste más rápido que el grupo N-FB.

Para analizar cómo se realizó este proceso de ajuste en la autoevaluación de los grupos a lo largo de las sesiones se realizó una ANOVA 2 (Grupo) x 9 (Sesiones) con la variable GA-Skill. Este análisis mostró un efecto principal de Grupo, $F(1, 17) = 17.04, p < .001, \eta^2_g = .39$, un efecto principal de Sesión, $F(8, 136) = 23.15, p < .001, \eta^2_g = .39$, y una doble interacción Clase x Sesión, $F(8, 136) = 3.41, p < .001, \eta^2_g = .09$. Estos resultados confirmaron que ambos grupos mejoraron en la precisión de sus evaluaciones a lo largo de las sesiones, y que el patrón de mejora en la evaluación fue diferente entre grupos. Concretamente, el grupo FB mostró una mayor precisión en sus evaluaciones en las últimas sesiones con respecto a las realizadas por el grupo N-FB.

Figura 1

Promedio de las evaluaciones grupales (GA-Skill) de cada grupo (FB vs N-FB) a lo largo de las sesiones (S1 a S9). Las barras representan el 95% del Intervalo de Confianza

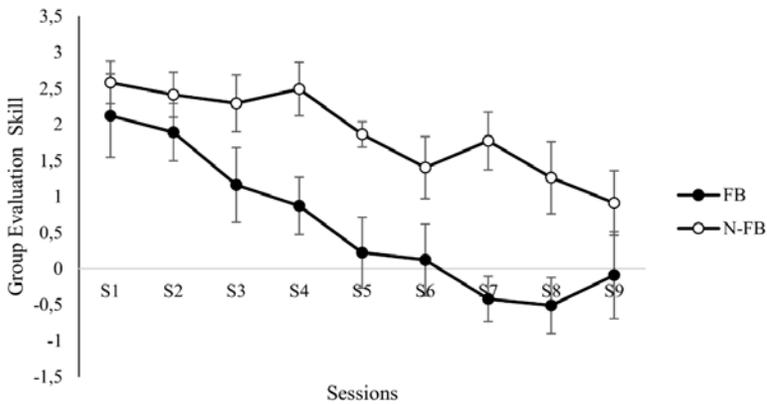
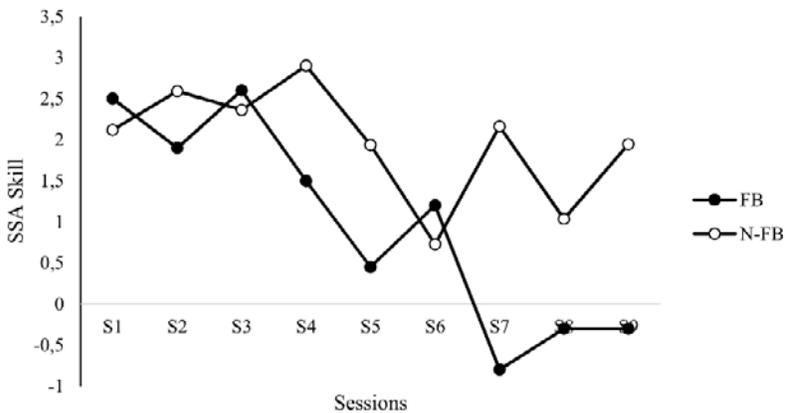


Figura 2

GSA-Skill de los subgrupos para cada grupo de clase a lo largo de las sesiones



La Figura 2, muestra la variable GSA-skill de los subgrupos de trabajo para cada grupo-clase a lo largo de las sesiones de exposición. De manera similar a lo anterior, los grupos al comienzo de las sesiones se autoevaluaban por encima de lo evaluado por el profesor (sobreevaluación) y a lo largo de las sesiones estas autoevaluaciones se fueron mostrando más precisas (más cercanas a la emitida por el profesor). No obs-

tante, nuevamente los subgrupos del Grupo 1 mostraron un ajuste de la evaluación más rápido que los alumnos de los subgrupos pertenecientes al Grupo 2. La Figura 2 compara por sesiones la GSA de 1 subgrupo para cada Grupo, esto hace que los datos solo puedan interpretarse de manera descriptiva y no puedan ser analizados por contraste de hipótesis.

Discusión y conclusiones

La presente investigación tenía la finalidad de analizar cómo la experiencia de evaluación y del feedback recibido sobre las evaluaciones podrían repercutir sobre la mejora de los alumnos a la hora de evaluar los logros de los demás y sus propios logros. La investigación desarrollada sigue las recomendaciones de estudios en el campo de la autoevaluación de proporcionar los criterios al alumnado y un instrumento para instaurar una autoevaluación más formativa (Abella García et al., 2020; Andrade, 2019; Siegesmund, 2017).

En base a los resultados obtenidos, se confirma la tendencia que tiene el alumnado a sobreestimar su capacidad, en comparación con la evaluación del profesor (Thawabie, 2017). Este hallazgo tiende a ser más común cuando el alumnado no es experto en la materia o no está familiarizado con el uso de estrategias de autoevaluación (Carroll, 2020). En cierto modo, esta tendencia se corrobora en nuestro estudio, ya que se observa cómo con el paso de las sesiones, el alumnado tiende a ser más preciso en comparación con las puntuaciones del profesorado (Yan et al., 2020). De esta manera, a medida que el alumnado va mejorando sus habilidades metacognitivas, centra su interés en reflexionar sobre sus producciones y es capaz de comparar su trabajo con los criterios propuestos, va siendo capaz de autorregularse, orientarse hacia unas metas definidas y aproximarse a una autoevaluación más precisa (Xiao & Yang, 2019). Estudios similares han mostrado que la instauración de evaluaciones formativas está estrechamente vinculada a la autorregulación del alumnado (Xiao & Yang, 2019).

Nuestro estudio también ha mostrado cómo la experiencia del alumnado para evaluar y autoevaluarse hace que mejoren estas habilidades asociadas a la evaluación, especialmente la autorregulación y la precisión. Al respecto, investigaciones como la realizada por Carroll (2020) con población australiana e internacional, encontró que la instauración de estrategias de autoevaluación y la retroalimentación supuso una mejora en el aprendizaje, especialmente para aquellos de bajo rendimiento.

Al respecto, una revisión sistemática reciente de Lerchenfeldt et al. (2019) señalaba que cuando los estudiantes conocían previamente los criterios de calificación de forma clara, mejoraban su capacidad para evaluarse a ellos mismos y a sus compañeros y favorecía su aprendizaje de la materia y trabajo en equipo. Este posicionamiento respalda la importancia de instaurar evaluaciones formativas en los procesos de enseñanza y aprendizaje del alumnado (Yan et al., 2021; Zainuddin et al., 2020), especialmente en educación superior.

De igual modo, se pudo determinar el importante papel ejercido por el feedback para mejorar la precisión de los estudiantes en la valoración de sus producciones. De hecho, diversos estudios sugieren la interacción entre retroalimentación y autorregulación (Clark, 2012). Este efecto en nuestro estudio se vio representado por la mejora significativa que mostró en la precisión de sus evaluaciones el grupo FB, el cual recibió retroalimentación basada en los criterios de evaluación, con respecto al grupo N-FB, que no recibió retroalimentación. Este hallazgo se sitúa en la línea de lo encontrado

por van der Kleij (2017), donde se aboga por comentarios estratégicos en las retroalimentaciones del profesor, que se ajusten a los criterios de evaluación ya conocidos por el alumnado. En contraposición, la investigación de Panadero et al. (2020) encontró que cuando el profesor ofrecía retroalimentación al alumnado, disminuían el uso de estrategias de autoevaluación realizadas por el alumnado, aunque eran capaces de autoevaluarse de forma más eficaz, gracias a los comentarios aportados.

A su vez, se sitúa en la línea de una revisión sistemática llevada a cabo por Yan et al. (2021), en la que analizan los factores que influyen en la implementación de evaluaciones formativas, desde la mirada del profesor. Este hallazgo debe ser considerado en el análisis de cualquier estudio sobre autoevaluación y mejora del aprendizaje del alumnado, pues según se planifiquen y desarrollen los procesos de enseñanza, así como según la visión que el profesor tenga en tal construcción, así serán las oportunidades que tendrán los alumnos para mejorar sus aprendizajes, autorregularse y orientarse a la consecución de los objetivos propuestos en el programa formativo (Xiao & Yang, 2019).

Otro aspecto importante en nuestro estudio es el uso de una rúbrica con unos criterios conocidos por el alumnado. En nuestro estudio, tanto alumnos como el propio profesor usaron la misma rúbrica como criterio de evaluación y eso repercutió en la mejora de la evaluación de los dos grupos. El uso de la rúbrica ha sido constatado en la literatura como uno de los factores que mejor contribuye, no solo para mejorar la precisión de la autoevaluación del alumnado, sino para mejorar su aprendizaje. Ejemplo de ello es el estudio de Su (2020) en el que se analizó el efecto de la rúbrica sobre el aprendizaje de habilidades de comprensión y ejecución relacionadas con el inglés en población china. Sus hallazgos determinaron como la rúbrica mejoró la comprensión y el aprendizaje de habilidades de interpretación y ejecución en inglés. Resultados similares muestra la investigación de Calle-Álvarez (2020) enfocada en la mejora de la escritura, la de Tur et al. (2019) orientada a mejorar el aprendizaje y la capacidad de reflexión de docentes en formación o el estudio de Zhang et al. (2019), en el que, dentro de una experiencia de aula invertida, la rúbrica supuso un mayor rendimiento de los estudiantes y de su conciencia metacognitiva. Además, otras investigaciones apuntan incluso hacia la instauración de co-rúbricas, para mejorar los protocolos de pensamiento en voz alta, dentro del aprendizaje autorregulado (Fraile et al., 2017).

Limitaciones y prospectiva

Una de las limitaciones del estudio es que está realizado con un grupo natural y eso hace que, aunque sea una muestra representativa de la realidad educativa universitaria, tenga un tamaño muestral reducido. Además, esto deviene en la siguiente limitación y es que, dado que los grupos de análisis son los subgrupos de clase, el análisis de GSA-Skill solo compara un subgrupo de FB con respecto a otro subgrupo de N-FB, permitiéndonos tan solo inferir el efecto de mejora de los datos descriptivos, pero no pudiendo realizar un contraste de hipótesis. Futuras investigaciones deberían realizarse ampliando la muestra a varios grupos naturales de clase.

Por otro lado, el presente estudio se centra en la repercusión que tiene la experiencia evaluando y la retroalimentación sobre la habilidad de evaluación, pero no se analiza cómo esta mejora en la habilidad de evaluación pudiera repercutir en una mejora en el propio rendimiento académico. Para poder hacer esta interpretación sería necesario un estudio en el que los alumnos tuvieran que hacer dos presentaciones, una previa a toda la experiencia de evaluación y feedback, y otra tras estas experiencias y com-

parar la mejora de la medida posterior respecto a la mediad previa (diseño pre-post con grupo control). Sería muy interesante que futuras investigaciones analizaran esta posible relación directa.

Referencias

- Abella García, V. A., Ausín Villaverde, V., Delgado Benito, V., & Casado Muñoz, R. (2020). Aprendizaje basado en proyectos y estrategias de evaluación formativas: Percepción de los estudiantes universitarios. *Revista Iberoamericana de Evaluación Educativa*, 13(1), 93-110. <https://dialnet.unirioja.es/descarga/articulo/7408493.pdf>
- Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, 27(4). <https://doi.org/10.3389/feduc.2019.00087>
- Brown, G. T., & Harris, L. R. (2014). The Future of Self-Assessment in Classroom Practice: Reframing Self-Assessment as a Core Competency. *Frontline Learning Research*, 2(1), 22-30. <http://dx.doi.org/10.14786/flr.v2i1.24>
- Calle-Álvarez, G. Y. (2020). La rúbrica de autoevaluación como estrategia didáctica de revisión de la escritura. *Revista de Investigación, Desarrollo e Innovación*, 10(2), 323-335. <https://doi.org/10.19053/20278306.v10.n2.2020.10628>
- Carroll, D. (2020). Observations of student accuracy in criteria-based self-assessment. *Assessment & Evaluation in Higher Education*, 45(8), 1088-1105. <https://doi.org/10.1080/02602938.2020.1727411>
- Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, 24(2), 205-249. <https://doi.org/10.1007/s10648-011-9191-6>
- Daura, F. T. (2017). Aprendizaje autorregulado e intervenciones docentes en la universidad. *Revista Educación*, 41(2), 56-74. <http://dx.doi.org/10.15517/revedu.v41i2.21396>
- Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, 53, 69-76. <https://doi.org/10.1016/j.stueduc.2017.03.003>
- Hawe, E., & Dixon, H. (2017). Assessment for learning: a catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, 42(8), 1181-1192.
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2020). Developing student competence through peer assessment: the role of feedback, self-regulation and evaluative judgement. *Higher Education*, 80(1), 137-156. <https://doi.org/10.1007/s10734-019-00469-2>
- Jónsson, Í. R., Smith, K., & Geirsdóttir, G. (2018). Shared language of feedback and assessment. Perception of teachers and students in three Icelandic secondary schools. *Studies In Educational Evaluation*, 56, 52-58. <http://dx.doi.org/10.1016/j.stueduc.2017.11.003>
- Kambourova, M. (2020). ¿Qué falta por comprender sobre el concepto autoevaluación (del aprendizaje) en educación superior? Una mirada diferente desde su historia. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 25(3), 640-658.
- Larruzea-Urkixo, N., & Ramírez, M. O. C. (2020). Diferencias individuales en aprendizaje autorregulado de estudiantes de los Grados de Educación: género, especialidad,

- notas y desempeño académico. *Revista de Investigación Educativa*, 38(2), 453-473. <https://doi.org/10.6018/rie.334301>
- Lerchenfeldt, S., Mi, M., & Eng, M. (2019). The utilization of peer feedback during collaborative learning in undergraduate medical education: a systematic review. *BMC medical education*, 19(1), 1-10. <https://doi.org/10.1186/s12909-019-1755-z>
- Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and practical connotations. When it happens, how is it acquired and what to do to develop it in our students. *Electronic Journal of Research in Educational Psychology*, 11(2), 551-576. <http://dx.doi.org/10.14204/ejrep.30.12200>
- Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: the effects of feedback, subject matter, year level, and gender. *Assessment in Education: Principles, Policy & Practice*, 27(6), 607-634. <https://doi.org/10.1080/0969594X.2020.1835823>
- Panadero, E., & Alqassab, M. (2019). An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading. *Assessment & Evaluation in Higher Education*, 44(8), 1253-1278. <https://doi.org/10.1080/02602938.2019.1600186>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98.
- Panadero, E., Brown, G. T., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28(4), 803-830. <https://doi.org/10.1007/s10648-015-9350-2>
- Panadero, E., Jonsson, A., & Strijbos, J. W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. En *Assessment for learning: Meeting the challenge of implementation* (pp. 311-326). Springer, Cham.
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89-101. http://dx.doi.org/10.1207/S15326985EP3602_4.
- Pui, P., Yuen, B., & Goh, H. (2020). Using a criterion-referenced rubric to enhance student learning: a case study in a critical thinking and writing module. *Higher Education Research & Development*, 1-14. <https://doi.org/10.1080/07294360.2020.1795811>
- Saiz, M. S. I., & Gómez, G. R. (2020). Aprendiendo a evaluar para aprender en la Educación Superior. *Revista Iberoamericana de Evaluación Educativa*, 13(1), 5-8. <https://dialnet.unirioja.es/descarga/articulo/7408491.pdf>
- Siegesmund, A. (2017). Using self-assessment to develop metacognition and self-regulated learners. *FEMS microbiology letters*, 364(11). <https://doi.org/10.1093/femsle/fnx096>
- Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer and collaborative assessment, *Assessment and Evaluation in Higher Education*, 18, 221-233. <https://doi.org/10.1080/0260293930180306>
- Su, W. (2020). Exploring how rubric training influences students' assessment and awareness of interpreting. *Language awareness*, 29(2), 178-196. <https://doi.org/10.1080/09658416.2020.1743713>

- Topping, K. J. (1998). Peer assessment students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.3102%2F00346543068003249>
- Thawabieh, A. M. (2017). A Comparison between Students' Self-Assessment and Teachers' Assessment. *Journal of curriculum and Teaching*, 6(1), 14-20. <http://dx.doi.org/10.5430/jct.v6n1p14>
- Tur, G., Urbina, S., & Forteza, D. (2019). Rubric-Based Formative Assessment in Process Eportfolio: Towards Self-Regulated Learning. *Digital Education Review*, 35, 18-35. <https://www.raco.cat/index.php/DER/article/view/360465>
- Van der Kleij, F. M., Adie, L. E., & Cumming, J. J. (2017). Using video technology to enable student voice in assessment feedback. *British Journal of Educational Technology*, 48, 1092-1105. <https://doi.org/10.1111/bjet.12536>
- Vicario-Molina, I., Martín-Pastor, E., Gómez-Gonçalves, A., & González-Rodero, L. (2020). Nuevos desafíos en la Educación Superior: análisis de resultados obtenidos y dificultades experimentadas en la realización del Trabajo Fin de Grado de estudiantes de los Grados de Maestro de la Universidad de Salamanca. *Revista Complutense de Educación*, 31(2), 185-194. <https://dx.doi.org/10.5209/rced.62003>
- Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34(2), 221-230. <https://doi.org/10.1080/02602930801955986>
- Wanner, T., & Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education*, 43(7), 1032-1047.
- World Medical Association. (2009). Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Journal of the indian medical association*, 107(6), 403-405.
- Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports students' self-regulation in English language learning. *System*, 81, 39-49. <https://doi.org/10.1016/j.system.2019.01.004>
- Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, 45(2), 224-238. <https://doi.org/10.1080/02602938.2019.1629390>
- Yan, Z., Brown, G. T., Lee, J. C. K., & Qiu, X. L. (2020). Student self-assessment: Why do they do it? *Educational Psychology*, 40(4), 509-532. <https://doi.org/10.1080/01443410.2019.1672038>
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*, 1-33. <https://doi.org/10.1080/0969594X.2021.1884042>
- Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin. Advance online publication*. <https://doi.org/10.1037/bul0000309>
- Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education*, 145, 103729. <https://doi.org/10.1016/j.compedu.2019.103729>

Zhang, Y., Chen, B. L., Ge, J., Hung, C. Y., & Mei, L. (2019). When is the best time to use rubrics in flipped learning? A study on students' learning achievement, meta-cognitive awareness, and cognitive load. *Interactive Learning Environments*, 27(8), 1207-1221. <https://doi.org/10.1080/10494820.2018.1553187>