

Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback

Ernesto Panadero^{1,2}, Daniel García-Pérez³, Javier Fernández Ruiz⁴, Juan Fraile⁵, Iván Sánchez-Iglesias⁶ & Gavin T. L. Brown⁷

¹ Facultad de Educación y Deportes, Universidad de Deusto, Bilbao, España.

² IKERBASQUE, Basque Foundation for Science, Bilbao, Spain.

³ Departamento de Investigación en Psicología y Educación, Facultad de Educación, Universidad Complutense de Madrid, Spain.

⁴ Departamento de Psicología, Sociología y Filosofía, Facultad de Educación, Universidad de León, Spain.

⁵ Universidad Francisco de Vitoria, Spain.

⁶ Departamento de Psicobiología y Metodología de las Ciencias del Comportamiento, Universidad Complutense de Madrid, Spain.

⁷ The University of Auckland, New Zealand.

Research funded by: Research funded by: Fundación BBVA call Investigadores y Creadores Culturales 2015 (project name Transición a la educación superior id. 122500); and by the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad) National I+D Call (Convocatoria Excelencia) project reference EDU2016-79714-P. First author funded by contribution of the Basque Government (Ref. IT1624-22) to the group Education Regulated Learning and Assessment.

Correspondence concerning this manuscript should be addressed to: Javier Fernández Ruiz. Despacho 223, Facultad de Educación. Universidad de León, León. 24007. Spain. E-mail: javier@fernandezruiz.com

Recommended citation: Panadero, E., García-Pérez, D., Fernández Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. L. (2023). Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: positive impact of rubrics vs. instructor feedback. *Educational Psychology*, 43(7), 756-779. <https://doi.org/10.1080/01443410.2023.2254015>

This is a post-peer-review, pre-copyedit version of an article published in Educational Psychology. The final authenticated version is available online at: <https://doi.org/10.1080/01443410.2023.2254015>. This manuscript may not exactly replicate the published version due to editorial changes and/or formatting and corrections during the final stage of publication. Interested readers are advised to consult the published version.

Abstract

This study explored the effects of three factors (feedback occasion, type of feedback, and year level) on self-efficacy and emotions when university students self-assessed. 126 higher education students from three different year levels self-assessed their performance on a writing task two times (before and after receiving feedback). Self-efficacy and emotions were measured via self-report at three times (a baseline before self-assessment, after self-assessment without feedback, and after self-assessment with feedback). Feedback was provided in one of three conditions (rubric vs. instructor's feedback vs. combined) to which the participants were randomly assigned. In general, rubric feedback showed the strongest effects on self-efficacy and emotions (i.e., increased positive and decreased negative emotions). Year level was significant for self-efficacy, showing that the effect of university schooling experience enhances this variable. This study shows the impact that feedback has on self-efficacy and emotions while self-assessing, with an especially positive effect for rubric feedback.

Keywords: self-assessment; self-efficacy; emotions; feedback; rubric; higher education.

Feedback and year level effects on university students' self-efficacy and emotions during self-assessment: Positive impact of rubrics vs. instructor feedback

Self-assessment “involves a wide variety of mechanisms and techniques through which students describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products” (Panadero et al., 2016 p. 804). Self-assessment is central for autonomy and self-regulation (e.g., Andrade, 2018) and has been shown to be a moderately strong predictor of students' academic achievement (Brown & Harris, 2013; Yan et al., 2022), as well as self-regulated learning and self-efficacy (Panadero et al., 2017; Yan, 2019). Feedback about academic performance, a strong predictor of achievement (Hattie & Timperley, 2007; Wisniewski et al., 2020), may have benefits for the validity and reliability of self-assessment. However, how feedback about performance contributes to achievement emotions and self-efficacy within self-assessment processes is little understood. Additionally, self-assessment can be influenced by educational experience and expertise in the task at hand (Brown & Harris, 2013; Falchikov & Boud, 1989), aspects that still need to be more systematically investigated. Therefore, this study aims to systematically extend our understanding of the effects of feedback on self-efficacy and achievement emotions in the context of self-assessment by tertiary students of three different year levels.

Self-assessment and its educational importance

Self-assessment is a vital academic skill students need to develop as it has a direct influence on students' academic performance and it is a central element on students' self-regulation (Boud, 1995; Kostons et al., 2012; Yan et al., 2022). Importantly, there is a distinction between formal and informal self-assessment, with the former being instructionally led by the teacher on the students, and the latter happening

as a natural process of the self-regulatory loop. Ideally, the first one will directly influence the second one, probably the rationale under which teachers ask their students to self-assess, but it is not always the case (for a detailed discussion on this Panadero & Alonso-Tapia, 2013). For that reason, the more researchers understand the effects of formal self-assessment, the better we will be able to integrate it in our courses.

Nevertheless, there are many aspects around self-assessment that remain unexplored -“the black box of self-assessment”- and need to be investigated (Yan & Brown, 2017). Importantly, we need to understand how students enact self-assessment, meaning identifying the specific strategies and criteria they use when self-assessing, and what are their reactions (Panadero et al., 2020, 2022). Also, there is limited research on how feedback influences self-assessment; in other words, what students do when they receive external feedback, how they include such information in their self-reflection, and how their subsequent actions are impacted. Only with such knowledge we will be able to tailor our self-assessment interventions, especially to help the students that struggle the most developing the skill.

Crucially, self-assessment does not happen in a vacuum as it is influenced by the instructional context (Boud, 1995; Harris & Brown, 2018; Tan, 2012), but this is also an aspect we need to better investigate. Panadero and colleagues (2016) review identified several key variables for which the effects of self-assessment have not been systematically explored. In the next section we elaborate in the ones that will be explored here which we believe to be of utter importance.

Potential educational influences on self-assessment: type of feedback, feedback occasion and year level (educational experience)

The impact of feedback, an integral part of assessment, has not been sufficiently studied. We adhere to the following definition:

Feedback is information that includes all or several components: students' current state, information about where they are, where they are headed and how to get there, and can be presented by different agents (i.e., peer, teacher, self, task itself, computer). This information is expected to have a stronger effect on performance and learning if it encourages students to engage in active processing (Lipnevich & Panadero, 2021 p. 25).

One of the key aspects of the definition is that feedback comes from different agents that can be internal or external to the learner. Butler and Winne's (1995) review pointed out that self-regulated learners go through loops of internal feedback that get adjusted through external feedback. Therefore, there is a crucial relationship between students' self-assessment and the external feedback they receive from other sources (e.g., teachers).

Unfortunately, as mentioned earlier, there is little knowledge about the influence of external feedback on students' self-assessment. Actually, we only know of four studies exploring the effects of external feedback on self-assessment actions (Panadero et al., 2012, 2020, 2022; Raaijmakers et al., 2019). Three of the studies did not show much promise for the effects of feedback on self-assessment as either the results were mostly non-significant (Panadero et al., 2012) or even negative (Panadero et al., 2020; Raaijmakers et al., 2019). However, the picture might be more complex as Panadero and colleagues (2020, 2022) found different results using a similar aim in secondary education and higher education students. In 2020 feedback reduced the use of both self-assessment strategies and criteria in secondary education students, while in 2022 feedback enhanced university students' self-assessment criteria and decreased self-assessment strategy use. Some of the main differences between both studies were in 2020 the design was descriptive with no manipulation of the type of feedback while in

2022 the design was a randomized controlled trial with three types of feedback (rubric vs. instructor vs. combined). The present study is a continuation of the 2022 study as we are analyzing new data from the same sample. Here we will explore the impact of feedback occasion -i.e., performing self-assessment before receiving feedback and after receiving feedback- and feedback type -i.e., rubric vs instructor vs combined- while performing self-assessment in two new variables: self-efficacy and emotional reactions. Next, we elaborate more on the reasons to study feedback in regard to self-assessment.

First, the impact of feedback can be evaluated by comparing how students enact self-assessment before and after receiving feedback. To our knowledge, there are only two previous studies investigating this matter which we already compared in the previous paragraph (Panadero et al., 2020, 2022). The delivery of external feedback seems then to have important effects on self-assessment that ask for further research to help instructors on when they should deliver feedback when requiring or expecting students to self-assess. We will have three measurement occasions of our dependent variables -i.e., self-efficacy and emotions- to explore the feedback occasion effect. The first measure will take place right after giving the instructions to the participants as to have a baseline for their self-efficacy and emotions; followed by a second measure right after the first self-assessment without feedback; and finished by a final third measure right after the second self-assessment with feedback.

Secondly, feedback can be of very different types, so it is important to understand if they have distinct effects on self-assessment. Historically, feedback researchers have a special interest in “organizing, categorizing and describing different aspects of feedback” (Panadero & Lipnevich, 2022 p. 3), in what has been called taxonomies or typologies of feedback. The rationale behind this is understanding if different types (e.g., oral vs. written) produce better academic results. Here we will

compare rubric vs. instructor vs. combined. Rubrics have been shown to have positive effects on self-assessment interventions and academic performance (Brookhart & Chen, 2015), calibration (Hawthorne et al., 2017), and even stronger effects on performance than other assessment tools such as exemplars (Lipnevich et al., 2014, 2022) or scripts (Panadero et al., 2012). Instructor's feedback is one of the most important influences on academic achievement (Hattie & Timperley, 2007; Wisniewski et al., 2020) especially if delivered in the form of comments (Koenka et al., 2019). Scholars have claimed that instructor's feedback is a key for self-assessment because students need to receive feedback on their self-assessment practices, as well as their learning outcomes (Andrade, 2018; Boud, 1995; Brown & Harris, 2014; Panadero et al., 2016). However, the empirical evidence on this effect is limited and, to our knowledge, instructor's content feedback has not been compared to other types of feedback on their effects on self-assessment. Understandably, it may be that the combination of a rubric and instructor comments will be of greater value than either type of feedback alone, thus the reason why our design also contains a combined condition. However, previous research has demonstrated that a rubric alone had better results than combining a rubric with exemplars (Lipnevich et al., 2014, 2022) and, therefore, the need to investigate the effects of a combined condition. In sum, we will explore the effect of three types of feedback (rubric vs. instructor vs. combined) in our participants' self-efficacy and emotions when they self-assess.

A final factor that could influence self-assessment and needs to be explored is educational experience operationalized here as year level. Research shows that with greater academic ability or performance, self-assessments are more likely to align with the judgments of teachers or test scores (Barnett & Hixon, 1997; Brown & Harris, 2013). Part of the reason for this is that self-assessment is more complex for novices

(Falchikov & Boud, 1989; Kostons et al., 2009, 2010). For example, Jax and colleagues (2019) found that with higher content knowledge the accuracy of self-assessment increased. Nonetheless, there is still a need to study in more detail the impact of educational level and domain skill especially as there are only a few studies directly exploring such influence. To our knowledge, only two studies have explored year level differences on self-assessment (Panadero et al., 2020, 2022). The results of these two studies could be summarized in that there are no effects of year level except for self-assessment strategies in secondary students (Panadero et al., 2020). However, we will continue exploring the effects of year level with the same data collection from Panadero and colleagues (2022) but here focusing on the effects on self-efficacy and emotional reactions.

Psychological covariates of Self-assessment: Self-efficacy and emotional reactions

Because self-assessment takes place within the mind and experience of an individual, it is likely that such processes are limited by or assisted by other psychological processes. Self-assessment coordinates with other self-related beliefs, such as self-efficacy –i.e. perception of one’s own ability to perform in a particular task– and self-regulation –i.e. cognitive, emotional, motivational and behavioral strategies deployed to achieve the students’ learning goals (e.g., Sitzmann et al., 2012). Here we will analyze the effects on self-efficacy and emotions.

The relationship of self-assessment and self-efficacy is based on the premise that it is probably bidirectional and complex, influenced by a variety of factors such as the learning context, student motivation, and individual differences among students. In terms of the direction of the effects, it is supposed that effective self-assessment can enhance self-efficacy, but at the same time self-efficacy might play a crucial role to successfully implement self-assessment. When students accurately assess their own

learning, they become more aware of their abilities and their progress. This can improve their confidence in their own skills, thereby increasing their self-efficacy. Likewise, students with higher self-efficacy may be more likely to engage in self-assessment, as they believe in their ability to successfully monitor and evaluate their own learning. This relationship has been attested in several studies. Yan and colleagues (2020) have shown self-efficacy is a strong predictor of self-reported use of self-assessment. A meta-analysis by Panadero et al. (2017) found a .73 effect size of self-assessment interventions and self-efficacy. Winstone and colleagues (2021) found that self-efficacy mediated the relationship between self-reported feedback use, which was greater among students who scored high in mastery approach goals, performance approach goals, and conscientiousness. Then, we need to start investigating the role of external feedback in the relationship between self-assessment and self-efficacy as external feedback holds a vital influence in self-assessment (Butler & Winne, 2019; Nicol, 2020). That was the aim of Panadero and colleagues (2020) who found no significant differences in their four factors' main effects (year level x feedback occasion x subject x gender), and only the interaction subject x feedback occasion resulted in a significant effect. Here we will investigate the results on self-efficacy from the study performed by Panadero and colleagues (2022) manipulating different feedback type, occasion, and year level.

The second psychological covariate of self-assessment that we will explore is emotions, as there seems to exist a relationship. One of the conclusions from Sitzmann et al.'s (2010) meta-analysis was that self-assessment had strong correlations with motivation and satisfaction which are affective outcomes, in other words, related to emotional reactions. In a similar vein, Panadero and colleagues (2017) found in their meta-analysis that self-assessment interventions had a -.65 effect size on negative maladaptive self-regulation, which is the one lead by negative emotions and avoidance

goals. Based on these and the extensive work on achievement emotions (Pekrun et al., 2002, 2018), it is highly likely that self-assessment depends on or contributes to emotional state. Indeed, in a study of university students' achievement emotions around a mid-term test, it was only when students got their scores that there was a systematic relationship to achievement emotions (Peterson et al., 2015). Therefore, we will also explore the effects on emotional reactions of feedback occasion, feedback type, and year level while performing self-assessment.

Research Aim and Questions

Our aim was to investigate the effects of feedback (occasion and type) on self-efficacy and emotions in the context of self-assessment by university students from different year levels. In order to accomplish this, we implemented a randomized experimental design, examining two feedback occasions -i.e., before feedback and after feedback-, three feedback conditions (rubric vs. instructor vs. combined) and three year levels (1st, 2nd and 3rd university undergraduates). This arrangement served to identify potential influences on the students' self-efficacy and emotional state during the self-assessment process. Significantly, we measured our dependent variables three times to establish a comprehensive baseline: prior to self-assessment, after self-assessment before feedback, and following self-assessment after feedback had been provided.

The study was organized around two research questions (RQ) and the corresponding hypotheses:

RQ1. What are the effects of feedback occasion, type of feedback, and year level on self-efficacy when students self-assess?

H1a. Self-efficacy will increase along the feedback occasions as students will increase their expertise in the task.

H1b. When participants receive feedback, those who receive the rubric (rubric condition) will show higher self-efficacy, as students will generate their own evaluation.

H1c. Students in higher year levels will report higher self-efficacy levels due to higher experience with the task.

RQ2. What are the effects of feedback occasion, type of feedback, and year level on emotions when students self-assess?

H2a. Positive emotions will increase, and negative emotions will decrease along the feedback occasions as students will increase their expertise in the task.

H2b. When participants receive feedback, those who receive the rubric (rubric condition) will show more positive emotions and less negative emotions as students will generate their own evaluation.

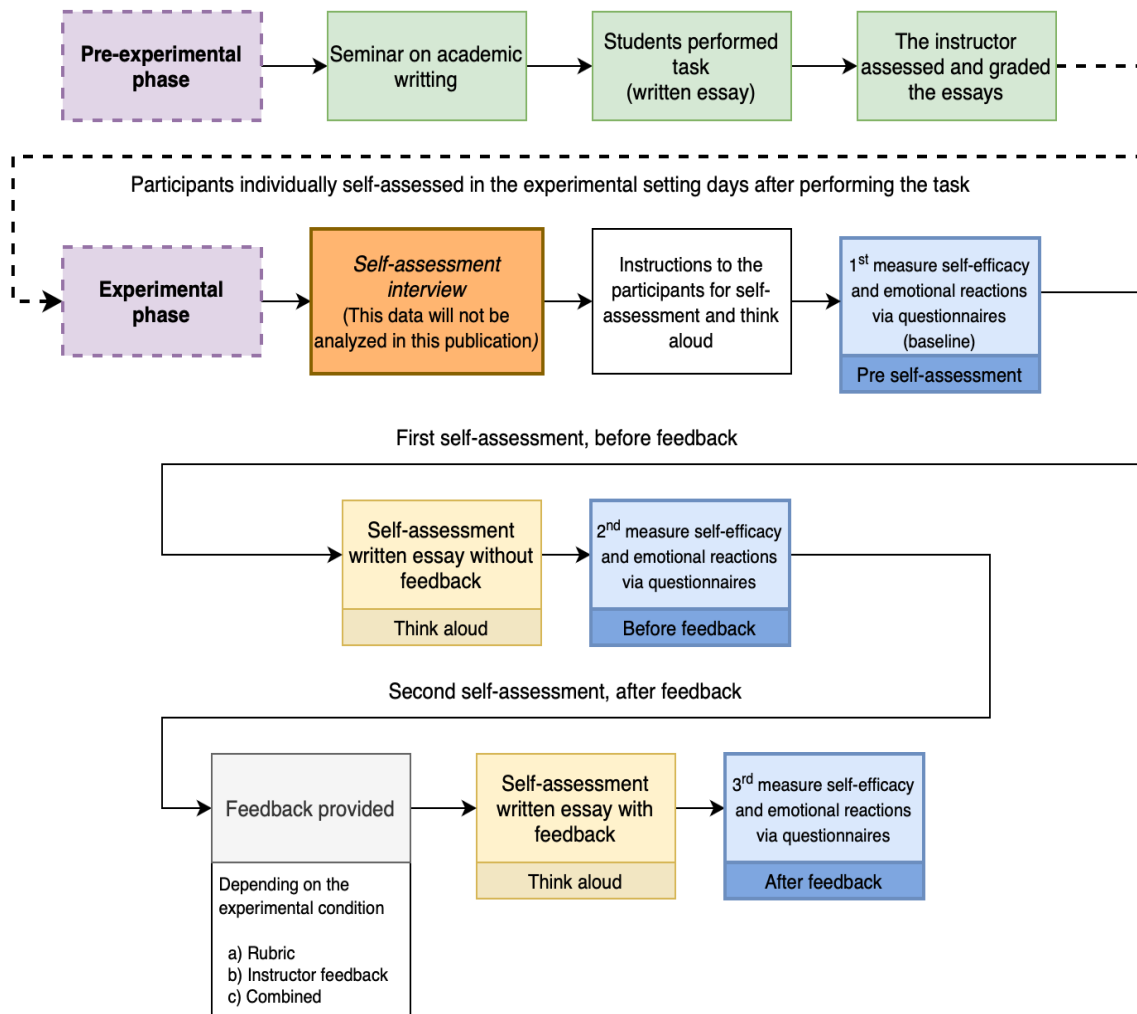
H2c. Students in higher year levels will report higher positive emotions and lower negative emotions due to higher experience with the task.

Method

This study employed a randomized control trial with a between-subjects design. As the university Ethics Committee did not grant permission for a control group, there was 'no treatment' control group, and the participants were randomly assigned to one of three alternative treatments. The procedure can be found in Figure 1.

Figure 1

Experimental procedure



Sample

The participants were 126 undergraduate psychology students (88.1% females) across first, second and third year of study (34.9%, 31.7%, and 33.3%, respectively) randomly assigned to one of three feedback conditions: rubric only ($n = 43$), instructor feedback ($n = 43$), and rubric + instructor feedback combined ($n = 40$). This is a convenience sample of volunteers from the university where the first author worked. Participants received credit in accordance with the faculty volunteering programme. We carried out a sensitivity analysis (G*Power 3.1.9.2; Faul et al., 2007) for a 3 x 3 ANOVA with a within-between interaction. Given a risk level of $\alpha = .005$, and a

statistical power of $1 - \beta = .800$, the current sample size would allow to detect an effect size $f = 0.127$. For a 3×3 ANOVA between factors, this effect size would be $f = 0.314$.

Data Collection and Instruments

Students' self-efficacy questionnaire (Appendix A). This instrument was created ad-hoc to measure students' perceptions of efficacy when writing an academic text. It consisted of eight items, using a seven-point rating scale ranging from 0 "absolutely disagree" to 6 "absolutely agree". The internal consistency of this instrument was computed for all three measurement moments: pre self-assessment, $\alpha = .78$; before feedback, $\alpha = .80$; and after feedback, $\alpha = .87$. To confirm the subjacent unidimensional structure of the questionnaire, we carried out a confirmatory factor analysis (CFA). As the assumption of multivariate normality was not met (Mardia Skewness = 181.21, $p < .001$; Mardia Kurtosis = 1.89, $p = .059$), a diagonally weighted least squares (DWLS) estimation method was used. Of the several fit indices employed, all but SRMR exceeded the recommended values (Chau, 1997; Schreiber et al., 2006): $\chi^2/df = 1.600$; RMSEA = .075, 90% CI [.013, .121]; CFI = .981; TLI = .974; SRMR = .097. Furthermore, the model showed positive item factor loadings ($ps < .001$), ranging from .375 to .807.

Emotions questionnaire (Appendix B). This instrument was created for this study, adopting 10 basic emotions from Pekrun et al. (2002); that is, four positive emotions (i.e., joy, hope, pride, relief) and six negative emotions (i.e., anxiety, nervousness, embarrassment, sadness, boredom, and anger). Participants answered in a scale from very low intensity (1) to maximum intensity (10). We used CFA to confirm that the underlying structure of the questionnaire consisted of two correlated factors (positive emotions and negative emotions). Once again, we utilized the DWLS estimation method due to the non-normal multivariate distribution (Mardia Skewness = 647.16, $p < .001$; Mardia Kurtosis = 6.99, $p < .001$).

The fit indices were $\chi^2/df = 1.300$; RMSEA = .051, 90% CI [.000, .090]; CFI = .975; TLI = .967; SRMR = .093. The factor analysis results revealed that the positive emotions factor exhibited significant and positive item factor loadings ($ps < .001$), ranging from .668 to .811. Similarly, the negative emotions factor displayed positive item factor loadings ($ps < .050$), ranging from .250 to .828. However, the item corresponding to "anger" had a non-significant factor loading ($\beta = .184, p = .040$). The correlation between the two factors was also non-significant, $r = -.214, p = .065$. Though the items were described individually, we calculated reliability indices for positive emotions ($\alpha = .93$) and negative emotions ($\alpha = .90$).

Rubric (Appendix C). The rubric was built from experts' writing models. It comprises the following assessment: (1) writing process, (2) structure and coherence, and (3) sentences, vocabulary and punctuation. The rubric presents three performance levels: low, average, and high. As these criteria was scored independently the rubric is analytic.

Instructor's feedback (Appendix D). The teacher offered written feedback on every essay, utilizing identical categories as those in the rubric. Regarding the "writing process" criterion, which wasn't entirely evident to the instructor, feedback was given by indicating whether certain strategies, like planning, were evident in the work. Alongside this, a grade between 0 to 10 points was assigned. The assessment of all essays was conducted by the second author, while the first author assessed a third of the essays, resulting in complete alignment with the rubric categories.

Procedure

This randomized experiment is part of a larger study; this report focuses on the specific effects on self-efficacy and emotional reactions as measured via self-report (See Figure 1). First, the participants attended a three hours group seminar on academic writing. At the end of the seminar, the students had to write a short essay answering the

question: “Why is the psychologist profession necessary?”. This topic was directly connected to the participants’ psychology programme. The participants wrote their essays on the computers and sent it to the research team. This essay did not have implications outside of the research experiment, but we emphasized its utility for the students’ academic perspective of the programme.

Days later, during the experimental phase, participants visited the laboratory setting individually, where they engaged face-to-face with one of the authors (second, third or fourth author) who used the same guide notes to ensure consistency and comparability among participants' experiences. First, participants were interviewed about their self-assessment practices by replying to questions exploring three main themes: (1) self-assessment training, (2) self-assessment processes, and (3) use of feedback for self-assessment. These data are not presented in this paper. Secondly, participants were provided with verbal instructions on how to self-assess their previous essay. These instructions were delivered by one of the authors who read the following text:

“You are going to perform two types of self-assessment. In the first one, I am going to give you back the essay you wrote in the writing workshop. You have to self-assess it, that is, evaluate your essay. Do everything you consider necessary during that self-assessment: strategies, techniques, etc. Remember that, according to the workshop content, you have to assess the quality of your text and explain why you believe it is of that quality”.

The researcher also wrote instructions to instruct the participants to think aloud, verbally reporting their thoughts, emotions, and motivational reactions as asked in the oral instructions. And, third, the participants filled the self-efficacy and emotions questionnaires.

Then students self-assessed their essays in the original form, in other words, without any feedback. To ensure that they remembered the instructions and in case they wanted to check what they were supposed to do we gave them a sheet containing the following information:

“The objective of this task is for you to self-assess your essay based on the recommendations given during the seminar. To do this, take into account the different aspects that the professor has indicated as relevant for academic writing.”

Purposefully, we did not set any limits, characteristics or give any indications to the students in terms of what their self-assessment should entail. The rationale was to observe the “natural occurrence” of self-assessment. Immediately after finishing the self-assessment, participants filled out the self-efficacy and achievement emotions scales a second time.

Then, participants were given feedback according to the condition they had been assigned to. In the rubric condition, participants received the rubric described in the previous section (Appendix C). In the instructor’s feedback condition, participants received written feedback from the seminar instructor, also described in the previous section (Appendix D). In the combined condition, participants received both the rubric and instructor’s feedback.

After reading the feedback, the participants were asked to self-assess again and reminded they had the instructions in the sheet given before their first self-assessment. Then they repeated the self-assessment while thinking aloud. Once they finished, they completed the self-efficacy and emotion questionnaires for the third time.

Data Analysis

We first computed the mean scores of the self-efficacy questionnaire, as well as the positive and negative emotions from the emotion questionnaire for each subject. Next, we conducted descriptive statistical analyses on these outcome variables. No missing values were found in our dataset. To study the effect of the factors (feedback occasion, condition, and year level) on self-efficacy and emotions three-way ANOVAs (mixed, as feedback occasion was a within-subjects factor) were calculated. The assumptions for the ANOVAs were previously checked, using Shapiro-Wilk tests and normal Q-Q plots for normality, and Levene tests for homoscedasticity. In the Feedback occasion x Condition x Year level on self-efficacy ANOVA, 3 out of 27 Shapiro-Wilk tests were significant. However, even in those cases, the normal Q-Q plots exhibited a reasonable fit to the reference diagonal. When the dependent variable was the mean of positive emotions, 5 out of 27 Shapiro-Wilk tests were significant. Nonetheless, the corresponding Q-Q plots demonstrated an adequate fit to a normal distribution. For the dependent variable of mean negative emotions, 2 out of 27 Shapiro-Wilk tests were significant. Once again, the Q-Q plots showed a good fit to a normal distribution. Concerning homoscedasticity, the Levene's tests were not significant ($ps > .05$) when conducting the ANOVAs for positive emotions and negative emotions. However, in the third measurement of self-efficacy (after feedback), the Levene's test was significant, $F = 2.12, p = .039$. This indicates a minor violation of the assumption of homoscedasticity, but ANOVA is fairly robust when sample sizes are equal. Nonetheless, we must interpret the findings with this consideration in mind.

For significant main effects, we used *post hoc* testing with Bonferroni correction for multiple comparisons. For significant interaction effects, simple effects were computed. For the confirmatory factor analysis (CFA) of the instruments shown in Data Collection and Instruments, we utilized R and the packages lavaan (version 0.6-3,

Rossee, 2012), and MVN (Korkmaz et al., 2014). Additionally, we employed SPSS 25 for all other analyses. The significance level for all inferential tests was set at $\alpha = .050$.

Results

RQ1. What are the effects of feedback occasion, type of feedback, and year level on self-efficacy when students self-assess?

A three-way mixed ANOVA was carried out to assess the effects of feedback occasion, condition, year level, and the interactions between them, on the self-efficacy scores (see Tables 1 and 2 for descriptive statistics).

Table 1

Descriptive Statistics for Self-Efficacy Scores, by Condition and Feedback Occasion

Condition	n	Pre self-assessment		Before feedback		After feedback	
		M	SD	M	SD	M	SD
Rubric	43	3.76	0.68	3.86	0.86	4.16	0.84
Instructor’s feedback	43	3.98	0.74	3.80	0.81	3.74	1.01
Combined	40	3.88	0.75	3.92	0.72	4.00	1.01
Total	126	3.87	0.72	3.86	0.80	3.96	0.97

Table 2

Descriptive Statistics for Self-Efficacy Scores, by Year Level and Feedback Occasion

Year level	n	Pre self-assessment		Before feedback		After feedback	
		M	SD	M	SD	M	SD
First year	44	3.65	0.69	3.70	0.73	3.72	1.04
Second year	40	3.82	0.73	3.75	0.87	3.94	0.89
Third year	42	4.16	0.66	4.13	0.74	4.24	0.89
Total	126	3.87	0.72	3.86	0.80	3.96	0.97

Table 3 shows the ANOVA tests summary for main effects and interactions.

Table 3

ANOVA Summary Table for Self-Efficacy Scores.

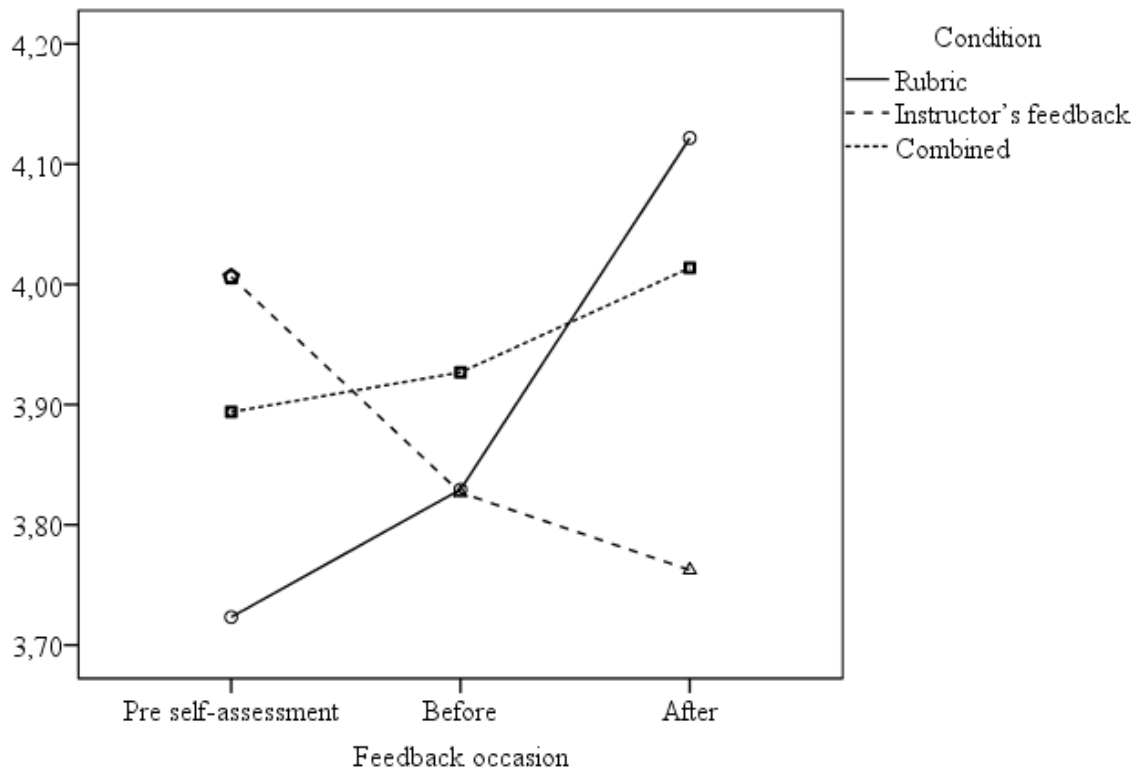
Effect	F	df	p	η^2
Condition	0.151	2, 115	.860	.003

Year level	3.586	2, 115	.031	.059
Feedback occasion	0.062	2, 114	.940	.001
Condition x Year level	0.374	4, 115	.827	.013
Feedback occasion x Condition	3.548	4, 230	.008	.058
Feedback occasion x Year level	0.706	4, 230	.589	.012
Feedback occasion x Condition x Year level	0.858	8, 230	.553	.029

No significant effect was found for Feedback occasion. A significant Feedback occasion x Condition interaction was observed. This interaction was interpreted computing the simple effects: Within each feedback occasion, there were no significant differences between conditions. However, in the rubric condition, the mean of self-efficacy scores is higher after feedback than before feedback ($p = .01$) and pre self-assessment ($p < .01$) occasions. In the other two conditions, no differences were found between feedback occasions (Figure 2).

Figure 2

Self-Efficacy scores, by Condition and Feedback occasion.



We found a significant effect of year level (a 5.9% of the variance of the scores can be accounted for the year level). The mean self-efficacy scores for the three year levels (Table 2) showed a significant difference in self-efficacy scores between the first- and third-year level ($p = .03$ in the *post hoc* tests). No significant effect was found for the interactions where year level was involved: Condition x Year level; Feedback occasion x Year level; and Feedback occasion x Year level x Condition.

Hypothesis 1a and 1b were partially supported. As expected, students using a rubric showed higher levels of self-efficacy when they received feedback than before. However, this increase was not observed in the other conditions. Hypothesis 1c was supported by the data. The more experienced students showed higher self-efficacy scores than the less experienced ones.

RQ2. What are the effects of feedback occasion, type of feedback, and year level on emotions when students self-assess?

Descriptive statistics of positive and negative emotions are reported for feedback condition (Table 4) and year level (Table 5) on each occasion of self-report.

Importantly, we report the results for the individual emotions (e.g., joy, anxiety) in Appendix E and F, in case the reader wants to explore such data.

Table 4

Descriptive Statistics for Scores of Positive and Negative Emotions Assessed, by Condition and Feedback Occasion

Emotions	Condition	Pre-assessment		Before feedback		After feedback	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Positive	Rubric	4.72	1.63	5.16	1.64	6.51	1.83
	Expert’s feedback	4.85	1.83	4.85	1.94	5.13	2.11
	Combined	5.39	1.50	5.67	1.54	6.01	1.72
	Total	4.98	1.67	5.22	1.74	5.88	1.97
Negative	Rubric	3.14	1.08	2.82	1.12	2.15	0.95
	Expert’s feedback	3.04	1.24	2.95	1.16	2.82	1.17
	Combined	2.84	1.28	2.54	1.24	2.24	1.12
	Total	3.01	1.20	2.78	1.17	2.41	1.12

N = 126; *n*_{Rubric} = 43; *n*_{Expert’s feedback} = 43; *n*_{Combined} = 40. Positive emotions: Joy, Hope, Pride, Relief. Negative emotions: Anxiety, Nervousness, Embarrassment, Sadness, Boredom, Anger.

Table 5

Descriptive Statistics for Scores of Positive and Negative Emotions Assessed, by Year Level and Feedback Occasion

Emotions	Year level	Pre-assessment		Before feedback		After feedback	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Positive	First year	4.84	1.61	5.27	1.76	5.85	2.25
	Second year	5.06	1.73	5.22	1.73	5.99	1.88
	Third year	5.05	1.72	5.16	1.77	5.81	1.78
	Total	4.98	1.67	5.22	1.74	5.88	1.97

Negative	First year	3.02	1.29	2.69	1.28	2.36	1.28
	Second year	2.89	1.09	2.74	1.10	2.36	0.99
	Third year	3.11	1.21	2.90	1.15	2.50	1.07
	Total	3.01	1.20	2.78	1.17	2.41	1.12

$N = 126$; $n_{1stYear} = 44$; $n_{2ndYear} = 40$; $n_{3rdYear} = 42$. Positive emotions: Joy, Hope, Pride, Relief. Negative emotions: Anxiety, Nervousness, Embarrassment, Sadness, Boredom, Anger.

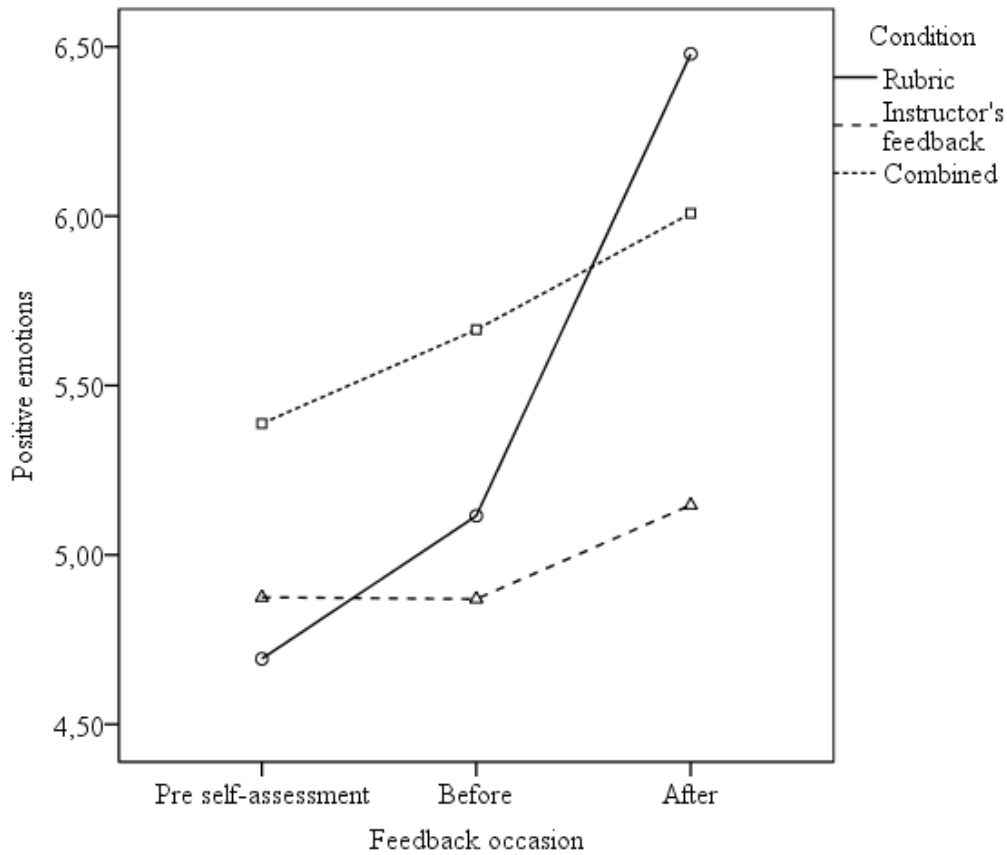
Positive emotions exhibited statistically significant effect based on Condition x Feedback occasion interaction (Table 6 and Figure 3). We computed the simple effects and found some significant differences to interpret this interaction. Within the third measurement occasion (after feedback), the rubric group scored higher than the instructor's feedback group ($p = .005$). In addition, within the Rubric group, we observed significant differences in positive emotions over time (pre self-assessment - before feedback, $p = .003$; pre self-assessment - after feedback, $p < .001$; before feedback - after feedback, $p < .001$).

Table 6*ANOVA Summary Table for Scores of Positive and Negative Emotions.*

	Effect	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Positive emotions	Condition	2.233	2, 117	.112	.037
	Year level	0.047	2, 117	.954	.001
	Feedback occasion	18.367	2, 116	< .001	.241
	Condition x Year level	1.265	4, 117	.288	.041
	Feedback occasion x Condition	4.968	4, 234	.001	.078
	Feedback occasion x Year level	1.054	4, 234	.380	.018
	Feedback occasion x Condition x Year level	0.546	8, 234	.821	.018
	Effect	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Negative emotions	Condition	1.363	2, 117	.260	.023
	Year level	0.246	2, 117	.782	.004
	Feedback occasion	24.626	2, 116	< .001	.298
	Condition x Year level	0.477	4, 117	.753	.016
	Feedback occasion x Condition	3.637	4, 234	.007	.059
	Feedback occasion x Year level	0.649	4, 234	.628	.011
	Feedback occasion x Condition x Year level	0.146	8, 234	.997	.005

Figure 3

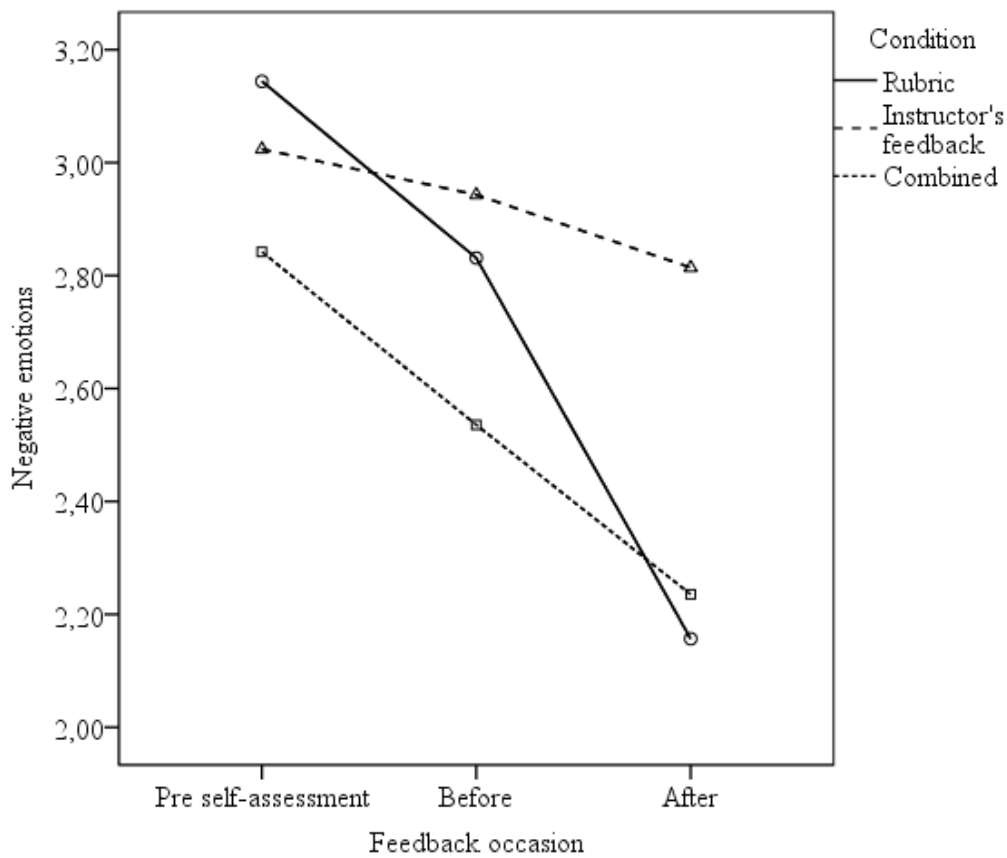
Mean Scores of Positive Emotions, by Condition and Feedback Occasion



We found similar results on negative emotions: significant Condition x Feedback occasion interaction (Table 6 and Figure 4). We found the following simple effects: Within the third measurement occasion (after feedback), the rubric group scored lower than the instructor's feedback group ($p = .021$). Within the Rubric group, we observed significant differences in negative emotions over time (pre self-assessment - before feedback, $p = .002$; pre self-assessment - after feedback, $p < .001$; before feedback - after feedback, $p < .001$). Finally, in the combined group, we observed significant differences between pre self-assessment - before feedback, $p = .003$; and between pre self-assessment - before feedback, $p < .001$) (but not before feedback - after feedback, $p = .070$).

Figure 4

Mean Scores of Negative Emotions, by Condition and Feedback Occasion



Hypothesis 2a was supported for positive emotions (students' scores increased with each feedback occasion) and for negative emotions (students' scores decreased). Hypothesis 2b was supported. As expected, when they had received feedback, students using a rubric showed higher positive emotions and lower negative emotions than the students receiving the instructor's feedback. The data did not support hypothesis 2c. Year level had no significant effect on either positive or negative emotions.

Discussion

This study explored the effects of three factors (i.e., feedback occasions, type of feedback, and year level -educational experience-) on self-efficacy and emotions when students self-assessed. The majority of our hypotheses were supported. To summarize

the findings, feedback conditions had similar scores for self-efficacy and emotions especially at the two first occasions of measurement. However, after feedback, self-efficacy and achievement emotions seemed to be more positive primarily in the rubric condition. Additionally, the small effect of year level on self-efficacy was independent of feedback condition, suggesting that greater experience and academic success are likely causes of this effect. Next, we will independently discuss the theory and previous research evidence regarding the two research questions.

Self-efficacy is highly connected with self-assessment as (a) self-efficacy is a self-assessment of one's own ability and (b) to evaluate and judge their efficacy students need some type of assessment of their own performance. One of the main claims has been that self-assessment might increase self-efficacy as students feel more competent being able to evaluate their own progress, an area of research largely explored in combination with the use of rubrics (e.g., Andrade et al., 2009). Such is the importance of the empirical evidence of this relationship that even a previous meta-analysis explored the effects of self-assessment interventions in the enhancement of self-efficacy, finding a $d = .73$ as a general effect size and also that self-efficacy increases more among girls (Panadero et al., 2017). This effect is logical because practicing self-assessment seems to be a contributor in the enhancement of self-efficacy. Now, what is the comparison of our results with similar previous research?

First, the interaction of feedback occasion and feedback type was significant which makes sense as only in the last measurement occasion (i.e., after feedback) the participants received different interventions. While no significant differences were observed in the pre-assessment and pre-feedback stages, indicating equal conditions prior to the intervention, the situation changed post-intervention. After the participants received various types of feedback, the rubric-based feedback demonstrated a positive

impact. Regarding the rubric positive effect on self-efficacy, this effect has been found in previous studies (e.g., Panadero et al., 2012), however other studies have not found significant effects (e.g., Andrade et al., 2009). When it comes to the effects of feedback occasions, this has been barely studied previously, and we only know of one study in which the feedback occasion was not significant (Panadero et al., 2020).

Importantly, our interventions from 2020 and the current one were brief, each lasting less than one hour. Thus, the length of the intervention could have influenced our results along with other factors such as the educational experience of the participants. For that reason, we also explored the effect of year level. Here we found that the most experienced students, expected to have greater domain knowledge for the writing task, reported higher self-efficacy in contrast with our results from 2020 with were no effect was found (Panadero et al., 2020). However, there are two important differences between both studies. First, our participants were university students and the ones in 2020 were secondary education students. And second, probably even more important, in 2020 participants performed different Spanish and mathematics activities in each year level, while in the present study the three year levels performed the same task. In general, our results somewhat contradict the main trend of academic self-efficacy decreasing as students' progress in the educational system (Pajares, 2008). Nevertheless, research has shown that self-assessment realism can be greater for students with greater expertise in a domain (Boud & Falchikov, 1989), hence it is highly likely that greater academic opportunities and success contribute to positive increases in self-efficacy in self-assessment as students age through the system.

Regarding the effects of self-assessment on emotions, theory indicates that self-assessment can generate stronger effects on emotional and motivational reactions than on cognitive processes (Sitzmann et al., 2010). This is logical as a large part of

assessing one's own work is dealing with uncertainty, success or failure, all relating to emotional reactions. Our results also showed that rubrics in interaction with feedback occasion is the most powerful intervention of the three, even than the combination of rubrics and instructor's feedback. The rubric appears to have served as an external regulatory tool, mitigating the emotional impact of self-assessment on students. This effect may be attributed to students having the opportunity to reflect on performance criteria and levels encapsulated in the rubric during their self-assessment, even in the absence of externally produced feedback. This finding resonates with previous research (Andrade & Du, 2007; Panadero et al., 2013, 2014). With respect to academic year level, it was observed that the most experienced students reported higher levels of negative emotions. This could likely be due to a more stringent and ambitious self-assessment approach, driven by their ability to attain higher performance levels and more precise self-evaluation (Dochy et al., 1999; Yan et al., 2020).

An interesting result is that the combined condition –i.e., rubric and instructor's feedback- was not superior to the instructor alone condition. We hypothesize two reasons. First, it could be that the students experience a high cognitive load as previous research has found when rubrics and exemplars are combined (Lipnevich et al., 2022). Thus, when the participants received the instructor's feedback and the rubric, they had to process two sources of information which might have cost the students to remain at the same self-efficacy and emotional levels. Second, it could be that the instructor's feedback negative influence prevails even in combination with a rubric. At the end, self-assessment based on a rubric is less confrontational than receiving external feedback, and such might have provoked lower self-efficacy and worse emotional reactions when facing written feedback from the instructor.

Limitations and Future Lines of Research

First, this study is a small-scale experimental study situated in one Spanish university. Further, the study took place within one discipline (psychology) and one type of academic activity (writing an essay). Hence, applicability to other sites, disciplines, or kinds of activity is unknown. Second, the lack of a control group is much less a concern, given that alternative treatments rather than null treatments were contrasted, but should also be kept in mind. Third, the participants received feedback only about their essays, not about the validity or realism of the self-assessment itself; perhaps feedback about the self-assessment might have led bigger effects on the quality of achievement emotions and self-efficacy (e.g., Andrade, 2018; Panadero et al., 2016).

We want to point out two future lines of research: (a) the use of technology such as physiological reaction equipment (Azevedo et al., 2018) would allow us to reach a deeper, objective and real time understanding of neuro-emotional and motivational reactions; and (b) it seems important to determine if there are differences in how students of different genders respond to different kinds of feedback.

Conclusions

In conclusion, it appears that rubric-based feedback contributes to gains in positive emotions and reductions in negative emotions along with increased self-efficacy. It is important to remark that rubrics present features that might help university students to feel more capable and at ease such as clear criteria and performance levels to which they can anchor their own performance and evaluative judgment expressed as self-assessment. Rubrics have shown a positive outcome here to contribute positively to student emotional responses and self-efficacy when self-assessing. Additionally, more experienced students obtained better self-efficacy while self-assessing, with no changes on positive emotions, and an increase of negative emotions.

References

- Andrade, H. (2018). Feedback in the context of self-assessment. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback* (pp. 376-408). Cambridge University Press.
- Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation In Higher Education, 32*(2), 159-181.
<https://doi.org/10.1080/02602930600801928>
- Azevedo, R., Taub, M., & Mudrick, N. V. (2018). Understanding and reasoning about real-time cognitive, affective, and metacognitive processes to foster self-regulation with advanced learning technologies. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 254-270). Routledge.
- Barnett, J. E., & Hixon, J. E. (1997). Effects of grade level and subject on student test score predictions. *The Journal of Educational Research, 90*(3), 170-174.
<https://doi.org/10.1080/00220671.1997.10543773>
- Boud, D. (1995). Assessment and learning: contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in Higher Education* (pp. 35-48). Kogan Page.
- Brookhart, S. M., & Chen, F. (2015) The quality and effectiveness of descriptive rubrics. *Educational Review, 67*(3), 343-368, <https://doi.org/10.1080/00131911.2014.929565>
- Brown, G., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan, (Ed.), *SAGE handbook of research on classroom assessment* (pp. 367-393). SAGE.
- Brown, G. T. L., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research, 3*, 22-30. <https://doi.org/10.14786/flr.v2i1.24>

Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback:

Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, 86(4), 606-629.

<https://doi.org/10.1111/bjep.12126>

Butler, R. (2011). Are positive illusions about academic competence always adaptive, under all circumstances: New results and future directions. *International Journal of Educational Research*, 50(4), 251-256.

<https://doi.org/10.1016/j.ijer.2011.08.006>

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245-281.

<https://doi.org/10.3102/00346543065003245>

Chau, P. Y. (1997). Reexamining a model for evaluating information center success using a structural equation modeling approach. *Decision Sciences*, 28, 309-334.

<https://doi.org/10.1111/j.1540-5915.1997.tb01313.x>

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395-430.

Faul, F., Erdfelder, E., Lang, A.-G. y Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. <https://doi.org/10.3758/BF03193146>

Harris, L. R., & Brown, G. T. (2018). *Using self-assessment to improve student learning*. Routledge.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>

- Hawthorne, K. A., Bol, L., & Pribesh, S. (2017). Can providing rubrics for writing tasks improve developing writers' calibration accuracy? *The Journal of Experimental Education*, 85(4), 689-708. <https://doi.org/10.1080/00220973.2017.1299081>
- Jax, J., Ahn, J. N., & Lin-Siegler, X. (2019). Using contrasting cases to improve self-assessment in physics learning. *Educational Psychology*, 39(6), 815-838. <https://doi.org/10.1080/01443410.2019.1577360>
- Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2019). A meta-analysis on the impact of grades and comments on academic motivation and achievement: A case for written feedback. *Educational Psychology*, 1-22. <https://doi.org/10.1080/01443410.2019.1659939>
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, 6(2), 151-162.
- Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 23(9), 1256-1265. <https://doi.org/10.1002/acp.1528>
- Kostons, D., van Gog, T., & Paas, F. (2010). Self-assessment and task selection in learner-controlled instruction: Differences between effective and ineffective learners. *Computers & Education*, 54(4), 932-940. <https://doi.org/10.1016/j.compedu.2009.09.025>
- Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, 42(4), 539-559.
- Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated*

- learning. Theory, research and applications* (pp. 111-168). Lawrence Erlbaum Associates.
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806-813.
<https://doi.org/10.1016/j.lindif.2012.04.007>
- Panadero, E., Brown, G. T. L., & 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., 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., García-Pérez, D., Fernández-Ruiz, J., Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. (2022). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education, 1*-21. <https://doi.org/10.1007/s10212-022-00639-4>
- 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. <https://doi.org/10.1016/j.edurev.2017.08.004>
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105.
https://doi.org/10.1207/S15326985EP3702_4

- Pekrun, R., Muis, K. R., Frenzel, A. C., & Goetz, T. (2018). *Emotions at school*.
Routledge.
- Peterson, E. R., Brown, G. T. L., & Jun, M. C. (2015). Achievement emotions in higher education: A diary study exploring emotions across an assessment event. *Contemporary Educational Psychology, 42*, 82-96.
<https://doi.org/10.1016/j.cedpsych.2015.05.002>
- Raaijmakers, S. F., Baars, M., Paas, F., van Merriënboer, J. J., & van Gog, T. (2019). Effects of self-assessment feedback on self-assessment and task-selection accuracy. *Metacognition and Learning, 1-22*. <https://doi.org/10.1007/s11409-019-09189-5>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software, 48*(2), 1-36.
<https://doi.org/10.18637/jss.v048.i02>
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99*, 323-338.
<https://doi.org/10.3200/JOER.99.6.323-338>
- Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*(2), 169–191.
- Tan, K. H. (2012). How Teachers Understand and Use Power in Alternative Assessment. *Education Research International, vol. 2012*, 1-12.
<https://doi.org/10.1155/2012/382465>

- Yan, Z. (2018). Student self-assessment practices: The role of gender, school level and goal orientation. *Assessment in Education: Principles, Policy & Practice*, 25(2), 183-199. <https://doi.org/10.1080/0969594X.2016.1218324>
- Yan, Z. (2019). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation In Higher Education*, 1-15. <https://doi.org/10.1080/02602938.2019.1629390>
- Yan, Z., & Brown, G. T. (2017). A cyclical self-assessment process: towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, 42(8), 1247-1262. <https://doi.org/10.1080/02602938.2016.1260091>
- Yan, Z., Brown, G. T. L., Lee, J. C. K., & Qiu, X. L. (2019). Student self-assessment: Why do they do it? *Educational Psychology*, 1-24. <https://doi.org/10.1080/01443410.2019.1672038>
- Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022, 2022/11/01/). Effects of self-assessment and peer-assessment interventions on academic performance: A meta-analysis. *Educational Research Review*, 37, 100484. <https://doi.org/https://doi.org/10.1016/j.edurev.2022.100484>
- Winstone, N. E., Hepper, E. G., & Nash, R. A. (2021). Individual differences in self-reported use of assessment feedback: The mediating role of feedback beliefs. *Educational Psychology*, 41(7), 844-862. <https://doi.org/10.1080/01443410.2019.1693510>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10(3087). <https://doi.org/10.3389/fpsyg.2019.03087>

Appendix A: Self-efficacy questionnaire

WRITING OF ACADEMIC TEXTS SELF-EFFICACY QUESTIONNAIRE

Express how capable or competent you feel about the following statements, using the scale shown below:

0	1	2	3	4	5	6
Very strongly disagree	Strongly disagree	Rather disagree	Indifferent	Rather agree	Strongly agree	Very strongly agree

	1st Occasion	2nd Occasion	3rd Occasion
I usually feel capable to write academic texts			
When I write an academic text I know what steps to follow			
I can write different types of texts following specific communicative purposes			
When I write an academic text, I am able to apply the principles of coherence and cohesion between paragraphs			
I am able to write clear texts, considering elements such as the length of the sentences and the vocabulary used			
Normally, I am able to quote and make the list of references following APA standards			
I am able to self-assess my own tasks			
I am able to self-assess my own academic texts			

Appendix B: Emotions questionnaire

ACADEMIC EMOTIONS QUESTIONNAIRE

Indicate the intensity in which you are experiencing the following emotions at this time. 1 indicates very low intensity, and 10 maximum intensity.

Emotions	1st Occasion	2nd Occasion	3rd Occasion
Happiness			
Hope			
Pride			
Relief			
Anxiety			
Nervousness			
Shame			
Sadness			
Boredom			
Anger			

Appendix C. Rubric

CATEGORY	LOW QUALITY	AVERAGE QUALITY	HIGH QUALITY
Writing process	I started writing the text without planning what I wanted to write. I have hardly reread what I was writing and, when I finished, I have not reviewed the text or I have only looked for misspellings.	2 options: a) Before writing, I have planned what I wanted to communicate. At the end, I have hardly reviewed the text or I have only looked for misspellings. b) I started writing without thinking much about what I wanted to tell. However, I reviewed the text several times, looking for all or some of these factors: Text structure, coherence and connection between paragraphs, clarity of the message, style, and spelling.	Before writing, I thoroughly planned what I wanted to tell and how I was going to do it. I reviewed while I was writing and, at the end, I also reviewed the full text at least once. While reviewing, I looked for all or some of these factors: Text structure, coherence and connection between paragraphs, clarity of the message, style, and spelling.
Text components: Structure & coherence /connection between paragraphs	There is no clear structure, with an introduction, a crux, and a closing. Lack of incorrect use of text connectors and/or discourse markers. Regarding paragraphs, one of these two happens: a) The text has only one or two paragraphs, without clear internal and external coherence. b) The text has many very short paragraphs, which makes it difficult to follow the argument line.	A structure is somehow present (introduction, crux and closure) but could be more clearly delimited. Connectors are most of the times used appropriately. However, there may be one or more of these flaws: Same paragraph includes different unorganized ideas. Same idea in two paragraphs when it could be in one. The paragraph where the argument is developed is too long; it could be divided. Connector/text markers are misused.	There is a very clear structure in the text: including opening, argument crux and closing. Ideas are connected and presented in well-organized paragraphs. Connectors and/or discourse markers are effectively used.
Text components: Sentences, vocabulary & punctuation.	Sentences are too long (over 40 words) or too short. Excessive use of text insertions within sentences. Punctuation is incorrect (e.g. lack of commas, the break the sentence). Too many colloquial expressions. Abuse of passive or impersonal tenses.	Most sentences are of adequate length, with a few too long or short or incomplete. Punctuation is correct, although there may be a few mistakes. The vocabulary is adequate, but different terms are used to refer to the central concept of the text. Some colloquial expression may appear.	The sentences are well constructed, usually following a simple structure, in an active language and a coherent use of the verbs. Punctuation is correct. The vocabulary is adequate, and the main terms are used with precision.

Appendix D. Instructor feedback (three samples)

GRADE: 3,5

The text structure has important flaws. It does not follow a coherent argument; on the contrary, ideas change abruptly in each paragraph. For instance, any of the first three paragraphs could actually be the introduction paragraph because each of them present different ideas as it was the introduction. Later, in the argument crux there are several ideas without connection. Finally, the previous to the last paragraph seems to be closing the text but, nonetheless, there is an additional paragraph after it. Furthermore, that previous to the last paragraph includes a new idea (about the methodology), which has not been mentioned before and it could be used as an argument in favour of Psychology.

To sum, even though a central message can be perceived (the multiple areas of application of Psychology), it is not developed nor transmitted effectively. Regarding grammar, highlighted in the text there are mistakes and comments in the footnotes.

GRADE: 6,5

The text has a quite clear structure, with a paragraph of introduction, three for crux and a closing paragraph. However, there are two arguments in the introduction, and one of them is not developed in order to refute it (the skepticism of certain people). In addition, the last paragraph includes a new idea that has not been discussed before and it does not recap and finish with the main message to be transmitted. In general, there is a correct use of connectors and discourse markers.

Regarding the style and grammar, in general, the construction of the sentences is correct, and the vocabulary is appropriate. Nevertheless, there are some mistakes in the sentence construction and some limitations in the vocabulary selection, which are highlighted in the text and commented in footnotes.

GRADE: 9

The text has an adequate argumentative structure, with an introductory paragraph, four for the argument crux and a closing paragraph. Connectors and discourse marks are properly used.

Regarding the text style, it is correct considering the vocabulary, the use of punctuation marks and the sentence construction. There are some minor mistakes highlighted in the text and commented in footnotes.

Appendix E. Descriptive Statistics for Scores of Individual Positive and Negative Emotions Assessed, by Condition and Feedback Occasion

Emotions	Condition	Pre-assessment		Before feedback		After feedback	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Joy	Rubric	5.21	1.95	5.44	1.94	6.74	2.08
	Expert's feedback	5.51	2.07	5.40	2.15	5.72	2.31
	Combined	6.15	1.49	6.33	1.54	6.83	1.71
Hope	Rubric	4.98	1.91	4.95	2.03	6.28	2.06
	Expert's feedback	5.02	2.36	4.72	2.46	5.02	2.57
	Combined	5.60	1.91	5.70	2.00	5.88	2.22
Pride	Rubric	4.42	2.00	4.63	2.01	6.21	2.29
	Expert's feedback	4.56	2.03	4.42	2.22	4.67	2.56
	Combined	5.23	2.03	5.38	2.01	5.53	2.34
Relief	Rubric	4.28	2.24	5.60	2.11	6.81	1.97
	Expert's feedback	4.33	1.86	4.88	2.06	5.09	2.41
	Combined	4.60	2.18	5.28	2.00	5.83	2.21
Anxiety	Rubric	3.86	1.79	3.40	1.77	2.42	1.61
	Expert's feedback	3.53	2.19	3.37	1.95	3.12	1.87
	Combined	3.38	1.89	3.00	1.85	2.38	1.41
Nervousness	Rubric	5.23	1.93	4.47	2.10	3.02	1.99
	Expert's feedback	4.65	2.38	4.56	2.24	3.84	2.01
	Combined	4.35	2.17	3.75	1.89	3.08	1.90
Embarrassment	Rubric	4.30	2.30	4.16	2.31	3.09	2.19
	Expert's feedback	4.58	2.31	4.70	2.18	4.40	2.21
	Combined	3.78	2.36	3.35	2.33	3.13	2.28
Sadness	Rubric	1.74	1.24	1.63	1.00	1.49	0.94
	Expert's feedback	1.70	1.10	1.65	1.04	2.14	1.77
	Combined	2.15	1.59	2.08	1.46	1.95	1.50
Boredom	Rubric	2.14	1.51	1.81	1.28	1.53	0.88
	Expert's feedback	2.35	1.57	2.05	1.36	1.95	1.33
	Combined	2.20	1.52	1.85	1.27	1.78	1.25
Anger	Rubric	1.28	0.70	1.28	0.81	1.25	0.69
	Expert's feedback	1.27	0.81	1.24	0.80	1.41	1.07
	Combined	1.08	0.27	1.08	0.27	1.03	0.16

$N = 126$; $n_{\text{Rubric}} = 43$; $n_{\text{Expert's feedback}} = 43$; $n_{\text{Combined}} = 40$.

Appendix F. Descriptive Statistics for Scores of Individual Positive and Negative Emotions Assessed, by Year Level and Feedback Occasion

Emotions	Year level	Pre-assessment		Before feedback		After feedback	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Joy	First year	5.45	2.04	5.70	2.02	6.30	2.48
	Second year	5.68	1.90	5.60	2.02	6.58	1.88
	Third year	5.71	1.74	5.81	1.80	6.40	1.90
Hope	First year	5.27	2.09	5.20	2.18	5.82	2.48
	Second year	5.25	1.86	5.18	2.05	5.85	2.13
	Third year	5.05	2.27	4.95	2.40	5.50	2.41
Pride	First year	4.34	1.89	4.61	2.09	5.34	2.65
	Second year	4.75	2.03	4.80	2.03	5.38	2.55
	Third year	5.10	2.15	4.98	2.23	5.69	2.23
Relief	First year	4.27	2.15	5.55	2.15	5.93	2.56
	Second year	4.58	2.06	5.30	1.96	6.18	2.17
	Third year	4.36	2.08	4.90	2.06	5.64	2.15
Anxiety	First year	3.45	1.95	2.89	1.74	2.48	1.68
	Second year	3.50	2.04	3.30	1.90	2.45	1.63
	Third year	3.83	1.92	3.62	1.90	3.00	1.67
Nervousness	First year	4.86	2.39	4.27	2.37	3.23	2.24
	Second year	4.63	2.12	4.10	2.02	3.25	1.92
	Third year	4.76	2.05	4.43	1.90	3.48	1.81
Embarrassment	First year	4.41	2.51	3.95	2.36	3.34	2.40
	Second year	4.15	2.29	4.35	2.29	3.93	2.31
	Third year	4.12	2.21	3.98	2.33	3.40	2.15
Sadness	First year	1.66	1.16	1.59	1.06	1.91	1.65
	Second year	1.80	1.07	1.85	1.08	1.75	1.24
	Third year	2.12	1.66	1.90	1.39	1.90	1.46
Boredom	First year	2.25	1.59	1.93	1.39	1.80	1.17
	Second year	1.88	1.36	1.63	1.05	1.50	0.91
	Third year	2.55	1.56	2.14	1.39	1.95	1.36
Anger	First year	1.25	0.71	1.28	0.82	1.30	0.88
	Second year	1.14	0.42	1.06	0.23	1.17	0.56
	Third year	1.23	0.74	1.26	0.79	1.23	0.81

$N = 126$; $n_{1stYear} = 44$; $n_{2ndYear} = 40$; $n_{3rdYear} = 42$.