How and When Presenting a Concept Map for Learning and an Accurate Self-evaluation?

A. Maillard, L. Motak, J. C. Sakdavong, C. Dupeyrat and N. Huet
CLLE-LTC CNRS UMR5263, Université Toulouse 2 Le Mirail, 5 allée Antonio Machado, Toulouse, France

Keywords: Self-regulated Learning, Self-evaluation, Concept Map, Cognitive Load.

Abstract: Self-evaluation is not an easy step for learners even if it is a decisive step in self-regulated learning. The goal of our study was to test concept maps effect on learning performance and self-evaluation accuracy. 136 students were assigned over five experimental groups in which the format used (consultation/construction) and the moment of use (simultaneous of the learning task vs. after the learning task) of concept map varied. Cognitive load was also measured in order to explain differences in performance and self-evaluation. Results suggested that participants in the consultation conditions have a more accurate self-evaluation and better performance than participants in the construction condition. More studies are required to identify more precisely what factors influence the efficiency of use conceptual map.

1 INTRODUCTION

Self-evaluation is a process which consists in detecting a difference between a specific learning goal and the current state of knowledge (Nelson & Narens, 1990). The accuracy of this estimation is essential because this is what enables learners to adopt the appropriate learning strategies and behaviours (Gama, 2004). However, numerous studies show that students experience difficulties to perform a correct self-evaluation (see Dunlosky and Nelson, 1994; Dunlosky & Nelson, 1992; Koriat, 1997).

Several tools have been developed in order to improve self-evaluation accuracy (Dunlosky and Rawson, 2012, Kornell and Son, 2009, Chi and al., 1989). However, their uses are specific (for definitions or for a well-guided condition) or their efficacy limited. More recently, Redford et al. (2012) have improved self-evaluation of comprehension by the aid of concept maps. Concept maps are schemas which display the relations between different concepts (see Nesbit and Adesope, 2006; Novak & Gowin, 1984).

According to Redford and his colleagues, learners who organize information themselves would exhibit a more accurate self-evaluation as compared to learners who merely consult an already defined map. Results of their experiments confirm this hypothesis by highlighting that self-evaluation is more accurate in the construction condition as compared to the rereading condition (Experiment 1) and to the consultation condition (Experiment 2). Thus, for the authors, presenting the same information twice does not enhance self-evaluation accuracy.

The Cognitive Load Theory (see Sweller, 1988) casts a new look at this lack of effect of concept map consultation. In particular, the "split attention effect" can offer explanations to this result. This "split attention effect" appears when people have to process different information for which the integration needs to be mentally performed in order to infer sense from the presented material (Tricot, 1998).

In parallel, in the Information Processing Theory perspective, Winne (2001) shows that self-regulated processes including self-evaluation, rely on cognitive resources. Thus, map consultation, learning and self-evaluation have a cognitive cost and the concurrent fulfilment of these three activities may overload learners' working memory, resulting in a bad integration of information and in difficulties to perform self-evaluation. On the contrary, the map construction condition involves only one source of information. Cognitive load of participants is consequently less important, which allows them to use all their resources to process, to integrate

Maillard A., Motak L., C. Sakdavong J., Dupeyrat C. and Huet N..
How and When Presenting a Concept Map for Learning and an Accurate Self-evaluation?.
DOI: 10.5220/0004414401880193
In *Proceedings of the 5th International Conference on Computer Supported Education* (CSEDU-2013), pages 188-193
ISBN: 978-989-8565-53-2

information but also to evaluate this degree of integration more easily.

Taken together, these studies (Redford et al, 2012; Tricot, 1998) suggest implicitly the issue of temporality in the presentation of concept maps in their effectiveness vis-à-vis self-evaluation and learning. Therefore, in this perspective of temporality, it is relevant to examine whether the construction and the consultation cannot be more efficient for learning and self-evaluation whether it is done after rather than during the learning task.

To answer this question, we based our experiment on Redford et al's (2012) study in which students had to construct or to consult a concept map during a learning task. We added two new conditions. The first condition was a consultation of a pre-conceived map after seeing a learning content (and not during). The second one was a construction condition after seeing a learning content. We added also a control condition in which no concept map was used at all. The learning task was word-processing and especially creation of styles.

In reference to the Cognitive Load Theory (Sweller, 1988), we predicted that learning would be better and self-evaluation more accurate for participants who use concept maps after the learning task as compared to participants who use concept maps during the learning task. In addition, we predicted that participants who construct the concept map would perform better and will have a more accurate self-evaluation than participants who merely consult the concept map.

In other words, organizing information oneself instead of just reading it would improve participants' learning. Indeed, concept map elaboration is a deep cognitive strategy (e.g., Weinstein and Mayer, 1986). Numerous authors (e.g., Pintrich and De Groot, 1990; Weinstein and Mayer, 1986) have highlighted that the use of deep strategies lead to higher performance and to a more accurate self-evaluation (e.g., Cassidy, 2006).

Finally, we expected an interaction effect between both, the modality and the time of map presentation. Thus, we predicted that the "concurrent consultation" condition would lead to the worst performance and to the less accurate self-evaluation whereas the "construction after the learning task" would lead to the best performance and to the more accurate self-evaluation.

This study was funded by the french ANR CONTINT program (ANR10-CORD-011-01).

2 METHOD

2.1 Participants

136 students (80 females, average age = 18.5, SD = 1.2) in first year of "Industries and Administration Management". No differences in the initial level of knowledge on word-processing were detected among the experimental groups ($p > .05$).

2.2 Familiarization Phase with Concept Maps

The familiarization phase was intended to introduce students to the "concept maps" designer tool that they would handle during the learning phase. Students were presented a screencast (screen + audio commentary) with the following information: the purpose and terms of the experiment, the definition of a concept map and its purpose, an example of concept map and how it should be read, the function and value of a concept map and a demonstration of the use of the concept maps designer tool. Then they had to do two training exercises including a feedback with examples of conceptual maps of the expected kind and of the inadequate kind. The tool used to build concept map was a simplified version of CoGui (http://www.lirmm.fr/cogui)

2.3 Learning Phase and Groups

The learning phase consisted of three videos. These videos all related to word processing and more specifically on how to create "styles" (with a word processor) and apply them to the document. The videos had an increasing level of difficulty. The software device was constructed so that learners cannot avoid any part of the videos or view them in a different order than the one proposed. During this learning phase, participants were randomly assigned in one of the five groups depicted as follows in the learning phase. The first group was instructed to watch videos and simultaneously build a concept map (N = 27). The second group was instructed to watch the videos and then build a concept map (N = 28). The third group watched the video and simultaneously consulted an expert conceptual map data (N = 27). The fourth group was viewing videos then consulted an expert concept map (N = 24). Finally, the fifth group (control group) was viewing videos simply without being presented concept maps (N = 30).

Three important instructions were given to each participant: the opportunity to "pause" the videos at any time, they were not allowed to watch video twice and to take notes. Participants were also informed that they will later undergo a test to assess their understanding: Test Phase and self-evaluation

2.3.1 Familiarization with Exercise Test

The exercise test itself was preceded by a familiarization period to enable to have a clear idea of what they would be asked to do.

The familiarization period (identical to the test) consisted of the three following steps:

- A performance prediction ("To what extent do you think you can correctly answer the following question?")
- Task instructions and execution ("You will be prompted to format the text without using styles")
- A post diction ("To what extent do you believe you have correctly answered the previous question?")

2.3.2 The Test

The exercise test consisted of the same three steps as the familiarization period. More specifically, during the task execution, participants were instructed to change the appearance of a text using styles. To do this, they first had to change the predefined styles according to specific instructions in 15 actions. One point per correct action was attributed. For measuring self-evaluation, participants answered on a scale ranging from 0 to 100. They estimated before and after the exercise how they thought they could or had correctly answered the question.

2.3.3 Self-evaluation Biases

Biases were measured by the difference between the actual performance of learners and their performance evaluation. For this, we used a linear regression analysis (Bouffard et al., 2006). First, we transformed the performance scores and measures of self-evaluation in Z scores. Then, we performed a regression of performance on self-assessment for standardized residuals. Standardized residuals equal to or greater than 1 indicated over-evaluation and standardized residuals equal or less than -1 indicated under-evaluation.

2.4 Measure of Cognitive Load

To measure cognitive load, we used the distinction done by Amadieu, Mariné and Laimey (2011) who measure cognitive load in two different ways. Specifically, the authors distinguish general cognitive load and cognitive "overload". For example, to measure cognitive load associated with the general understanding of the videos, participants were asked to rate on a 9-point scale: "The cognitive effort to understand videos was:" very small (1), very important (9) (Paas, 1992). To measure cognitive "overload", participants asked to rate on a 9-point scale: "Indicate how much it was difficult for you to understand videos", very easy (1), very difficult (9). The same questions were asked to measure the cognitive load associated with consultation / construction of concept maps.

2.5 Procedure

The first two phases of the device were identical for all participants. The first phase measured the level of knowledge of participants before any learning; the second one was to familiarize participants with the concept maps. In the third phase, called learning, students were assigned randomly to one of five groups of our experimental manipulation, watched the learning videos and then did the familiarization and test exercises. At the end of the experiment, all participants responded to questions measuring cognitive load.

3 RESULTS

3.1 Performance

A one-way analysis of variance (ANOVA) was conducted with group as the independent variable and with performance as dependent variable. Results showed a significant effect of the group, $F(4, 131) = 2.72$, $p = .03$, $\eta^2 p = .08$. Post-hoc test (Tukey) detected a significant difference between the control group and the simultaneous construction group, $p = .04$, showing that the control group ($M = 10.6$) performed better than the simultaneous construction group ($M = 6.0$). The other groups did not significantly differ between them.

A two-way ANOVA excluding control group, was conducted in order to detect eventual interaction effect between modality and moment. Results detected an effect of modality on performance, $F(1, 102) = 5.54$, $p = .02$, $\eta^2 p = .05$. Participants in the consultation group ($M = 9.8$) performed better than participants in the construction group ($M = 7.0$). Neither moment effect nor interaction effect was found.

3.2 Measure of Self-evaluation

3.2.1 Prediction Biases

In order to detect biases differences between our five groups, a one-way ANOVA was run on standardized residuals that detected a main effect of the group, $F(4, 130) = 3.00$, $p < .05$, $\eta^2p = .08$. Post-hoc test (Tukey) detected a significant difference between the delayed construction group and the simultaneous construction ($p < .05$) showing that participants in the delayed construction group tended to overestimate their performance while participants in the simultaneous construction tended to underestimate their performance (Table 1). Moreover, participants in the control group tended to underestimate their performances while participants in the delayed construction group tended to overestimate their performance ($p < .05$).

A two-way Moment x Modality ANOVA detected an interaction, $F(1, 102) = 4.63$, $p < .05$, $\eta^2p = .04$. Whatever the moment, consultation conditions enabled an accurate self-evaluation in both the delayed and the simultaneous condition. Simple effect analysis showed that mean bias did not differ between both consultation conditions. However, participants tended more to over-evaluation in the delayed construction condition than to under-evaluation in the simultaneous construction condition, $t(53) = 2.39$, $p = .02$.

3.2.2 Post-diction Biases

One-way ANOVA run on standardized residuals of post diction detected a main effect of the group, $F(4, 130) = 2.23$, $p = .05$, $n^2p = .07$. Post-hoc test detected a marginal difference ($p = .06$) between the delayed construction group and the simultaneous consultation group, suggesting that participants in the delayed construction group tended to overestimate their performance while participants in the simultaneous consultation group tended to underestimate their performance (see Table 1).

Two-way ANOVA only detected an effect of the moment of presentation, $F(1, 102) = 4.17$, $p = .04$, $\eta^2p = .04$, showing that the delayed condition led to a higher evaluation while the simultaneous condition led to a lower evaluation. Interestingly, participants for whom self-evaluation was the more accurate were those in the delayed consultation condition. Neither modality nor interaction effect were found.

Table 1: Descriptive statistics means (and standard Deviation) of standardized residuals of self-evaluation bias.

Groups	N	Prediction	Post-diction
DConstr	28	.43 (1.05)	36 (.92)
SConstr	27	-.25 (1.08)	-.14 (.90)
DConsult	24	-.08 (.91)	.01(1.14)
SConsult	27	.07 (.97)	.33(1.14)
Control	30	-.18 (.87)	.09 (.81)

Note: D for Delayed; S for Simultaneous; Constr for Construction; Consult for Consultation.

3.3 Cognitive Load

3.3.1 Mental Load

A one way ANOVA with group as factor and mental load to videos understanding as dependent variable was computed. Results showed a significant group effect $F(1,131) = 6.97$; $p < .001$; $\eta^2p = .18$. Post-hoc test (Tukey) detected a significant difference between delayed construction group and all the other four groups. The analysis showed that the mental load to understand videos was lower for participants in the delayed construction group than for the other groups.

A Modality x Moment ANOVA performed on mental load to videos understanding showed an effect of the modality $F(1, 102) = 2.56$, $p < .05$, $\eta^2p = .11$, an effect of moment $F(1, 102) = 18.31$, $p < .001$, $\eta^2p = .15$ and an interaction effect, $F(1, 102) = 5.97$, $p < .05$, $\eta^2p = .055$. Simple effect analysis showed that the mental load to understand videos was less important for participants in the delayed construction group ($M = 2.86$) than for the participants in the delayed consultation group ($M = 4.46$), $t(50) = -2.61$, $p < .01$. Moreover, the mental load was lower for the participants in the delayed construction group ($M = 2.86$) than for the participants in the simultaneous construction group ($M = 5.52$), $t(53) = -5.29$, $p < .001$.

Same analysis was conducted with the mental load felt to construct/consult the concept maps and detected an interaction effect between the two factors, $F(1, 102) = 7.35$, $p < .01$, $\eta^2p = .07$. Simple effect analysis showed that when the map was used in a delayed way, participants in the construction group felt a mental load significantly less important ($M = 2.18$) than participants in the consultation group ($M = 4.42$), $t(50) = -4.34$, $p < .001$. Analyses also showed, that participants felt a less important mental load when the construction was done in a delayed way ($M = 2.18$) than when the construction was done in a simultaneous way (M = 4.56), $t(53)$ = -5.99, p < .001.

Finally, for all participants, the cognitive load felt to understand videos was negatively correlated with the performance, r = - .21, p < .05, with the prediction, r = -.31, p < .001 and with the post diction, r = -.29, p < .001. The cognitive load felt to construct/consult maps was negatively correlated with the performance r = -.24, p = .014, and marginally with the post diction, r = -.18, p = .072.

3.3.2 Mental Overload

A one-way ANOVA with group as factor and mental overload to videos understanding as dependent variable was computed. Results showed a significant group effect $F(1, 131) = 7.38$; $p < .001$; η^2p =.18. Post-hoc test (Tukey) detected a significant difference between delayed construction group and all the other four groups. The analysis showed that the mental overload to understand videos was lower for participants in the delayed construction group than for the other groups.

A Modality x Moment ANOVA performed on mental overload felt to understand learning videos found significant effect of modality $F(1, 102)$ = 6.08, p <.05, η^2p =.06, an effect of moment $F(1, 102)$ = 8.20, p <.001, η^2p =.07 and an interaction effect, $F(1, 102)$ = 15.62, p < .001, η^2p = .13. Simple effect analysis showed that participants in the simultaneous consultation group felt a less important mental load (M = 4.85) than the participants in the simultaneous construction group (M = 5.89), $t(52)$ = 2.27, p = .03. Analysis also showed that participants in the delayed construction felt significantly a less important mental load (M = 4.96) than participants in the simultaneous construction (M = 5.89), $t(53)$ = -2.11, p = .04.

Finally a Modality x Moment ANOVA conducted on mental overload felt to construct/consult concept maps showed a significant interaction effect, $F(1, 102)$ = 6.67, p < .05, η^2p = .06. Simple effect analysis showed that participants in the simultaneous consultation group felt a significantly less important mental overload (M = 4.30) than participants in the simultaneous construction group (M = 5.89), $t(52)$ = 3.35, p < .001.

The difficulty felt to understand videos was negatively correlated with the performance, r = -.23, p = .01, with the prediction, r = -.41, p < .001, and with the post diction, r = -.20, p = .02. Similarly, cognitive overload felt to construct/consult maps was negatively correlated with the performance r = -.23, p = .02, with the post diction, r = -.27, p = .01, and was marginally correlated with the prediction, r = -.19, p = .06.

4 CONCLUSIONS

The aim of this study was to determine what the condition concept maps' presentation is the most relevant for learning and self-evaluation. Our first hypothesis assumed that concept maps construction would allow learners to have better learning performance and self-evaluation as compared to those who had to consult concept maps. Secondly we assumed that using a map after the learning task would be more efficient than using a map during the learning task. Finally we supposed that the "concurrent consultation" condition would lead to the worst performance and to the less accurate self-evaluation whereas the "construction after the learning task" would lead to the best performance and to the more accurate self-evaluation.

Results contradict the first and the third hypothesis and did not confirm the second one. Indeed, they showed that the consultation conditions were more efficient to improve learning and self-evaluation accuracy. More specifically, construction conditions decrease performance and self-evaluation accuracy as compared to the control group.

Regarding the cognitive load and overload, we thought that a higher cognitive load in the simultaneous and /or consultation conditions could explain lower performance and more inaccurate self-evaluation. This hypothesis is confirmed because participants found easier to understand videos when they had to construct a map rather than they had to consult it. Nonetheless, this is only right when map were used after the learning task. However, learners found easier to consult simultaneously a map than construct it simultaneously. These results confirm Stull and Mayer (2007) results. According to these authors, construction condition adds a task to the learning and cognitively overloads learners.

Moreover, these results are in contradiction with previous study (see Redford et al., 2012). They might be explained by the fact effective simultaneity between maps and videos was difficult to reach.

Finally, regarding to the cognitive load felt by learners, it could be relevant to check the concept map quality and numbers of created links. This point could explain why learners found easier to understand videos when the have to construct a map.

To conclude, we can recommend to not overloading learners as they are learning and this by limiting the tasks number during learning; present help system as concept map not during learning, and privilege the maps consultation rather than the maps construction.

REFERENCES

Amadieu, F., Marine, C. & Laimay, C., 2011. The attention-guiding effect and cognitive load in the compréhension of animations. *Computers in Human Behavior*, *27*, pp.36-40.

Bouffard, T., Vezeau, C., Chouinard, R. & Marcotte, G., 2006. L'illusion d'incompétence et les facteurs associés chez l'élève du primaire. *Revue Française de Pédagogie, 155*, pp.9-20.

Cassidy, S., 2006. Learning style and student self-assessment skill. *Education + Training, 48*, 170-177.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P. & Glaser, R., 1989. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, pp.145-182.

Dunlosky, J., & Nelson, T. O., 1992. Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* pp.374–380.

Dunlosky, J., & Nelson, T. O., 1994. Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33*(4), pp.545-565.

Dunlosky, J., & Rawson, K. A., 2012. Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), pp.271-280.

Gama, C.A., 2004. *Integrating Metacognition Instruction in Interactive Learning Environments.* Thèse de doctorat, Université de Sussex, Grande Bretagne.

Koriat, A., 1997. Monitoring one's own knowledge during study : A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, pp.349-370.

Kornell, N. & Son, L. K., 2009. Learners'choices and beliefs about self-testing. *Memory, 17*(5), pp.493-501.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G.H. Bower ed., 1995. *The Psychology of Learning and Motivation: Advances in Research and Theory*. Academic Press, San Diego. (Vol. 26,) pp. 125-169

Nesbit, J. C., Adesope, O. O., 2006. Learning with concept and knowledge maps: a meta-analysis. *Review of Educational Research*, *76* (3), pp.413-448.

Novak, J. D., & Gowin, D. B., 1984. *Learning how to learn.* Cambridge University Press, New York.

Paas, F., 1992. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, 84, pp.429–434.

Pintrich, P. R., & De Groot, E., 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, pp.33-40.

Redford,J. S., Thiede, K. W., Wiley, J. & Griffin, T.D., 2012. Concept mapping improves metacomprehensionaccuracy among 7th graders. *Learning and Instruction, 22*, pp.262-270.

Stull, A., Mayer, R. E., 2007, Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphics organizers. *Journal of Educational Psychology*, 99(4), pp.808-820.

Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* pp.257-285.

Tricot, A., 1998. Charge cognitive et apprentissage. Une présentation des travaux de John Sweller. *Revue de Psychologie de l'Education, 1*, pp.37-64.

Weinstein, C. E., & Mayer, R. E., 1986. The teaching of learning strategies. In M. Wittrock 2008. *Handbook of research on teaching*, Macmillan. New York.

Winne, P. H., 2001. Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk eds., 2001. *Self-regulated learning and academic achievement: Theoretical perspectives*. NJ: Erlbaum, Mahwah.