

Running head: Rubrics vs. scripts

Running head: Rúbricas y guiones

**Rubrics vs. self-assessment scripts: Effects on first year university students' self-regulation and performance**

Ernesto Panadero<sup>1,2</sup>, Jesús Alonso Tapia<sup>1</sup> and Juan Antonio Huertas<sup>1</sup>

<sup>1</sup>*Universidad Autónoma de Madrid*; <sup>2</sup>*University of Oulu (Finland)*

**Recommended citation:**

(for the English version: **preferred choice**)

Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014). Rubrics vs. self-assessment scripts: Effects on first year university students' self-regulation and performance. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37(1), 149-183. doi:10.1080/02103702.2014.881655

(for the Spanish version)

Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014). Rúbricas y guiones de autoevaluación: Efectos sobre la autorregulación y el rendimiento de estudiantes universitarios de primer año. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37(1), 149-183. doi:10.1080/02103702.2014.881655

This is a pre-print of an article published in *Infancia y Aprendizaje*. Personal use is permitted, but it cannot be uploaded in an Open Source repository. The permission from the publisher must be obtained for any other commercial purpose. This article 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 official published version.

*Abstract*

The objective of this study was to analyze the effects of rubrics and self-assessment scripts on self-regulation, learning, self-efficacy and goal activation. A total of 85 university students enrolled in a psychology course on Motivation and Emotion participated in the study. During a semester, they summarized texts and designed conceptual maps of each topic unit with the help of rubrics or scripts. Self-regulation, self-efficacy and goal activation were measured through questionnaires, and learning through the conceptual maps and summaries. Results showed that the use of rubrics increased learning but decreased the students' learning self-regulation more than the use of scripts. No significant differences were found for self-efficacy and students preferred using rubrics, reporting they focus more on learning when using these. Results are partially in line with previous research. Theoretical and educational implications are discussed.

*Keywords:* Self-regulation, self-evaluation, self-efficacy, self-assessment scripts, rubrics, formative assessment.

**Rúbricas y guiones de autoevaluación: efectos sobre la autorregulación y el rendimiento de estudiantes universitarios de primer año**

Ernesto Panadero<sup>1,2</sup>, Jesús Alonso Tapia<sup>1</sup> and Juan Antonio Huertas<sup>1</sup>

<sup>1</sup>Universidad Autónoma de Madrid; <sup>2</sup>Universitat Autònoma de Barcelona

*Resumen*

En este estudio se analizan los efectos del uso de rúbricas y guiones de autoevaluación en la autorregulación, aprendizaje, autoeficacia y la activación de metas. Participaron 85 estudiantes universitarios de psicología matriculados en el curso Motivación y Emoción. Durante un semestre resumieron textos y crearon mapas conceptuales de cada unidad didáctica impartida con la ayuda de rúbricas o guiones. Se midió la autorregulación, la autoeficacia y la activación de metas mediante cuestionarios, mientras que el aprendizaje se valoró a partir de los resúmenes y los mapas conceptuales. Se encontró que el uso de las rúbricas aumentó el aprendizaje pero disminuyó la autorregulación orientada al aprendizaje. No hubo diferencias en autoeficacia y los alumnos prefirieron el uso de rúbricas, e informaron de que éstas les orientaron más a metas de aprendizaje. Los resultados están parcialmente en consonancia con estudios anteriores. Se discuten las implicaciones teóricas y educativas.

*Palabras clave:* Autorregulación, autoevaluación, autoeficacia, guiones de autoevaluación, rúbricas, evaluación formativa.

As students move forward in the educational system, there is less guidance from teachers, they therefore need to be more proactive and take full responsibility for their own learning to complete their training successfully especially during higher education (Torenbeek, Jansen, & Hofman, 2010). Therefore, students need to use a variety of learning strategies that demonstrate self-regulation, which refers to “self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals” (Zimmerman, 2000, p. 14). Self-regulation is then essential for success at higher education where students need to adapt to more demanding learning situations (Pintrich, 2004; Wolters, 1998). However, to develop self-regulation students need room for practice, modelling, and feedback, which requires direct instruction that should be included in the university programmes (Zimmerman & Kitsantas, 2005).

A crucial skill to develop for being a successful self-regulated student is self-assessment (Puustinen & Pulkkinen, 2001). For students to be able to exert control over their performance, they need to be able to self-assess the process they are following and the final outcome reached to improve their execution in the future while using some criterion to assess whether they are progressing or not (Winne & Hadwin, 1998). If students are not able to self-assess, they cannot cope with their learning difficulties and, as a consequence, they do not progress as expected (Alonso-Tapia & Panadero, 2010; Zimmerman & Moylan, 2009).

Based on that premise the study of self-assessment and its effects on self-regulation and learning has increased in recent years (Andrade & Valtcheva, 2009; Taras, 2010). Alonso-Tapia and Panadero (2010) proposed a classification of interventions that promote self-assessment, including research on: (a) the effects of *rubrics* (Jonsson & Svingby, 2007), (b) the effects of *scripts, prompts and cues* (Bannert, 2009), and (c) self-assessment understood as *self-grading* (Falchikov & Boud,

1989). Rubrics and scripts are making the most important contributions whereas the results of the third are discouraging (Andrade & Valcheva, 2009; Falchikov & Boud, 1989; Sadler & Good, 2006). Therefore this study will explore differences between these two tools.

### **Rubric and Scripts: tools for self-assessment**

Rubrics are documents “that articulate the expectations for an assignment by listing the criteria or what counts, and describing levels of quality from excellent to poor” (Reddy & Andrade, 2010, p. 435). When using a rubric, students should compare their final outcome against the level of achievement showed in the rubric to assess and score their performance. Rubrics have also been used to enhance inter-judge reliability (Jonsson & Svingby, 2007).

Scripts contain specific, structured steps to complete a task, including the assessment criteria that are presented as questions which the students have to answer to themselves (Alonso-Tapia & Panadero, 2010). These questions offer the opportunity to reflect about their performance and to execute the task as an expert would do. An example would be on a script to write summaries: “Does my summary transmit the main idea from the text?”.

### **Prior research on the effect of rubrics and scripts**

When rubrics are well designed, they enhanced performance (Andrade & Valcheva, 2009; Jonsson & Svingby, 2007) As for scripts, their effect on learning has been studied, usually in laboratory settings, showing that students using them improved

their learning (Bannert, 2009; Berthold, Nückles, & Renkl, 2007; Kramarski & Dudai, 2009; Kramarski & Michalsky, 2009, 2010).

Prior research has found compelling evidence that script use increases self-regulation as they promote metacognitive reflections, self-monitoring, and self-assessment (Bannert, 2009; Kramarski & Dudai, 2009). Regarding rubrics, only two studies have explored their effect, actually comparing their effect on self-regulation to that of the scripts (Panadero, Alonso-Tapia, & Huertas, 2012; Panadero, Alonso-Tapia, & Reche, 2012). Panadero et al. found that scripts enhanced learning self-regulation more than rubrics, and rubrics produced superior results to those in the control group. Regarding performance/avoidance self-regulation, which represents self-regulatory actions motivated by performance and avoidance goals that are detrimental for learning (Alonso-Tapia, Panadero & Ruiz, 2012), there is an advantage of using rubrics as they decreased performance/avoidance self-regulation, more than scripts did.

Previous research has supported the hypothesis that promoting self-assessment may increase the perception of self-efficacy, which conducts to higher self-regulation and learning (Pajares, 2008). However, mixed results have been found, ranging from one study with no significant results (Panadero, Alonso-Tapia, & Reche, 2012) to two studies showing that the use of rubrics in interaction with other variables (gender and type of feedback) increases self-efficacy (Alonso-Tapia & Panadero, 2010; Andrade, Du, & Wang, 2008; Panadero, Alonso-Tapia, & Huertas, 2012).

In conclusion, there are two main differences between rubrics and scripts that promote different effects. First, rubrics provide differentiation of performance whereas scripts do not. Second, rubrics contain a set of samples describing different qualitative execution levels, whereas the scripts do not offer those samples, as the students have to create their answers based on the script questions. These two differences might

influence students' self-regulation and learning in different ways. Rubrics emphasize the assessment of the final product that might focus students' attention on grades. On the contrary, scripts emphasize the assessment of the process, which might enhance students' reflective actions. Due to this fact, rubrics might make students worry more about the score and activate performance and/or avoidance goals instead of learning goals, while scripts might focus students on learning goals during the forethought and execution phases. However, previous results (Panadero, Alonso-Tapia, & Huertas, 2012; Panadero, Alonso-Tapia, & Reche, 2012) were not conclusive perhaps due to the method of measure used. So, new evidence is needed on this point.

In consequence, we designed the present study to address some unanswered questions in previous research that explored the comparison between rubrics and scripts. In Panadero, Alonso-Tapia, & Huertas (2012), the participants, secondary students, carried out an academic task in a laboratory setting. It was unclear whether these results would be replicated with different students working in a natural context. In the second study (Panadero, Alonso-Tapia, & Reche, 2012), the participants were university students in a pre-service teacher course, in which they had to design multimedia presentations. The question afterwards was whether these results would be the same in a different natural context.

In this study, first year psychology students in a higher education setting allowed us to test ideas from Halonen et al. (2003) regarding the implementation of rubrics on psychology training. Two skills were chosen to be trained. First, summarizing text that has shown significantly improves students' writing (Mateos, Villalón, de Dios, & Martín, 2007). Second, conceptual mapping, this is a learning strategy that increases students' performance and an effective technique to evaluate students' knowledge (Berry & Chew, 2008; Jacobs-Lawson & Hershey, 2002).

In summary, the objective of this study is to explore the relative effects of rubrics and scripts on first year psychology students' learning, self-regulation, self-efficacy and goal activation. Our hypothesis is that, though both strategies will enhance self-regulation, scripts will be superior, that is, they will do it more effectively than rubrics, as they center student's attention on process rather than on outcome, which is what rubrics do. As for effects on learning and self-efficacy, given that self-regulation tends to be positively associated to learning and self-efficacy, it could be hypothesized again a superior effect of scripts. However, due to the inconsistency of previous results, the confidence on this hypothesis is lower than in the case of the hypothesis on the effects on self-regulation. Finally, scripts are expected to promote the adoption of learning goals, while rubrics are expected to promote performance goals due to their scoring feature.

## **Method**

### *Participants*

Participation in the study was voluntary for the sample of 85 first-year psychology undergraduate students enrolled in the course "Motivation and Emotion" for their first time. Most participants were females (88.2%). The mean age was 18.9 years ( $SD = .9$ ). Students had been randomly assigned to each group by university administration services when they enrolled into the course. Therefore, the two intact sections formed the two conditions also randomly assigned: 41 students (48.2%) in the rubric group and 44 students (51.8%) in the script group. Demographically there were no differences among conditions with the majority of students being natives (rubric 90%, script 86%) and the rest being international students.

### *Design*

This study is a quasi-experimental design with two experimental conditions (rubrics vs. scripts) for the dependent variables: learning and self-efficacy. In addition, for the dependent variable self-regulation the design is a 2 (pre and post) x 2 (tool: rubric or script).

### *Materials*

We used several instruments for assessing the three dependent variables.

- a) *Learning*. To design the summary/conceptual maps rubrics and scripts, expert models of these tasks were analyzed. Then, the summaries and conceptual maps that the students created were analyzed to evaluate the intervention effects on learning through the same rubrics and scripts used by the students.
- b) *Self-regulation*. Two different instruments were used following prior research suggestions (Boekaerts & Corno, 2005; Samuelstuen & Bråten, 2007):
  - *Emotion and Motivation Self-regulation Questionnaire (EMSR-Q)* (Alonso-Tapia, Panadero, & Ruiz, 2012). This questionnaire includes 20 items to be answered in a five-point Likert scale. They are included in five basic scales grouped in two general scales, *Learning self-regulation*, with 8 items and a reliability index (Cronbach's  $\alpha$ ) of .78, and *Performance/avoidance self-regulation*, with 12 items and a reliability index  $\alpha = .86$ . The first scale includes self-messages or actions orientated to learning goals, as for example: "I will plan the activity before starting to execute it". The highest the value in this scale, the more positive is self-regulation for learning. The second scale includes self-messages or actions showing lack of self-regulation or orientated to performance, as for example: "I am getting nervous. I don't



know how to do it". The highest the values in this scale, the more negative actions detrimental for learning occur.

- *Specific Self-Regulation Questionnaire (SSR-Q)*. This questionnaire was created for this study due to its specificity as the items needed to refer to the activity being performed. It includes two scales. The first scale has seven items whose content refers to specific self-regulatory actions related to the creation of a complex conceptual map summarizing a high amount of text content. For example, "What shall I include in this conceptual map?" Its reliability is over the recommended level ( $\alpha = .83$ ). The second has also 7 items whose content refers to specific self-regulatory actions related to writing a summary of a long text. For example, "Shall I change this part so my summary will be easier to understand?" Its reliability is over the recommended level ( $\alpha = .76$ ). Items have to be answered in a five point Likert scale.
- c) *Self-efficacy*. This questionnaire was created for this study due to its specificity. It includes two scales: a) *Summary self-efficacy scale* (6 items,  $\alpha = .74$ ), with items like: "Usually, I am able to transmit the main ideas in my summary", and b) *Conceptual-map self-efficacy scale* (5 items,  $\alpha = .77$ ), with items like: "Usually, I establish relationships between concepts in different columns of my conceptual map". Items are answered in a seven point Likert scale.
- d) *'Perceived level of help' and 'goal orientation promoted by the tools'*. These two variables were measured through 8 items answered in a seven point Likert scale and created for this study. These items were analysed individually because they referred to different aspects of the task that have identity on their own. So, they will be presented in the "results" section.

*Procedure*

“Motivation and Emotion” is a one semester compulsory course in the first year of the psychology program. Two intact sections enrolled in the course participated in this study. The same teacher taught the course following the same instructional structure in both groups. The course was organized around six didactic units. The students had to create a summary and a conceptual map for each didactic unit. At the beginning of the semester, the students received the corresponding tool, rubrics or scripts, and instructors modeled their use. For each summary and conceptual map they designed, they had to self-assess using either rubrics or scripts. Right after the presentation the students filled the self-regulation and the goal orientation questionnaires. Four weeks later the students received one-on-one feedback about their two first summaries and conceptual maps. They were also modeled again on the use of the tools. At the end of the semester the students submitted the rest of summaries and conceptual maps for evaluation and in the same session filled the self-regulation, self-efficacy, perception of the tool and goals promoted by the tool questionnaires. Final scores were given two weeks later.

*Analysis Procedure*

For the variable with only one measure, factorial ANOVAs were calculated (learning). For the variables with pre-post measure (self-regulation and self-efficacy) ANOVAs with repeated measures were calculated.

**Results<sup>1</sup>***Intervention effects on Learning*

Two researchers evaluated independently the portfolios summaries and

conceptual maps using the rubrics given to the students ( $Kappa = .91$ ). The rubric condition outperformed the script condition, see Table I for means and standards deviations. *Summaries:  $F(1, 81) = 4.42; p < .05; \eta^2 = .097$ , Rubric  $M = 3.11$ , Script  $M = 2.86$ . Conceptual maps:  $F(1, 81) = 9.25; p < .01; \eta^2 = .184$ , Rubric  $M = 2.75$ , Script  $M = 2.29$ .*

TABLE I

*Intervention effects on Self-Regulation*

*Learning self-regulation scale.* The interaction between type of treatment and occasion of measurement was significant,  $F(1, 83) = 5.31, p < .05, \eta^2 = .06$ , Rubric  $M = 43.28$ , Script  $M = 41.21$ . As can be seen in Figure 1, the rubric group had a higher learning self-regulation before intervention than the script group (Rubric  $M = 44.63$ , Script  $M = 40.89$ , Dif.:  $3.75, p < .01$ ). However, the rubric group decreased significantly on their learning self-regulation after the intervention (Rubric pre  $M = 44.63$ , Rubric post  $M = 41.93$ , Dif.:  $-2.7, p < .01$ ), whereas the script group did not vary significantly. In the post-measure, both groups were equal in their self-regulation. It seems, then, that *rubrics have a negative effect on learning self-regulation that scripts do not have.* However, as there was no control group, it cannot be concluded from these data that the use of scripts benefits self-regulation, only that prevented the decrease produced by rubrics.

FIGURE 1

*Performance/avoidance self-regulation scale.* It is important to remember that high scores on this scale indicate a lack of adequate self-regulation when the student struggles with difficulties. The interaction between type of treatment -rubric vs. script-

and occasion of measurement -pre post- was significant,  $F(1, 83) = 8.86, p < .001, \eta^2 = .096$ , Rubric  $M = 43.8$ , Script  $M = 45.92$ . As can be seen in Figure 2, both groups reported the same level before intervention. However, performance/avoidance self-regulation decreased after intervention in the rubric group (Rubric pre  $M = 45.19$ , Rubric post  $M = 42.41$  Dif.:  $-2.78, p = .06$ ), whereas such messages increased in the script group (Script pre  $M = 44.23$ , Script post  $M = 47.61$ , Dif.:  $3.38, p < .05$ ), being the difference between groups significant at the post measure (Rubric  $M = 42.41$ , Script  $M = 47.61$ , Dif.:  $5.2, p < .05$ ). In summary, both groups started with the same level of performance/avoidance self-regulation, but after the intervention the script students experienced an increase while the rubric students a decrease, which implies that *scripts have a negative effect on performance/avoidance self-regulation while rubrics do not.*

#### FIGURE 2

*Summary self-regulation specific scale.* The interaction between type of treatment and the occasion of measurement fell slightly short of the standard limits of significance,  $F(1, 83) = 3.644, p = .06, \eta^2 = .042$ , Rubric  $M = 20.36$ , Script  $M = 20.11$ . In Figure 3 it can be seen that, at the pre measure, the rubric group reported a higher level of specific self-regulation for summarizing while at the post measure the script group reported higher non-significant level. In conclusion, the meaning of the almost significant effect is that scripts “tend” to benefit learning self-regulation, increasing the degree of self-messages related to supervising the task process, whereas rubrics tend to have the opposite effect.

#### FIGURE 3

*Conceptual map specific self-regulation scale.* The interaction between type of

treatment and occasion of measurement was significant,  $F(1, 83) = 4.70, p < .05, \eta^2 = .054$ , Rubric  $M = 25.01$ , Script  $M = 24.88$ . As it can be seen in the Figure 4, whereas the rubric group shows a significant decrease in the level of specific self-regulation (Rubric pre  $M = 26.15$ , Rubric post  $M = 23.88$ , Dif.: 2.27;  $p < .001$ ), the script group did not shows a significant change. This effect means that rubrics “tend” to hinder specific conceptual-map self-regulation, decreasing the degree of self-messages related to supervising the task process, whereas scripts do not have this effect.

#### FIGURE 4

##### *Intervention effects on self-efficacy*

Regarding self-efficacy, intervention had no significant effects, neither for summaries,  $F(1, 83) = 1.18; p = .280, \eta^2 = .014$ , Rubric  $M = 22.39$ , Script  $M = 21.59$ , nor for conceptual maps,  $F(1, 83) = 0.001; p = .971, \eta^2 = .000$ , Rubric  $M = 16.73$ , Script  $M = 16.70$ .

##### *Intervention effects on students' perception of the usefulness of instruments*

The ‘perceived help’ was measured in two ways. First, during the ‘Feedback session’, students were asked about their perception of the following four aspects in a scale from 1 to 5.

- *Difficulty to perform the summary.* No significant effect,  $F(1, 38) = .482; p = .492, \eta^2 = .013$ .
- *Difficulty to perform the conceptual map.* No significant effect,  $F(1, 38) = 1.16; p = .287, \eta^2 = .030$ .
- *Perceived help of the instrument to perform the summary.* Significant effect favouring the rubric,  $F(1, 22) = 6.32; p < .05, \eta^2 = .231$ .

- *Perceived help of the instrument to perform the conceptual map.* Significant effect favouring the rubric,  $F(1,22) = 4.41$ ;  $p < .05$ ,  $\eta^2 = .167$ .

Second, in the last questionnaire students were asked how much their self-assessment instrument (rubric or script) helped them to perform the following eight self-regulatory activities:

- *Planning the task.* There was no significant effect,  $F(1, 83) = 3.064$ ;  $p = .084$ ;  $\eta^2 = .036$ .
- *During the task (performance).* Difference was significant, being the rubric better perceived,  $F(1, 83) = 12.3$ ;  $p < .001$ ;  $\eta^2 = .129$ .
- *At the end of the task.* Difference was significant, being the rubric better perceived,  $F(1, 83) = 14.08$ ;  $p < .001$ ;  $\eta^2 = .145$ .
- *To self-assess.* Difference was significant, being the rubric better perceived,  $F(1, 83) = 3.815$ ;  $p = .054$ ;  $\eta^2 = .044$ .
- *To remain calmed.* No significant effect,  $F(1, 83) = .934$ ;  $p = .337$ .
- *To focus on learning.* Difference was significant, being the rubric better perceived,  $F(1, 83) = 3.79$ ;  $p = .055$ ;  $\eta^2 = .044$ .
- *To focus on the grade.* There was no significant effect,  $F(1, 83) = .254$ ;  $p = .615$ .
- *To correct mistakes.* Difference was significant, being the rubric better perceived,  $F(1, 83) = 7.31$ ;  $p < .01$ ;  $\eta^2 = .081$ .

## Discussion

The main objective of this study was to compare the effects of two self-assessment tools -rubrics and scripts- in helping university students in a natural context to learn, to self-regulate, to increase their self-efficacy and how they influence the activation of different goals.

*Learning*

The use of rubrics was better for the creation of summaries and conceptual maps than the use of scripts. The fact that the rubrics enhanced learning in the present study is in line with prior research (Andrade et al., 2008; Sadler & Good, 2006). However, why did scripts not enhance learning at the same level as prior research (e.g. Berthold et al., 2007, Kramarski & Michalsky, 2009, 2010) has also found? There can be several explanations.

The first explanation has to do with the possibility that the effects of the self-assessment tools may be mediated by other variables such as the nature and structure of the task. In a previous study comparing the effect of both tools on performance while working on a highly structured task, designing multimedia material, and working with small groups (Panadero, Alonso-Tapia, & Reche, 2012), neither the rubric nor the script produced a better performance when compared to a control group. However, in the present study, the learning activities were less structured and the groups were bigger in terms of sample size, which may have had an impact on the intervention effect.

The second explanation of why rubrics seem to be a better strategy for enhancing performance than scripts may lie on the fact that rubrics give information regarding the students' advances and how these relate to the score. This may have produced a sort of *coaching*, because the rubric offers the final product that should be reached, while students using the script were confronted with the expert process with no information about the final product they should achieve. As previously suggested (Kyndt, Dochy, Struyven, & Cascallar, 2010), familiarity with task quality criteria in highly complex situations leads to a deeper level of processing. Consequently, rubrics could have been a better tool to reach an acceptable version of the final product.

Nonetheless, it is crucial to point out that the conceptual maps and summaries were evaluated and scored using the same rubric given to the student. Using script as the scoring tool for the evaluators does not seem as feasible as using rubrics, because the script emphasizes process assessment to which the teachers cannot access as they evaluate the final product. In addition, rubrics have demonstrated utility as a scoring tool that increases scorers' reliability (Jonsson & Svingby, 2007). Therefore, though the rubric had been the correct tool to grade the students' portfolio, giving the rubric to the students seems to be a sort of coaching biasing the evaluation of learning effects. Perhaps a more appropriate measurement of the learning effects of the self-assessment tools would be evaluating the students without the tools in a task other than the one trained for so that they would have to transfer what they have learned.

### *Self-regulation*

The use of rubrics seems to be more detrimental than scripts for the students' *learning self-regulation* as this type of positive self-regulation decreased after the intervention. In this sense, scripts are preferable to rubrics as they have been found in previous research to significantly increase learning self-regulation allowing the students to a better understanding of the procedure followed and the product obtained (Kramarski & Dudai, 2009; Panadero, Alonso-Tapia, & Huertas, 2012).

The increase of the level of *performance/avoidance self-regulation* in the script condition is an unexpected result that differs from prior research. In Panadero, Alonso-Tapia and Huertas' (2012) examination of a secondary education sample, they found no significant effect, whereas with university students (Panadero, Alonso-Tapia, & Reche, 2012) the use of rubrics decreased performance/avoidance self-regulation. As scripts do not have a scoring feature, students might have felt unsure about their performance, a



fact that could have triggered their negative thoughts and emotions (Panadero, 2011). Therefore, rubrics seem to be a better tool than scripts for decreasing avoidance/performance self-regulation in the case of university students. This might be because university students are more worried than secondary students about their grades. Accordingly, due to the rubric scoring feature, students might be able to check the quality of their performance when comparing it to the rubric and experience less negative thoughts (Andrade & Valtcheva, 2009). In conclusion, if results on learning and performance/avoidance self-regulation are considered together, it seems that there is a tendency for self-assessment tools to improve or hinder different characteristics of self-regulation. Therefore, it can be maintained that the success of a self-assessment intervention on self-regulation may depend on the type of task, measure and context (Panadero, Alonso-Tapia, & Reche, 2012). Also, rubrics might be preferable to scripts for tasks where it is crucial to identify different levels of accomplishment (Reitmeier & Vrchota, 2009), when the students are novel in the task so that they can have a clearer understanding of how the product should look like (e.g. foundational courses), or when the task has a high level of specificity and the steps to be followed do not allow for flexibility on the performance –reflection about the procedure is not required–. Scripts might be preferable when the students are familiar with the task and, therefore, they can reflect about the content and the procedure to perform the task (e.g. more advanced courses) or even the final product might be open to discussion and the students need to have a clear argument for the procedure they follow (e.g. a critique paper). As a consequence the teachers should choose what tool is more helpful to achieve the established goals.

*Self-efficacy*

The self-assessment tools did not have a significant different impact on self-efficacy, a finding which was in line with prior research in higher education (Panadero, Alonso-Tapia, & Reche, 2012). However, with secondary education students two studies have found that rubrics increase self-efficacy in interaction with gender (Andrade, Wang, Du, & Akawi, 2009), and that rubrics boost self-efficacy in interaction with mastery feedback (Panadero, Alonso-Tapia, & Huertas, 2012). Thus, it seems that rubrics, in interaction with other variables, increase the self-efficacy of secondary education students, whereas in higher education this effect has not been found.

A first possible explanation is that it may be necessary to receive feedback for self-efficacy to improve because students' reflect about their capacity to perform a task when they face situations in which they are confronted with unexpected results (van Dinther, Dochy, & Segers, 2010). Therefore, the more the students received feedback about their performance, the higher the chances of reflecting about their capacity, a reflection that may have an impact on their self-efficacy. In the present study the students received feedback three times over the course of a semester, which may not be an intense enough intervention to make the students reflect about their capacity to perform the task.

A second explanation might be that university students' self-efficacy is less easily influenced than secondary education students' (as the former have a longer educational experience and a more comprehensive concept about their self-efficacy), and thus less dependable on their performance on one particular course. Therefore the intervention might need to be longer in time to produce an impact on university students' self-efficacy.

A third explanation is that, as no control group participated in this study, no results were available as to whether the intervention had been successful when

compared to a baseline group. Even if a control group was used in Panadero, Alonso-Tapia and Reche (2012) and no significant differences were found between rubric, script and control group, in that study the students did not receive feedback along the semester. Therefore, neither this study (because of the lack of a control group) nor the just mentioned by Panadero et al. (because of the lack of feedback which is crucial for self-efficacy interventions; van Dinther et al., 2010) can be considered conclusive regarding the impact of self-assessment tools on university students' self-efficacy.

*Subjective perception of Instrument usefulness and of goals activation*

In the 'Feedback session' data there was not a significant effect of the self-assessment instrument on the students' perception of the difficulty for writing summaries and designing conceptual maps. Actually, the low effect sizes of that effect show little connection between the use of the rubrics and scripts and how difficult the students perceive the tasks, which may depend on other variables (e.g. previous experience). However, there were significant differences in the students' perception of the self-assessment instrument usefulness. The students using rubrics reported that their instrument helped them to write the summaries and to design the conceptual maps more than students using scripts. Thus, the rubric was perceived as a better instrument, what seems logical as the use of rubrics seems to have a "coaching" effect.

When the perception of instruments usefulness was measured by questionnaire, there were significant differences in five out of the eight factors measured favouring the use of the rubric –a sixth item felt slightly short of the standard limits of significance–. The use of rubrics was perceived as more helpful in the three phases of the task (planning, performance and self-reflection), and also to self-assess, to focus on learning, to correct mistakes, and –almost significant– to remain calm. In the light of these results, it seems

that students using rubrics perceived them as more helpful than the students using scripts perceived the latter, and also that the use of rubrics promoted the election of learning goals more than the use of scripts did. These effects may be due to the security the students felt based on the rubric grading system or to the fact that the rubrics are easier to use as they present the assessment standards, in comparison to the reflective questions provided by the script.

In relation to the subjective usefulness for goal activation, students seem to consider that rubrics are better than scripts because, contrary to our expectations, they do not focus students in the grade more than the use of scripts does. However, these results have to be considered with caution, as they rely on a self-report measure based on two items.

In sum, our results explore some of these self-assessment instruments effects. On one hand, considering the objective effects, the rubrics seem to hinder more learning self-regulation, and scripts “tend” to favour it, in line with previous research (Alonso-Tapia & Panadero, 2010; Bannert, 2009). Rubrics seem to favour performance, though this effect may be due to coaching. So we suggest the use of transfer activities to assess the effects of both instruments. On the other hand, considering the subjective effects, the students seem to perceive rubrics as more helpful, and so they feel comfortable with its use, as previous research points out (Andrade & Du, 2005; Reynolds-Keefer, 2010).

### *Limitations*

There are several limitations to this study. First, self-regulation was only measured through questionnaires. Though two types of questionnaires have been used as recommended (Boekaerts & Corno, 2005), it would be preferable to use also, for example, *thinking aloud protocols* (Alonso-Tapia & Panadero, 2010). However, unless

students work on computers, it is difficult to use different measures when research is carried out in natural contexts and there is research that support the approach used in this study (Samuelstuen & Bråten, 2007).

Second, the size of the experimental conditions was of medium size and, even though there was control over variables to test the comparability of the groups, nesting effect should be kept in mind.

Third, although the initial design included a control group, problems along the intervention invalidated this data. For this reason the gain promoted by the use of the self-assessment instruments could not be compared against a group not using them, which is especially relevant for self-efficacy. In consequence, future research should continue exploring the impact of using self-assessment tools in university students'.

Fourth, research does not provide convergent results and it is necessary a systematic study of the moderating variables effects on interventions using rubrics and scripts (type of task, type of students, etc.).

In conclusion, the comparison of the relative effects of rubrics and scripts has shown that both have advantages and limitations when used with university students. Rubrics seem to decrease negative self-messages and increase learning supporting Halonen et al. (2003) framework for working on psychology courses using rubrics, although they tend to hinder learning self-regulation, whereas scripts seems to result in higher levels of learning self-regulation. It is then crucial that teachers reflect about which of the two tools is better for the instructional goals of the task.

### *Notes*

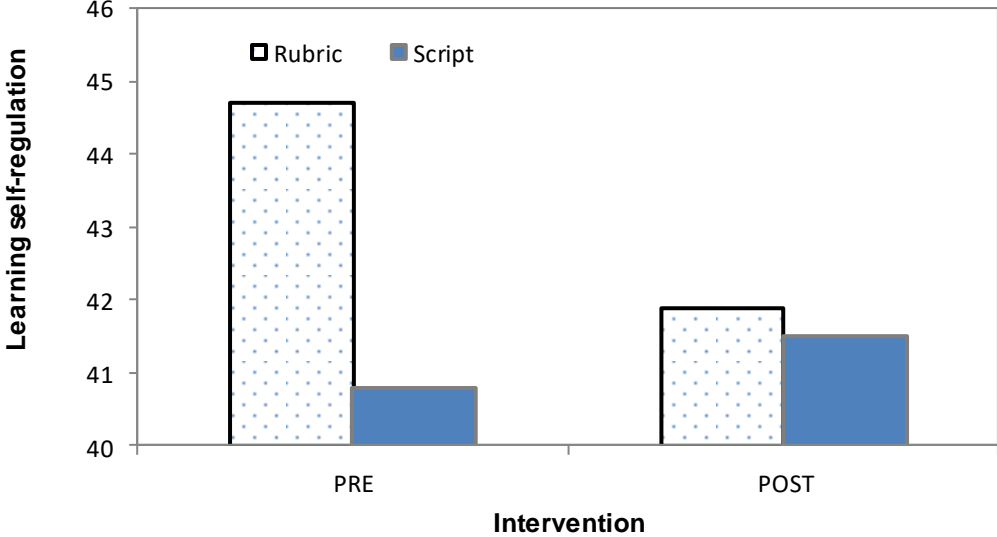
<sup>1</sup> Participants' goal orientation was controlled measured using the LEMEX questionnaire (Alonso-Tapia, Huertas & Ruiz, 2010). The results showed that both groups were equalled in terms of the participants' goal orientations.

Table I

*Means and standards deviations*

Tool	Learning		Self-regulation								Self-efficacy		
	Summ.	CM	Learning Sr.		Perf./Av. Sr.		Spec. Sr. CM		Spec. Sr. Su.		Summ.	CM.	
			Pre	Post	Pre	Post	Pre	Post	Pre	Post			
Rubric	<i>M</i>	3.11	2.75	44.63	41.93	45.19	42.41	26.15	23.88	20.83	19.90	22.39	16.73
	<i>SD</i>	.35	.45	1.10	1.10	1.82	1.72	.78	.96	4.45	6.07	3.53	3.25
Script	<i>M</i>	2.86	2.29	40.89	41.54	44.23	47.61	24.77	25.00	19.73	20.50	21.59	16.70
	<i>SD</i>	.34	.44	1.06	1.06	1.76	1.66	.75	.93	3.98	4.30	3.24	3.51

Note. Summ. = Summary; CM = Conceptual Map; Sr. = Self-regulation; Perf./Av. Sr. = Performance/Avoidance self-regulation; Spec. Sr. CM = Specific Self-regulation for Conceptual Maps; Spec. Sr. Su. = Specific Self-regulation for Summary.

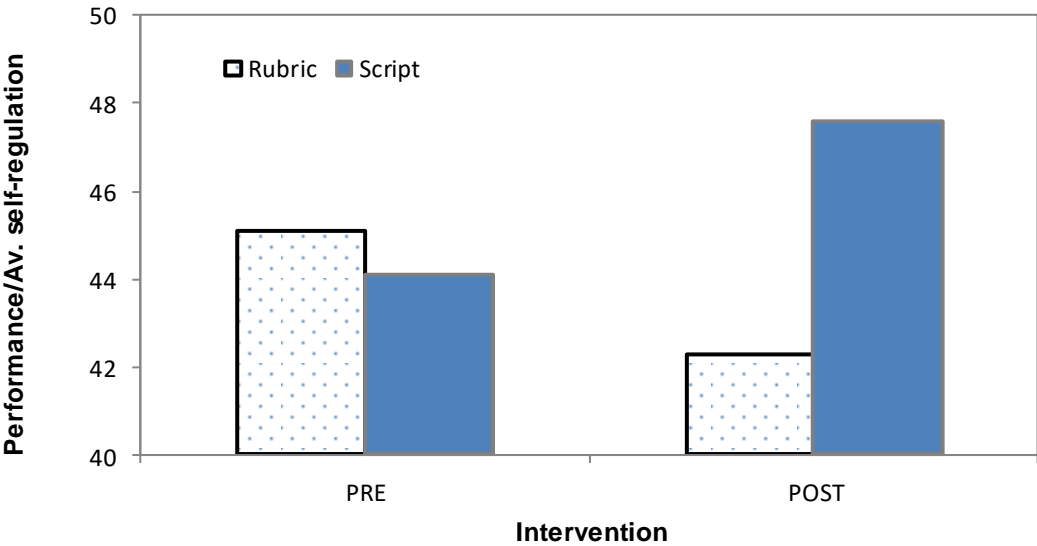


---

Figure 1

Interaction effect between type of self-assessment tool and occasion on Learning self-regulation

---

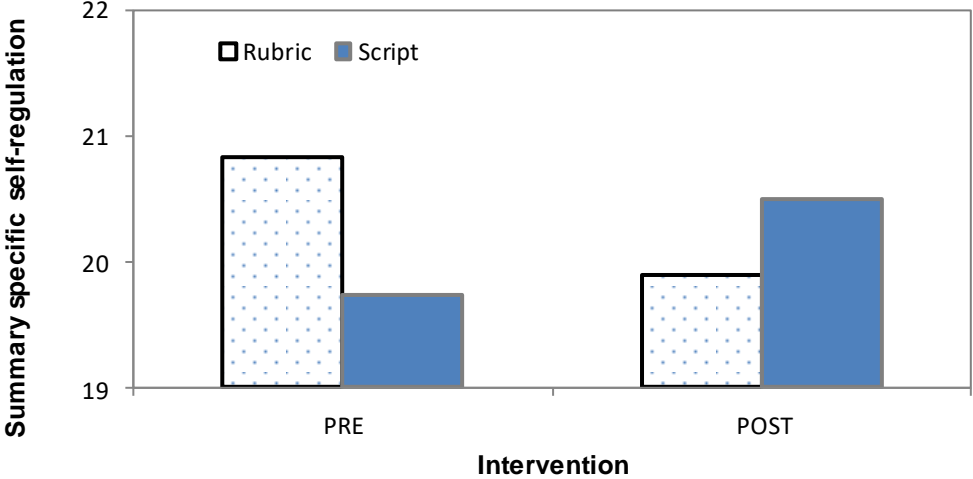


---

Figure 2  
Interaction effect between type of self-assessment tool and occasion on  
Performance/avoidance self-regulation

---

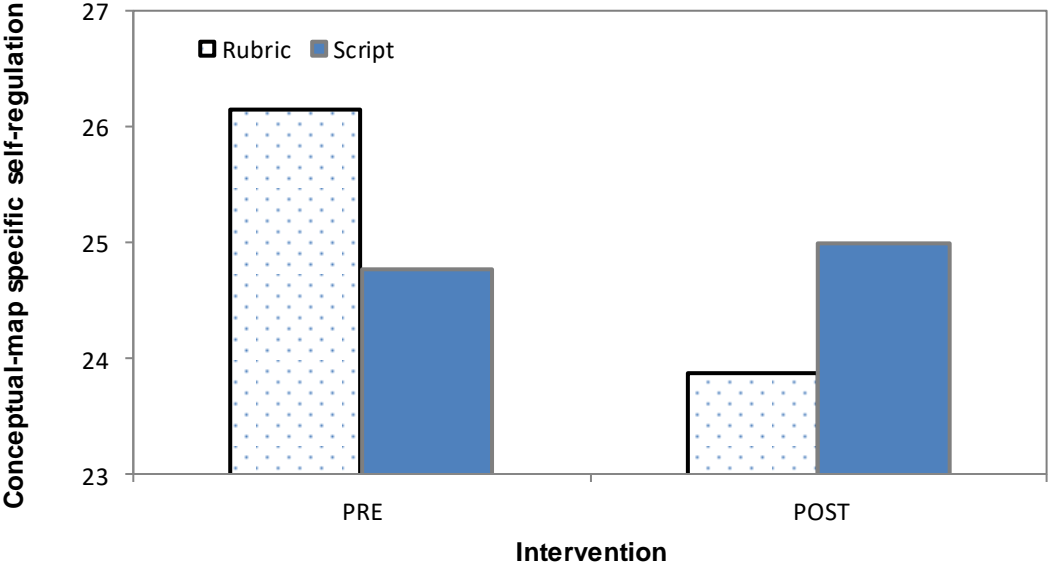




---

Figure 3  
Interaction effect between type of self-assessment tool and occasion on Summary specific self-regulation

---



---

Figure 4  
Interaction effect between type of self-assessment tool and occasion on Conceptual Maps  
specific self-regulation

---

## Appendix A

### Rubric: how to write a summary

Score Criteria	4	3	2	1
<b>The main idea from the original text is included</b>	The summary starts with the main idea of the original text	It contains the main idea but not at the beginning	What is proposed on the summary as the main idea is not	The summary does not contain the main idea
<b>The secondary ideas that support the main idea are included</b>	The rest of important ideas are in the summary with economy of words	The rest of important ideas are in the summary but it could have been done with less words	Some of the important ideas are in the summary but not all and it could have been done with less words	The summary contains elements that are superficial and there are too many words

(Continue).

## Appendix B

### Self-assessment script: how to write a summary

- Does my summary transmit the main idea from the text? Is it at the beginning of my summary?
- Also, are the important ideas also on my summary?
- Have I selected the main ideas from the text to make them explicit on my summary?
- Have I thought about my purpose for the summary? What is my goal?

(Continue).

### **Rúbricas y guiones de autoevaluación: efectos sobre la autorregulación y el rendimiento de estudiantes universitarios de primer año**

A medida que los estudiantes van avanzando en el sistema educativo, los profesores dejan de ser una guía para ellos, y deben por tanto ser más proactivos y responsabilizarse de su propio aprendizaje si quieren obtener una formación satisfactoria, especialmente en la educación superior (Torenbeek, Jansen, & Hofman, 2010). Así, los estudiantes deben usar varias estrategias de aprendizaje basadas en la autorregulación, que se entiende como “los pensamientos, sentimientos y acciones autogenerados y que se planifican y adaptan cíclicamente al logro de metas personales” (Zimmerman, 2000, p. 14). La autorregulación es por tanto esencial para el éxito en la educación superior, que exige a los estudiantes que se adapten a situaciones de aprendizaje cada vez más demandantes (Pintrich, 2004; Wolters, 1998). Sin embargo, para desarrollar la autorregulación los estudiantes deben crear espacios para la práctica, el modelado y la retroalimentación, lo que requeriría que los docentes universitarios incluyeran estos elementos como parte de los programas de instrucción directa (Zimmerman & Kitsantas, 2005).

El desarrollo de la habilidad de autoevaluación es crucial para que el estudiante aprenda a autorregularse (Puustinen & Pulkkinen, 2001). Para que los estudiantes puedan controlar su rendimiento deben ser capaces de autoevaluar el proceso en el que están inmersos y el resultado final alcanzado, y así mejorar su ejecución en el futuro, valiéndose de un criterio que les permita averiguar si están progresando o no (Winne & Hadwin, 1998). Si no aprenden a autoevaluarse no podrán hacer frente a las dificultades que se les presenten en el proceso de aprendizaje y, en consecuencia, no progresarán como esperan (Alonso-Tapia & Panadero, 2010; Zimmerman & Moylan, 2009).

Basándose en esa premisa, el estudio de la autoevaluación y sus efectos en la autorregulación y el aprendizaje se ha impulsado en los últimos años (Andrade & Valcheva, 2009; Taras, 2010). Alonso-Tapia y Panadero (2010) propusieron una clasificación de las intervenciones que fomentan la autoevaluación, incluyendo las investigaciones sobre: (a) los efectos de las *rúbricas* (Jonsson & Svingby, 2007), (b) los efectos de los *guiones*, *indicios* y *pistas* (Bannert, 2009), y (c) la autoevaluación entendida como *autocalificación* (Falchikov & Boud, 1989). Los estudios que han hecho las aportaciones más importantes se están realizando en el área de las rúbricas y los guiones, mientras que las aportaciones de los estudios que trabajan en la tercera línea son algo desalentadoras (Andrade & Valcheva, 2009; Falchikov & Boud, 1989; Sadler & Good, 2006). Por tanto, en el presente estudio exploraremos las diferencias entre las dos primeras.

### **Rúbricas y guiones: herramientas para la autoevaluación**

Las rúbricas son documentos “que articulan las expectativas para una tarea, enumerando los criterios, o qué es lo que cuenta, y describiendo niveles de calidad, de excelente a pobre”

(Reddy & Andrade, 2010, p. 435). Al usar una rúbrica, los estudiantes deberían comparar su resultado final con el nivel de logro especificado en la rúbrica, evaluando y puntuando así su rendimiento. Las rúbricas también se han usado para mejorar la fiabilidad interjueces (Jonsson & Svingby, 2007).

Los guiones especifican pasos concretos y estructurados para la realización de una tarea, incluyendo los criterios de evaluación, que se presentan como preguntas que los estudiantes tienen que contestarse a sí mismos (Alonso-Tapia & Panadero, 2010). Esas preguntas les dan la oportunidad de reflexionar sobre su rendimiento, y de llevar a cabo la tarea como lo haría un experto. Un ejemplo de un guión para escribir resúmenes sería el siguiente: “¿Transmite mi guión la idea principal del texto?”.

### **Investigación previa sobre los efectos de las rúbricas y los guiones**

Cuando están bien diseñadas, las rúbricas mejoran el rendimiento (Andrade & Valtcheva, 2009; Jonsson & Svingby, 2007). En cuanto a los guiones, su efecto sobre el aprendizaje se ha estudiado por lo general en laboratorio, mostrando que mejoraron el aprendizaje de los estudiantes que los utilizaron (Bannert, 2009; Berthold, Nückles, & Renkl, 2007; Kramarski & Dudai, 2009; Kramarski & Michalsky, 2009, 2010).

La investigación previa arroja pruebas convincentes de que el uso de guiones incrementa la autorregulación, ya que fomentan la reflexión metacognitiva, la autosupervisión y la autoevaluación (Bannert, 2009; Kramarski & Dudai, 2009). Respecto a las rúbricas, solo dos estudios han investigado sus efectos, comparando de hecho su efecto sobre la autorregulación con el de los guiones (Panadero, Alonso-Tapia, & Huertas, 2012; Panadero, Alonso-Tapia, & Reche, 2012). Panadero et al. encontraron que los guiones mejoraron la autorregulación del aprendizaje más que las rúbricas, y las rúbricas produjeron resultados superiores a los del grupo control. Respecto a la autorregulación de rendimiento/evitación, es decir, acciones de autorregulación motivadas por el rendimiento y por metas de evitación que son perjudiciales para el aprendizaje (Alonso-Tapia, Panadero & Ruiz, en prensa), se encontró que el uso de rúbricas aportaba una ventaja, ya que redujeron la autorregulación de rendimiento/evitación más que los guiones.

La investigación previa ha prestado apoyo a la hipótesis de que fomentar la autoevaluación puede incrementar la percepción de autoeficacia, lo que lleva a una mejor autorregulación y un mayor aprendizaje (Pajares, 2008). Sin embargo, se han obtenido resultados encontrados, desde un estudio que no obtuvo resultados significativos (Panadero, Alonso-Tapia, & Reche, 2012) a dos estudios que encontraron que el uso de rúbricas, en interacción con otras variables (sexo y tipo de retroalimentación) incrementó la autoeficacia (Alonso-Tapia & Panadero, 2010; Andrade, Du, & Wang, 2008; Panadero, Alonso-Tapia, & Huertas, 2012).

En definitiva, hay dos diferencias principales entre las rúbricas y los guiones y los efectos de cada uno de ellos. En primer lugar, las rúbricas aportan una diferenciación del rendimiento,

algo que no hacen los guiones. En segundo lugar, las rúbricas contienen un conjunto de ejemplos que describen distintos niveles cualitativos de ejecución, mientras que los guiones no aportan dichos ejemplos, ya que los estudiantes deben crear sus propias respuestas, basadas en las preguntas que se incluyen en los guiones. Estas diferencias se plasman en distintas vías de influencia sobre la autorregulación y el aprendizaje de los estudiantes. Las rúbricas enfatizan la evaluación del producto final, lo que puede hacer que los estudiantes presten su atención a las calificaciones. Por el contrario, los guiones enfatizan la evaluación del proceso, lo que puede mejorar las acciones reflexivas de los estudiantes. Debido a ello, las rúbricas pueden hacer que los estudiantes se preocupen más de la puntuación y activen metas de rendimiento y/o evitación en lugar de metas de aprendizaje, mientras que los guiones pueden enfocar más a los estudiantes hacia metas de aprendizaje en las fases de planificación y ejecución. Sin embargo, los resultados previos (Panadero, Alonso-Tapia, & Huertas, 2012; Panadero, Alonso-Tapia, & Reche, 2012) no fueron concluyentes, tal vez por el método de medida empleado. Es por tanto necesario recabar más pruebas sobre este punto.

Así, diseñamos el presente estudio para responder algunas preguntas, sin respuesta en investigaciones previas, que exploran la comparación entre rúbricas y guiones. En Panadero, Alonso-Tapia y Huertas (2012) los participantes, estudiantes de secundaria, realizaron una tarea académica en un laboratorio. No quedó claro si estos resultados podrían replicarse con otros estudiantes en un contexto natural. En el segundo estudio (Panadero, Alonso-Tapia, & Reche, 2012) los participantes eran estudiantes universitarios de un curso de formación de profesorado en prácticas, en las que tenían que diseñar presentaciones multimedia. En este caso, la cuestión es si los resultados serían los mismos en un contexto natural diferente.

En este estudio, la utilización de una muestra de estudiantes de psicología de primer año en un contexto de educación superior nos permitió poner a prueba las ideas de Halonen et al. (2003) respecto a la introducción de rúbricas en la enseñanza de la psicología. Se escogieron dos habilidades a entrenar. En primer lugar, se ha comprobado que el resumen de textos mejora significativamente la escritura de los estudiantes (Mateos, Villalón, de Dios, & Martín, 2007). En segundo lugar, los mapas conceptuales; una técnica que mejora el rendimiento de los estudiantes, y que es efectiva para evaluar su conocimiento (Berry & Chew, 2008; Jacobs-Lawson & Hershey, 2002).

En síntesis, el objetivo de este estudio es explorar los efectos relativos de las rúbricas y los guiones sobre el aprendizaje, la autorregulación, la autoeficacia y la activación de metas de estudiantes de primer año de psicología. Nuestra hipótesis es que, aunque ambas estrategias mejorarán la autorregulación, los guiones serán superiores, es decir, lo harán de manera más efectiva que las rúbricas, ya que dirigen la atención del estudiante hacia el proceso más que hacia el resultado, que es lo que hacen las rúbricas. Respecto a los efectos sobre el aprendizaje y la autoeficacia, dado que la autorregulación tiende a asociarse positivamente con el aprendizaje y la autoeficacia, podría hipotetizarse de nuevo un efecto superior de los guiones.

## Método

### *Participantes*

La participación en este estudio era voluntaria para la muestra de 85 estudiantes de grado de psicología que se habían inscrito en la asignatura “Motivación y Emoción” por primera vez. La mayoría de los participantes eran mujeres (88.2%). La media de edad era de 18.9 años ( $DT = .9$ ). Los servicios administrativos de la universidad asignaron a los estudiantes a cada grupo de forma aleatoria cuando se inscribieron en la asignatura. Por tanto, se dejaron ambas secciones intactas para formar las dos condiciones, también asignadas aleatoriamente: 41 estudiantes (48.2%) en el grupo de rúbricas, y 44 estudiantes (51.8%) en el grupo de guiones. No había diferencias demográficas entre ambas condiciones, siendo la mayoría de los estudiantes nativos (rúbricas 90%, guiones 86%) y el resto estudiantes de otros países.

### *Diseño*

Este estudio es un diseño cuasi-experimental con dos condiciones experimentales (rúbricas y guiones) para las variables dependientes: aprendizaje y autoeficacia. Además, para la variable dependiente autorregulación el diseño es un 2 (pre y post) x 2 (herramienta: rúbrica o guión).

### *Materiales*

Empleamos varios instrumentos para evaluar las tres variables dependientes.

- a) *Aprendizaje*. Para diseñar las rúbricas y guiones del resumen y de los mapas conceptuales se analizaron modelos expertos de estas tareas. Después se analizaron los resúmenes y mapas conceptuales creados por los estudiantes para evaluar los efectos de la intervención sobre el aprendizaje, usando las mismas rúbricas y los mismos guiones empleados por los estudiantes.
- b) *Autorregulación*. Se usaron dos instrumentos distintos, siguiendo las sugerencias aportadas por estudios previos (Boekaerts & Corno, 2005; Samuelstuen & Bråten, 2007):
  - *Cuestionario de Autorregulación de la Emoción y la Motivación (Emotion and Motivation Self-regulation Questionnaire, EMSR-Q, Alonso-Tapia, Panadero, & Ruiz, in press)*. Este cuestionario incluye 20 ítems a responder en una escala Likert de cinco puntos. Se inscriben en cinco escalas básicas, agrupadas en dos escalas generales, *Autorregulación del aprendizaje*, con 8 ítems y un índice de fiabilidad (alfa de Cronbach) de .78, y *Autorregulación de rendimiento/evitación*, con 12 ítems y un índice de fiabilidad  $\alpha = .86$ . La primera escala incluye mensajes autodirigidos o acciones orientados a metas de aprendizaje, como por ejemplo: “Planificaré la actividad antes de comenzar a ejecutarla”. Cuanto más alto sea el valor en esta escala, más positiva es la autorregulación del aprendizaje. La segunda escala incluye mensajes autodirigidos o acciones que muestran ausencia de autorregulación o que están orientados al rendimiento, como por ejemplo: “Me

estoy poniendo nervioso. No sé cómo hacerlo”. Cuanto más alto sea el valor en esta escala, más acciones negativas para el aprendizaje están ocurriendo.

- *Cuestionario Específico de Autorregulación (Specific Self-Regulation Questionnaire, SSR-Q)*. Este cuestionario se creó para este estudio, ya que los ítems debían referirse específicamente a la actividad que se estaba llevando a cabo. Incluye dos escalas. La primera tiene siete ítems cuyo contenido se refiere a las acciones autorreguladoras específicamente ligadas a la creación de un mapa conceptual complejo que resuma una gran cantidad de contenido textual. Por ejemplo, “¿Qué debo incluir en este mapa conceptual?” Su fiabilidad se situó por encima del nivel recomendado ( $\alpha = .83$ ). La segunda tiene también 7 ítems cuyo contenido se refiere a las acciones autorreguladoras específicamente ligadas a la escritura de un resumen de un texto extenso. Por ejemplo, “¿Debo cambiar esta parte para que mi resumen sea más fácil de entender?”. Su fiabilidad se sitúa por encima del nivel recomendado ( $\alpha = .76$ ). La respuesta a los ítems se da usando una escala Likert de cinco puntos.
- c) *Autoeficacia*. Este cuestionario se creó para este estudio por su especificidad. Incluye dos escalas: a) *Escala de autoeficacia para el resumen* (6 ítems,  $\alpha = .74$ ), con ítems como: “Por lo general, soy capaz de transmitir las ideas principales en mi resumen”, y b) *Escala de autoeficacia para el mapa conceptual* (5 ítems,  $\alpha = .77$ ), con ítems como: “Por lo general, establezco relaciones entre conceptos en columnas distintas de mi mapa conceptual”. La respuesta a los ítems se da usando una escala Likert de siete puntos.
- d) *‘Nivel de ayuda percibido’ y ‘orientación a metas fomentada por las herramientas’*. Estas dos variables se midieron mediante 8 ítems, a los que se respondía con una escala Likert de siete puntos, y que se crearon específicamente para este estudio. Los ítems se analizaron individualmente porque se referían a distintos aspectos de la tarea que tienen identidad propia. Por ello los presentaremos en la sección de resultados.

### *Procedimiento*

“Motivación y Emoción” es una asignatura semestral obligatoria del primer año del programa de psicología. Los dos grupos inscritos en la asignatura participaron intactos en este estudio. El mismo profesor dio las clases a ambos grupos siguiendo la misma estructura instruccional. La asignatura se organiza alrededor de seis unidades didácticas. Los estudiantes debían crear un sumario y un mapa conceptual para cada unidad didáctica. Al principio del semestre, los estudiantes recibieron la herramienta correspondiente, rúbricas o guiones, y los profesores modelaron su uso. Para cada sumario y mapa conceptual que diseñaron debían autoevaluarse, usando rúbricas o guiones según el caso. Inmediatamente después de la presentación de la



asignatura, los estudiantes rellenaron los cuestionarios de autorregulación y orientación a metas. Cuatro semanas más tarde recibieron retroalimentación, de forma individual, sobre sus dos primeros resúmenes y mapas conceptuales. Se les modeló de nuevo en el uso de las herramientas. Al final del semestre los estudiantes entregaron el resto de resúmenes y mapas conceptuales para su evaluación, y en la misma sesión rellenaron los cuestionarios sobre autorregulación, autoeficacia, percepción de la herramienta y metas fomentadas por la herramienta. Se les dieron las puntuaciones finales dos semanas más tarde.

#### *Procedimiento de análisis*

Para la variable con una sola medida (aprendizaje) se calcularon ANOVAs factoriales. Para las variables con medidas pre-post (autorregulación y autoeficacia) se calcularon ANOVAs de medidas repetidas.

### **Resultados<sup>1</sup>**

#### *Efectos de la intervención sobre el aprendizaje*

Dos investigadores evaluaron independientemente los portafolios de resúmenes y mapas conceptuales utilizando las rúbricas entregadas a los estudiantes ( $Kappa = .91$ ). La condición rúbrica obtuvo mejores resultados que la condición guión (ver tabla I para las medias y desviaciones estándar). *Resúmenes*:  $F(1, 81) = 4.42$ ;  $p < .05$ ;  $\eta^2 = .097$ , Rúbrica  $M = 3.11$ , Guión  $M = 2.86$ . *Mapas conceptuales*:  $F(1, 81) = 9.25$ ;  $p < .01$ ;  $\eta^2 = .184$ , Rúbrica  $M = 2.75$ , Guión  $M = 2.29$ .

TABLA I

#### *Efectos de la intervención sobre la autorregulación*

*Escala de autorregulación del aprendizaje*. La interacción entre tipo de tratamiento y momento de medida fue significativa,  $F(1, 83) = 5.31$ ,  $p < .05$ ,  $\eta^2 = .06$ , Rúbrica  $M = 43.28$ , Guión  $M = 41.21$ . Como se puede ver en la Figura 1, el grupo de rúbricas tenía una autorregulación del aprendizaje antes de la intervención más alta que el grupo de guiones (Rúbrica  $M = 44.63$ , Guión  $M = 40.89$ , Dif.: 3.75,  $p < .01$ ). Sin embargo, la autorregulación del aprendizaje del grupo de rúbricas descendió significativamente después de la intervención (Rúbrica pre  $M = 44.63$ , Rúbrica post  $M = 41.93$ , Dif.: -2.7,  $p < .01$ ), mientras que el grupo de guiones no varió significativamente. En la medida post, ambos grupos fueron iguales en su autorregulación. Parece, pues, que *las rúbricas tienen un efecto negativo sobre la autorregulación del aprendizaje que no tienen los guiones*. Sin embargo, como no había grupo control, no se puede concluir de estos datos que el uso de guiones beneficie la autorregulación, solo que evitó el descenso producido por el uso de rúbricas.

FIGURA 1

*Escala de autorregulación de rendimiento/evitación.* Es importante recordar que las puntuaciones altas en esta escala indican una falta de autorregulación adecuada cuando el estudiante se enfrenta a dificultades. La interacción entre el tipo de tratamiento (rúbricas o guiones) y el momento de medida (pre-post) fue significativa,  $F(1, 83) = 8.86, p < .001, \eta^2 = .096$ , Rúbrica  $M = 43.8$ , Guión  $M = 45.92$ . Como se puede ver en la Figura 2, ambos grupos mostraron el mismo nivel antes de la intervención. Sin embargo, la autorregulación de rendimiento/evitación descendió después de la intervención en el grupo de rúbricas (Rúbrica pre  $M = 45.19$ , Rúbrica post  $M = 42.41$  Dif.:  $-2.78, p = .06$ ), mientras que el uso de este tipo de mensajes disminuyó en el grupo de guiones (Guión pre  $M = 44.23$ , Guión post  $M = 47.61$ , Dif.:  $3.38, p < .05$ ), siendo la diferencia entre grupos significativa en la medida post (Rúbrica  $M = 42.41$ , Guión  $M = 47.61$ , Dif.:  $5.2, p < .05$ ). En resumen, ambos grupos comenzaron con el mismo nivel de autorregulación de rendimiento/evitación, pero después de la intervención los estudiantes del grupo de guiones experimentaron un incremento, mientras que los del grupo de rúbricas experimentaron un descenso, lo que implica que *los guiones tienen un efecto negativo sobre la autorregulación de rendimiento/evitación, mientras que las rúbricas no.*

#### FIGURA 2

*Escala específica de autorregulación para los resúmenes.* La interacción entre el tipo de tratamiento y el momento de medida quedó ligeramente por debajo de los límites estándar de significación,  $F(1, 83) = 3.644, p = .06, \eta^2 = .042$ , Rúbrica  $M = 20.36$ , Guión  $M = 20.11$ . En la Figura 3 se puede ver que, en la medida pre, el grupo de rúbricas mostró un nivel de autorregulación específica para el resumen más alto, mientras que en la medida post el grupo de guiones mostró un nivel no significativamente más alto. Es decir, el significado de este efecto casi significativo es que los guiones “tienden” a beneficiar la autorregulación del aprendizaje, aumentando la cantidad de mensajes autodirigidos relacionados con la supervisión del proceso de la tarea, mientras que las rúbricas tienden a tener el efecto opuesto.

#### FIGURA 3

*Escala específica de autorregulación para los mapas conceptuales.* La interacción entre el tipo de tratamiento y el momento de medida fue significativa,  $F(1, 83) = 4.70, p < .05, \eta^2 = .054$ , Rúbrica  $M = 25.01$ , Guión  $M = 24.88$ . Como puede verse en la Figura 4, mientras que el grupo de rúbricas muestra un descenso significativo en el nivel de autorregulación específica (Rúbrica pre  $M = 26.15$ , Rúbrica post  $M = 23.88$ , Dif.:  $2.27; p < .001$ ), el grupo de guiones no muestra un cambio significativo. Este efecto quiere decir que las rúbricas “tienden” a obstaculizar la autorregulación específica para los mapas conceptuales, disminuyendo la cantidad de mensajes autodirigidos relacionados con la supervisión del proceso de la tarea, mientras que los guiones no tienen este efecto.

#### FIGURA 4

*Efectos de la intervención sobre la autoeficacia*

Respecto a la autoeficacia, la intervención no tuvo efectos significativos, ni para los resúmenes,  $F(1, 83) = 1.18$ ;  $p = .280$ ,  $\eta^2 = .014$ , Rúbrica  $M = 22.39$ , Guión  $M = 21.59$ , ni para los mapas conceptuales,  $F(1, 83) = 0.001$ ;  $p = .971$ ,  $\eta^2 = .000$ , Rúbrica  $M = 16.73$ , Guión  $M = 16.70$ .

*Efectos de la intervención sobre la percepción de los estudiantes acerca de la utilidad de los instrumentos*

La ‘ayuda percibida’ se midió de dos maneras. En primer lugar, durante la ‘sesión de retroalimentación’ se preguntó a los estudiantes sobre su percepción acerca de los siguientes cuatro aspectos, en una escala del 1 al 5.

- *Dificultad para realizar el sumario.* Sin efecto significativo,  $F(1, 38) = .482$ ;  $p = .492$ ,  $\eta^2 = .013$ .
- *Dificultad para realizar el mapa conceptual.* Sin efecto significativo,  $F(1, 38) = 1.16$ ;  $p = .287$ ,  $\eta^2 = .030$ .
- *Ayuda percibida del instrumento para realizar el sumario.* Efecto significativo a favor de la rúbrica,  $F(1, 22) = 6.32$ ;  $p < .05$ ,  $\eta^2 = .231$ .
- *Ayuda percibida del instrumento para realizar el mapa conceptual.* Efecto significativo a favor de la rúbrica,  $F(1,22) = 4.41$ ;  $p < .05$ ,  $\eta^2 = .167$ .

En segundo lugar, en el último cuestionario se pidió a los estudiantes que evaluaran cuánto les había ayudado su instrumento de autoevaluación (rúbrica o guión) en la realización de las siguientes seis actividades autorreguladoras:

- *Al planificar la tarea.* No hubo efecto significativo,  $F(1, 83) = 3.064$ ;  $p = .084$ ;  $\eta^2 = .036$ .
- *Durante la tarea (ejecución).* La diferencia fue significativa, siendo la rúbrica mejor percibida,  $F(1, 83) = 12.3$ ;  $p < .001$ ;  $\eta^2 = .129$ .
- *Al finalizar la tarea.* La diferencia fue significativa, siendo la rúbrica mejor percibida,  $F(1, 83) = 14.08$ ;  $p < .001$ ;  $\eta^2 = .145$ .
- *Para autoevaluarse.* La diferencia fue significativa, siendo la rúbrica mejor percibida,  $F(1, 83) = 3.815$ ;  $p = .054$ ;  $\eta^2 = .044$ .
- *Para permanecer tranquilo.* No hubo efecto significativo,  $F(1, 83) = .934$ ;  $p = .337$ .
- *Para focalizarse sobre el aprendizaje.* La diferencia fue significativa, siendo la rúbrica mejor percibida,  $F(1, 83) = 3.79$ ;  $p = .055$ ;  $\eta^2 = .044$ .
- *Para focalizarse sobre la calificación.* No hubo efecto significativo,  $F(1, 83) = .254$ ;  $p = .615$ .
- *Para corregir errores.* La diferencia fue significativa, siendo la rúbrica mejor percibida,  $F(1, 83) = 7.31$ ;  $p < .01$ ;  $\eta^2 = .081$ .

## Discusión

El objetivo principal de este estudio era comparar los efectos de dos herramientas de autoevaluación (rúbricas y guiones) sobre el aprendizaje, la autorregulación, el incremento de la autoeficacia y la activación de diferentes metas de estudiantes universitarios en un contexto natural.

### *Aprendizaje*

El uso de rúbricas fue mejor para la creación de sumarios y mapas conceptuales que el uso de guiones. Las rúbricas mejoraron el aprendizaje en el presente estudio, en consonancia con investigaciones previas (Andrade et al., 2008; Sadler & Good, 2006). Sin embargo, ¿por qué no mejoró el aprendizaje de forma similar con el uso de guiones, como otras investigaciones (p. ej. Berthold et al., 2007, Kramarski & Michalsky, 2009, 2010) han encontrado anteriormente? Puede haber varias explicaciones.

La primera tiene que ver con la posibilidad de que los efectos de las herramientas de autoevaluación estén mediados por otras variables, como la naturaleza y la estructura de la tarea. En un estudio previo que comparaba el efecto de ambas herramientas sobre el rendimiento durante el trabajo en una tarea altamente estructurada, el diseño de material multimedia, y trabajando con grupos pequeños (Panadero, Alonso-Tapia, & Reche, 2012), ni la rúbrica ni el guión produjeron un mejor rendimiento en comparación con un grupo control. Sin embargo, en el presente estudio las actividades de aprendizaje estaban menos estructuradas, y el tamaño de la muestra de los grupos era mayor, lo que puede haber tenido influencia en el efecto de la intervención.

La segunda explicación sobre por qué las rúbricas parecen una mejor estrategia para mejorar el rendimiento que los guiones puede encontrarse en el hecho de que las rúbricas aportan información sobre los avances de los estudiantes, y sobre cómo estos se relacionan con la puntuación. Esto puede producir una suerte de *instrucción*, porque la rúbrica muestra el producto final que debe alcanzarse, mientras que los estudiantes que usan el guión se enfrentan al proceso experto sin tener información sobre cómo debe ser el producto final que deben lograr. Como se ha sugerido con anterioridad (Kyndt, Dochy, Struyven, & Cascallar, 2010), la familiaridad con los criterios de calidad de la tarea en situaciones muy complejas lleva a un nivel de procesamiento más profundo. Consecuentemente, las rúbricas pueden haber sido mejores herramientas para llegar a una versión aceptable del producto final.

Sin embargo, es muy importante apuntar que los mapas conceptuales y los resúmenes se evaluaron y puntuaron utilizando la misma rúbrica que recibió el estudiante. Que los evaluadores utilizaran un guión como herramienta de puntuación no parecía tan factible como que utilizaran rúbricas, porque el guión pone el énfasis sobre la evaluación del proceso, al que no tienen acceso los profesores, ya que evalúan el producto final. Además, en estudios previos se ha comprobado la utilidad de las rúbricas como una herramienta de puntuación que eleva la fiabilidad de los evaluadores (Jonsson & Svingby, 2007). Por tanto, aunque la rúbrica era la

herramienta adecuada para calificar el portafolio de los estudiantes, darles la rúbrica puede haber actuado como un entrenamiento que introdujo un sesgo en la evaluación de los efectos sobre el aprendizaje. Tal vez una forma más apropiada de medir los efectos sobre el aprendizaje de las herramientas de autoevaluación hubiera sido evaluar a los estudiantes sin las herramientas en otra tarea diferente para la que no habían sido entrenados, de forma que tuvieran que transferir sus conocimientos.

### *Autorregulación*

El uso de las rúbricas parece ser más perjudicial que el de los guiones para la *autorregulación del aprendizaje*, ya que este tipo de autorregulación positiva disminuyó después de la intervención. En este sentido, los guiones son preferibles a las rúbricas porque se ha encontrado en investigaciones anteriores que incrementan significativamente la autorregulación, permitiendo a los estudiantes una mejor comprensión del procedimiento seguido y del producto obtenido (Kramarski & Dudai, 2009; Panadero, Alonso-Tapia, & Huertas, 2012).

El aumento en el nivel de *autorregulación de rendimiento/evitación* en la condición guión es un resultado inesperado que contrasta con las investigaciones previas. En el estudio que llevaron a cabo Panadero, Alonso-Tapia y Huertas (2012) sobre una muestra de educación secundaria no se encontró un efecto significativo, mientras que en estudiantes universitarios (Panadero, Alonso-Tapia, & Reche, 2012) el uso de rúbricas disminuyó la autorregulación de rendimiento/evitación. Dado que los guiones no tienen una faceta de puntuación, los estudiantes pueden haberse sentido inseguros respecto a su propio rendimiento, lo que podría haber disparado sus pensamientos y emociones negativas (Panadero, 2011). Por tanto, las rúbricas parecen ser una mejor herramienta que los guiones para disminuir la autorregulación de rendimiento/evitación en el caso de los estudiantes universitarios. Esto se puede deber a que los estudiantes universitarios se preocupan más por sus calificaciones que los estudiantes de secundaria. Así, la faceta de puntuación de la rúbrica permite a los estudiantes comprobar la calidad de su rendimiento (comparándolo con la rúbrica), experimentando menos pensamientos negativos (Andrade & Valtcheva, 2009). En resumen, si se toman conjuntamente los resultados de aprendizaje y autorregulación de rendimiento/evitación parece que las herramientas de autoevaluación tienden a fomentar o dificultar distintas características de la autorregulación. Por tanto, se puede argumentar que el éxito de una intervención sobre la autorregulación mediante la autoevaluación puede depender del tipo de tarea, medida y contexto (Panadero, Alonso-Tapia, & Reche, 2012). Asimismo, las rúbricas pueden resultar más adecuadas que los guiones para tareas en las que es importante identificar distintos niveles de rendimiento (Reitmeier & Vrchota, 2009), cuando los estudiantes no han realizado nunca la tarea, de manera que puedan llegar a una mejor comprensión de cómo debería ser el resultado final (p. ej. cursos de introducción), o cuando la tarea es de un alto nivel de especificidad y los pasos a seguir no permiten flexibilidad en el rendimiento (no se

requiere reflexión sobre el procedimiento). Los guiones puede ser más adecuados cuando los estudiantes están familiarizados con la tarea y, por tanto, puede reflexionar sobre su contenido y el procedimiento necesario para llevarla a cabo (p. ej. cursos más avanzados), o incluso el producto final es modificable, por lo que los estudiantes deben argumentar claramente qué procedimiento deben seguir (p. ej. un artículo de reflexión). Por lo tanto, los profesores deberían escoger la herramienta más útil para lograr las metas deseadas.

### *Autoeficacia*

Las herramientas de autoevaluación no tuvieron un efecto significativo sobre la autoeficacia, un resultado que concuerda con las investigaciones previas en el nivel de educación superior (Panadero, Alonso-Tapia, & Reche, 2012). Sin embargo, dos estudios con estudiantes de educación secundaria han encontrado que las rúbricas aumentan la autoeficacia en interacción con el sexo (Andrade, Wang, Du, & Akawi, 2009), y que las rúbricas aumentan la autoeficacia en interacción con la retroalimentación (Panadero, Alonso-Tapia, & Huertas, 2012). Por tanto parece que las rúbricas, en interacción con otras variables, aumentan la autoeficacia de los estudiantes de secundaria, mientras que en la educación superior este efecto no se ha demostrado.

Una primera explicación posible es que puede ser necesario recibir retroalimentación para que mejore la autoeficacia, ya que de esa forma los estudiantes reflexionan sobre su capacidad para realizar una tarea cuando se enfrentan a situaciones en las que pueden encontrar resultados inesperados (van Dinther, Dochy, & Segers, 2010). Así, cuanta más retroalimentación sobre su rendimiento reciban los estudiantes, mayores serán las posibilidades de que reflexionen sobre su capacidad, reflexiones que pueden ejercer influencia sobre su autoeficacia. En el presente estudio los estudiantes recibieron retroalimentación tres veces a lo largo de un curso semestral; tal vez esta intervención no haya tenido la intensidad suficiente como para hacerles reflexionar sobre su capacidad para realizar la tarea.

Una segunda explicación puede ser que la autoeficacia de los estudiantes universitarios sea más difícilmente modificable que la de los estudiantes de educación secundaria (ya que la experiencia educativa de los primeros es mayor, y su concepto sobre su autoeficacia es más comprensivo), y por tanto depende menos de su rendimiento en una asignatura concreta. Por tanto, tal vez la intervención deba durar más para producir un impacto en la autoeficacia de los estudiantes universitarios.

Una tercera explicación es que, como este estudio no contaba con un grupo control, no había datos que permitieran conocer si la intervención había tenido éxito en comparación con la línea base. Incluso aunque en Panadero, Alonso-Tapia y Reche (2012) se incluyó un grupo control y no se encontraron diferencias significativas entre rúbrica, guión y grupo control, en dicho estudio los estudiantes no recibieron retroalimentación en todo el semestre. Por tanto, ni el presente estudio (por la falta de un grupo control), ni el mencionado de Panadero et al. (por

la falta de retroalimentación, de crucial importancia en las intervenciones sobre la autoeficacia; van Dinther et al., 2010) pueden considerarse concluyentes al respecto de la influencia de las herramientas de autoevaluación sobre la autoeficacia de los estudiantes universitarios.

*Percepción subjetiva de la utilidad del instrumento y de la activación de metas*

En los datos de la ‘sesión de retroalimentación’ no se encuentra un efecto significativo del instrumento de autoevaluación sobre la percepción de los estudiantes acerca de la dificultad de escribir resúmenes y de diseñar mapas conceptuales. En realidad, los bajos niveles de ese efecto denotan una débil conexión entre el uso de rúbricas y guiones y cuán difíciles perciben las tareas los estudiantes, por lo que sus percepciones pueden depender de otras variables (p. ej. experiencia previa). Sin embargo, había diferencias significativas en la percepción de los estudiantes acerca de la utilidad de los instrumentos de autoevaluación. Los estudiantes que usaron las rúbricas manifestaron que su instrumento les ayudó a escribir los resúmenes y a diseñar los mapas conceptuales más que los estudiantes que usaron guiones. Por tanto, se percibió la rúbrica como un mejor instrumento, lo que parece lógico, ya que el uso de rúbricas parece tener un efecto de “entrenamiento”.

Cuando se midió mediante cuestionarios la percepción acerca de la utilidad de los instrumentos, se encontraron diferencias significativas en cinco de los ocho factores medidos a favor de las rúbricas (un sexto ítem cayó ligeramente por debajo de los límites estándar de significación). El uso de rúbricas se percibió como más útil en las tres fases de la tarea (planificación, ejecución y autorreflexión), y también para autoevaluarse, focalizarse en el aprendizaje, corregir errores y (casi significativo) permanecer tranquilo. A la luz de estos resultados, parece que los estudiantes que usaron las rúbricas percibieron la herramienta utilizada como más útil que los estudiantes que usaron guiones, y también que el uso de rúbricas fomentó más la elección de metas de aprendizaje que el uso de guiones. Estos efectos pueden deberse a la seguridad que aporta a los estudiantes el sistema de calificación basado en rúbricas, o al hecho de que las rúbricas son más fáciles de usar, porque presentan los criterios de evaluación, que los guiones, que presentan preguntas de reflexión.

En cuanto a la utilidad subjetiva para la activación de metas, los estudiantes parecen considerar que las rúbricas son mejores que los guiones porque, en contra de nuestras expectativas, no focalizan la atención de los estudiantes sobre la calificación más que el uso de guiones. Sin embargo, estos resultados también deben tomarse con cautela, ya que están basados en una medida de dos ítems realizada mediante autoinforme.

Resumiendo, nuestros resultados indagaron sobre los efectos de estos instrumentos de autoevaluación. Por una parte, respecto a los efectos objetivos, las rúbricas parecen dificultar más la autorregulación del aprendizaje, y los guiones “tienden” más a favorecerla, en línea con las investigaciones previas (Alonso-Tapia & Panadero, 2010; Bannert, 2009). Las

rúbricas parecen favorecer el rendimiento, aunque este efecto puede deberse al entrenamiento. Sugerimos pues el uso de actividades de transferencia para evaluar los efectos de ambos instrumentos. Por otra parte, respecto a los efectos subjetivos, los estudiantes parecen percibir las rúbricas como más útiles, y por tanto se encuentran más cómodos usándolas, como otros estudios han apuntado anteriormente (Andrade & Du, 2005; Reynolds-Keefer, 2010).

### *Limitaciones*

El presente estudio tiene varias limitaciones. En primer lugar, la autorregulación solo se midió mediante cuestionarios. Aunque se emplearon dos tipos de cuestionario, siguiendo las recomendaciones previas (Boekaerts & Corno, 2005), sería preferible usar también *protocolos en voz alta* (Alonso-Tapia & Panadero, 2010), por ejemplo. Sin embargo, a no ser que los estudiantes trabajaran con computadoras, sería difícil realizar las diferentes medidas en contextos naturales, y la investigación previa avala el método seguido en este estudio (Samuelstuen & Bråten, 2007).

En segundo lugar, el tamaño del efecto de las condiciones experimentales era mediano y, incluso aunque se controlaron las variables para comprobar la comparabilidad de los grupos, hay que tener en cuenta el efecto de anidamiento.

En tercer lugar, aunque el diseño inicial incluía un grupo de control, algunos problemas durante la intervención invalidaron esos datos. Por ello, no se pudo comparar la ganancia provocada por los instrumentos de autoevaluación respecto a un grupo que no los utilizara, algo especialmente importante en el caso de la autoeficacia. Por tanto, en el futuro se debería continuar explorando el impacto del uso de herramientas de autoevaluación en estudiantes universitarios.

En cuarto lugar, los resultados de las investigaciones disponibles sobre este tema no son convergentes, lo que indica que es necesario un estudio sistemático de los efectos de las variables moderadoras en las intervenciones que utilizan rúbricas y guiones (tipo de tarea, tipo de estudiantes, etc.).

Para concluir, la comparación de los efectos relativos de rúbricas y guiones en estudiantes universitarios ha mostrado que ambas tienen ventajas y limitaciones. Las rúbricas parecen disminuir los mensajes negativos autodirigidos y aumentar el aprendizaje, en consonancia con el trabajo de Halonen et al. (2003) sobre el uso de rúbricas en asignaturas de psicología, aunque tienden a dificultar la autorregulación del aprendizaje, mientras que los guiones parecen favorecer niveles más altos de autorregulación del aprendizaje. Es por tanto crucial que los profesores reflexionen sobre cuál de las dos herramientas es más indicada para las metas de la tarea.



Tabla I

*Medias y desviaciones estándar*

Herramienta	Aprendizaje		Autorregulación						Autoeficacia				
	Res.	MC	AR		AR		AR Esp. MC		AR Esp. Res.		Res.	MC	
			Aprendizaje	Rend./Ev.	Pre	Post	Pre	Post	Pre	Post			
Rúbrica	<i>M</i>	3.11	2.75	44.63	41.93	45.19	42.41	26.15	23.88	20.83	19.90	22.39	16.73
	<i>DT</i>	.35	.45	1.10	1.10	1.82	1.72	.78	.96	4.45	6.07	3.53	3.25
Guión	<i>M</i>	2.86	2.29	40.89	41.54	44.23	47.61	24.77	25.00	19.73	20.50	21.59	16.70
	<i>DT</i>	.34	.44	1.06	1.06	1.76	1.66	.75	.93	3.98	4.30	3.24	3.51

Nota. Res. = Resumen; MC = Mapa Conceptual; AR. = Autorregulación; AR Rend./Ev. = Autorregulación Rendimiento/Evitación; AR Esp. MC = Autorregulación Específica para Mapas Conceptuales; AR Esp. Res. = Autorregulación Específica para Resúmenes.

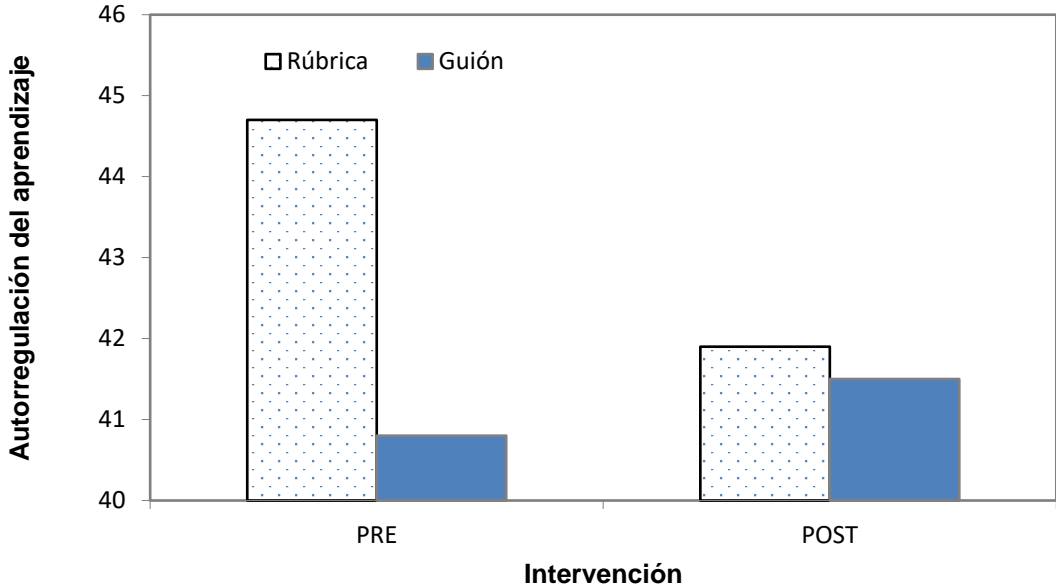
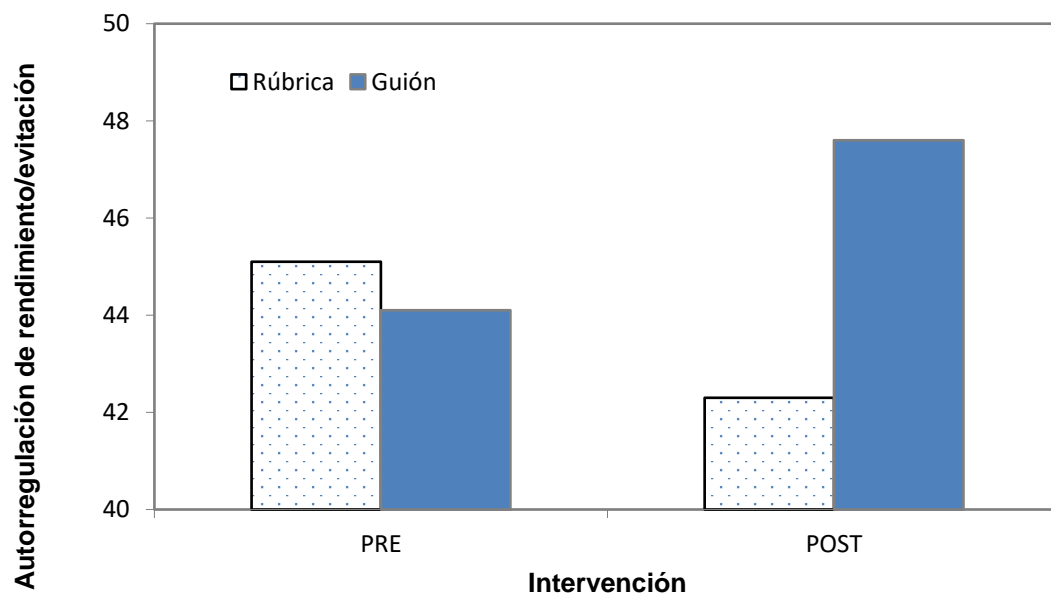


Figura 1  
Efecto de la interacción entre el tipo de herramienta de autoevaluación y el momento de medida sobre la autorregulación del aprendizaje

---

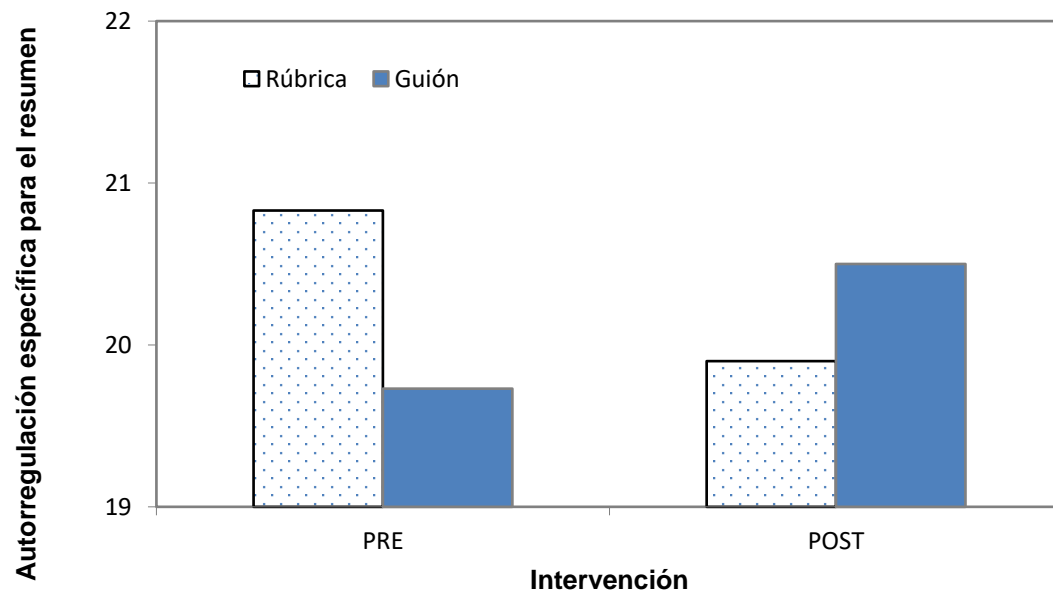


---

Figura 2

Efecto de la interacción entre el tipo de herramienta de autoevaluación y el momento de medida sobre la autorregulación de rendimiento/evitación

---

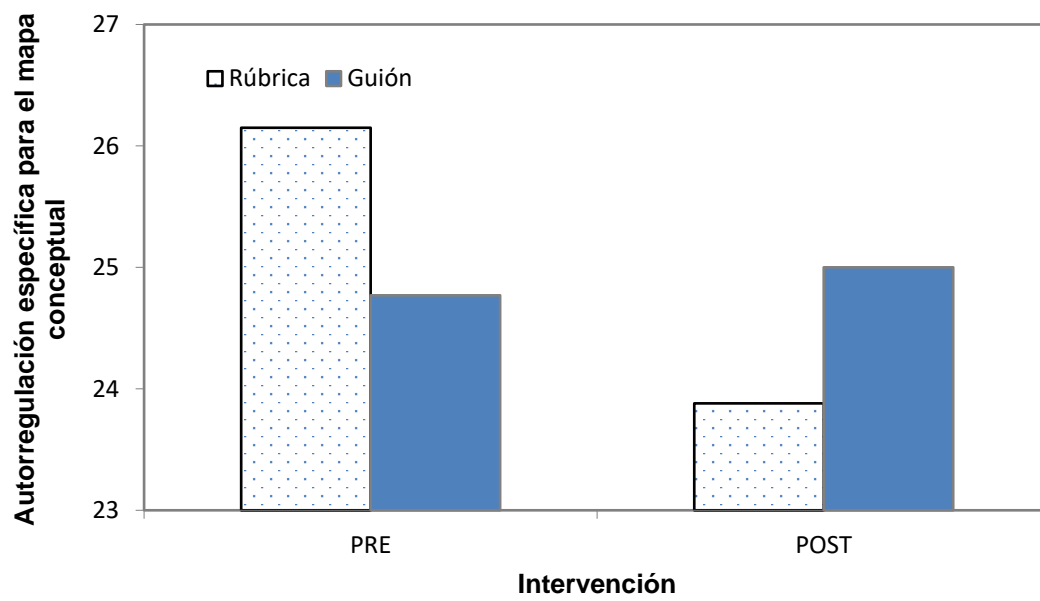


---

Figura 3

Efecto de la interacción entre el tipo de herramienta de autoevaluación y el momento de medida sobre la autorregulación específica del resumen

---



---

Figura 4

Efecto de la interacción entre el tipo de herramienta de autoevaluación y el momento de medida sobre la autorregulación específica del mapa conceptual

---

## Apéndice A

### Rúbrica: cómo escribir un resumen

Puntuación Criterios	4	3	2	1
<b>Se incluye la idea principal del texto original</b>	El resumen comienza con la idea principal del texto original	Contiene la idea principal pero no al principio	Lo que se propone en el resumen como la idea principal en realidad no lo es	El resumen no contiene la idea principal
<b>Se incluyen las ideas secundarias que apoyan a la idea principal</b>	El resto de ideas importantes están reflejadas en el resumen con pocas palabras	El resto de ideas importantes están en el resumen pero se podrían haber reflejado con menos palabras	Algunas de las ideas importantes están en el resumen, pero no todas, y se podrían haber reflejado con menos palabras	El resumen contiene elementos superficiales y tiene demasiadas palabras

(Continúa).

## Apéndice B

### Guión de autoevaluación: cómo escribir un resumen

- ¿Transmite mi resumen la idea principal del texto? ¿Está esta idea al principio del resumen?
- Asimismo, ¿están las ideas importantes también en mi resumen?
- ¿He escogido las ideas principales del texto y las he reflejado de forma explícita en mi resumen?
- ¿He pensado en mi intención al escribir el resumen? ¿Cuál es mi meta?

(Continúa).

## Notas

<sup>1</sup>La orientación a metas de los participantes se midió mediante el cuestionario LEMEX (Alonso-Tapia, Huertas & Ruiz, 2010). Los resultados mostraron que ambos grupos estaban equilibrados en cuanto a la orientación a metas de sus participantes.