

# Effects of concurrent monitoring on cognitive load and performance as a function of task complexity

Citation for published version (APA):

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology*, 25(4), 584-587. <https://doi.org/10.1002/acp.1726>

**DOI:**

[10.1002/acp.1726](https://doi.org/10.1002/acp.1726)

**Document status and date:**

Published: 01/07/2011

**Document Version:**

Peer reviewed version

**Document license:**

CC BY-SA

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[pure-support@ou.nl](mailto:pure-support@ou.nl)

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 22 Mar. 2025

**Open Universiteit**  
[www.ou.nl](http://www.ou.nl)



Running head: COGNITIVE LOAD AND SELF-REGULATED LEARNING

Please refer to this article as van Gog, T., Kester, L., & Paas, F. (in press). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology*. doi: 10.1002/acp.1726

Effects of Concurrent Monitoring on Cognitive Load and Performance as a Function of Task  
Complexity

Tamara van Gog<sup>a,b</sup>, Liesbeth Kester<sup>b</sup>, and Fred Paas<sup>a,b</sup>

<sup>a</sup> Institute of Psychology, Erasmus University Rotterdam, The Netherlands

<sup>b</sup> Center for Learning Sciences and Technologies & Netherlands Laboratory for Lifelong  
Learning, Open University of The Netherlands

Author Note:

Correspondence concerning this manuscript should be addressed to Tamara van Gog, Center for Learning Sciences and Technologies, Open University of The Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands. T: +31 45 5762276; F: +31 45 5762907; E: [tamara.vangog@ou.nl](mailto:tamara.vangog@ou.nl)

*Acknowledgement:* This work was funded by the Netherlands Organization for Scientific Research (Veni Grants 451-08-003 and 451-07-007).

## Abstract

For self-regulated learning to be effective, students or trainees need to be able to accurately monitor their performance while they are working on a task, use the outcomes as input for self-assessment of that performance after completing the task, and select an appropriate new learning task in response to that assessment. From a cognitive load perspective, monitoring can be seen as a secondary task that may become hard to maintain and hamper performance on the primary task under high load conditions. The experiment presented here investigated the effects of concurrent performance monitoring on cognitive load and performance as a function of task complexity. Results showed that monitoring significantly decreased performance and tended to increase cognitive load on complex, but not on simple tasks. The findings are discussed in terms of theoretical consequences and instructional design for self-regulated learning.

## A Cognitive Load Perspective on Self-Regulated Learning: Concurrent Monitoring Negatively Affects Performance on Complex Tasks

A major aim of many contemporary educational and training programs is to foster students' or employees' self-regulated learning skills (Zimmerman, 1990). Research has shown that this aim is not likely to be achieved in a 'learning by doing' manner, that is, by providing learners with a high amount of control over their learning process (e.g., what tasks they work on or what information they study, in what order, and for how long) without any additional support. Rather, learners require some scaffolds such as metacognitive prompts or tutoring (Azevedo & Cromley, 2004; Van den Boom, Paas, & Van Merriënboer, 2007). Moreover, next to effects on the development of self-regulation skills, research has shown that whereas providing students with control over their learning process may have beneficial effects on their motivation or involvement, it often has no or detrimental effects on learning outcomes, especially for novice learners (see e.g., Azevedo, Moos, Greene, Winters, & Cromley, 2008; Corbalan, Kester, & Van Merriënboer, 2006; Niemic, Sikorski, & Walberg, 1996). These effects regarding learning outcomes are, however, not entirely surprising if we look at the cognitive demands imposed by self-regulated learning.

In theory, self-regulated learning provides the opportunity for personalized instruction that is adaptive to individual learner's needs. Such adaptive instruction has been shown to be more effective and efficient than fixed (one size fits all) or non-adaptive instruction (Camp, Paas, Rikers, & Van Merriënboer, 2001; Kalyuga & Sweller, 2004; Salden, Paas, & Van Merriënboer, 2006). In adaptive instruction, a learner performs a task, the system or teacher monitors and assesses performance on that task on different aspects (e.g., time taken, strategy used, number and types of errors), and then selects a new task that is appropriate in content and complexity level to help the learner improve certain aspects of his or her performance. For self-regulated learning to be truly adaptive and effective, learners *themselves* need to be

able to accurately monitor their performance while they are working on a task, use the outcomes as input for self-assessment of that performance after completing the task, and select an appropriate new learning task in response to that assessment<sup>1</sup>.

When learners are not able to monitor their performance while they are working on a task, they will not have a good memory representation of their performance process after completing it. Indeed, studies applying retrospective verbal protocols suggest learners are not very good at monitoring; they often show a rather poor recollection of the task performance process (see e.g., Van Gog, Paas, Van Merriënboer, & Witte, 2005). This is problematic, as without accurate monitoring, accurately self-assessing performance is difficult, and inaccurate self-assessment in turn may negatively affect task selection. As a consequence, learners may be devoting time and effort to learning tasks that are not at all adaptive to their needs. The reason why monitoring is so difficult, may be that most learning tasks are complex and impose a high cognitive load, especially for novices.

### *Monitoring and Cognitive Load*

Many learning tasks are complex, that is, they impose a high intrinsic cognitive load on working memory (Sweller, Van Merriënboer, & Paas, 1998). Intrinsic cognitive load depends on task complexity, because it is determined by the number of interacting information elements that have to be related, controlled, and kept active in working memory during task performance. It also depends on the expertise of the task performer: As a result of learning, elements are combined into cognitive schemata stored in long-term memory that can be retrieved and handled as a single element in working memory, thereby decreasing the intrinsic load of the task (Sweller et al., 1998).

The need to monitor performance during self-regulated learning can be seen as a secondary task. Under dual-task conditions, accurate performance of the secondary task or of both the primary and the secondary task becomes hard to maintain under high load conditions

(see e.g., Brünken, Plass, & Leutner, 2003). That is, under conditions of high task complexity, or high intrinsic load, little resources are available for processes that impose additional cognitive demands, such as concurrently monitoring performance. Learners can increase their effort to accommodate to the dual tasks demands to the extent that cognitive resources are still available. When the limit of cognitive capacity has been reached, however, learners need to divide their resources between performing the primary and the secondary task (cf. Kanfer & Ackerman, 1989), and as a result, monitoring, task performance, or both, may be hampered. Under conditions of low task complexity, or low intrinsic cognitive load, on the other hand, additional cognitive demands can be easily accommodated as ample resources are available, and the low complexity of the primary task may also make secondary task less complex. In sum, when tasks impose high intrinsic load, which most learning tasks do, the need to monitor performance may: a) lead to low quality monitoring (secondary task), and therefore, a poor recollection of performance on which to base self-assessment, and/or b) hamper performance of the learning task (primary task).

A pilot study was conducted to study the hypothesis that concurrent performance monitoring increases cognitive load and decreases performance on high intrinsic load but not on low intrinsic load tasks (reported in Van Gog & Paas, 2009). A mixed factorial design was used with task complexity as between-subjects factor and monitoring as within-subjects factor: participants first had to work on a puzzle without the instruction to monitor their performance, then with that instruction. This order was deliberately not counterbalanced, as the instruction to monitor on the first task might influence later task performance even when this instruction would not be given with the second task. Moreover, this order could not affect cognitive load to the advantage of the hypothesis, because intrinsic cognitive load tends to decline with increasing practice of a task (Sweller et al., 1998). Cognitive load was measured using the nine-point subjective mental effort rating scale developed by Paas (1992), which is

widely used in educational research (for an overview see Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Van Gog & Paas, 2008). Mental effort reflects the actual cognitive load, that is, the cognitive capacity that is allocated by the individual to accommodate the demands imposed by the task (Paas et al., 2003). The results suggested that monitoring indeed resulted in a trend towards higher cognitive load ( $p = .07$ ) and significantly lower performance on complex (9x9 Sudoku puzzles), but not on simple tasks (4x4 Sudoku puzzles). These results were in line with the hypothesis, however, they were far from unequivocal because of several reasons. First of all, this pilot had a relatively small number of participants ( $n = 19$  in the complex,  $n = 12$  in the simple tasks condition). Secondly, the order of the tasks was not counterbalanced. So the possibility that the findings concerning the complex tasks were due to potential differences in task difficulty between the tasks on which performance had to be monitored and the tasks on which this was not the case. That is, there are different levels of difficulty *within* 4x4 and 9x9 Sudokus, and even though the pairs of 4x4 and 9x9 puzzles that were used were at the same, standard level of difficulty according to the source, it cannot be ruled out that differences between the puzzles caused these results. Finally, although it is highly unlikely as it would have equally affected the simple tasks condition, it cannot be ruled out with this within-subjects design that participants did not provide a higher rating on the second task simply because they received an additional instruction. Therefore, the present study uses a design with monitoring as between-subjects factor and task complexity as within subjects factor, which might provide a stronger test of this hypothesis that concurrent performance monitoring increases cognitive load and decreases performance on complex tasks (i.e., high in intrinsic load) but not on simple tasks (i.e., low in intrinsic load).

## Method

### *Participants*



Fifty-eight Dutch secondary education students participated in this study (32 male; age  $M = 16.91$ ,  $SD = 0.76$ ). All participants were familiar with the rules of Sudoku but were relative novices (i.e., they did not make more than 1 or 2 Sudoku puzzles per week).

### *Materials*

*Demographic questionnaire.* This short questionnaire asked participants to indicate their age and gender.

*Tasks.* The tasks consisted of two Sudoku puzzles, one simple and one complex. Sudoku puzzles consist of a grid with several regions that has to be filled with numbers so that every row, column, and region contains only one instance of each number. The simple Sudoku puzzle (low in intrinsic load) consisted of a 4x4 grid with four 2x2 regions (mini-grids). Four cells were already filled in. The complex Sudoku puzzle (high in intrinsic load) consisted of a 9x9 grid with nine 3x3 regions (mini-grids). Thirty cells were already filled in.

*Mental effort rating scale.* Invested mental effort was measured using the 9-point subjective rating scale developed by Paas (1992). The scale ranged from (1) very, very low mental effort, to (9) very, very high mental effort.

### *Design and Procedure*

Participants were randomly assigned to either the Monitoring ( $n = 27$ ) or the No Monitoring condition ( $n = 31$ ). Task complexity was a within-subjects variable, and the order was counterbalanced (i.e., Simple-Complex or Complex-Simple). Participants first filled out the demographic questionnaire. Then, they worked on the two Sudoku puzzles, for which they were given maximally 2 minutes per puzzle (to cancel out potential interaction of time on task with mental effort measures). Given participants' expertise level, this was not enough time to solve the Complex puzzle, but participants were instructed to try and complete as much of the puzzle as possible. In the monitoring condition, participants received the additional instruction to monitor what they were doing, that is, to keep track of what they were doing, in what order,

and why (e.g., which rules they applied or which strategies they used). Start and stop times were indicated by the experimenter. Immediately after completing each puzzle, participants rated their invested mental effort on the 9-point rating scale.

### *Data Analysis*

Participants' performance was rated by counting the number of cells they correctly filled in. This resulted in a maximum score of 12 on the Simple Sudoku puzzle, and a maximum score of 51 on the Complex puzzle.

Next to performance and mental effort scores, efficiency of participants' task performance was analyzed by using the computational approach for combining measures of performance with measures of mental effort invested to attain this performance developed by Paas and Van Merriënboer (1993; see also Van Gog & Paas, 2008). Mean standardized performance (P) and mental effort (E) scores are entered into the following formula:

$$Efficiency = \frac{zP - zE}{\sqrt{2}}$$
 . A combination of equal/higher performance with lower/equal mental effort results in a higher efficiency score, whereas a combination of lower/equal performance and equal/higher mental effort results in a lower efficiency score.

### Results

The manipulation of intrinsic cognitive load (caused by the number of interacting information elements a task contains) was successful: mean mental effort invested in the Simple puzzle was 2.31 ( $SD = 1.86$ ), whereas in the Complex puzzle this was 6.87 ( $SD = 1.73$ ).

ANOVAs were used to test these directional hypotheses (i.e.,  $p$  is divided by 2). In line with our hypothesis, the instruction to monitor did not affect mental effort ratings on the Simple puzzle ( $M_{no} = 2.16$ ,  $SD = 1.86$ ;  $M_{monitoring} = 2.48$ ,  $SD = 1.88$ ),  $F(1,56) < 1$ ,  $ns$ , whereas mental effort ratings were significantly higher in the Monitoring condition on the Complex

puzzle ( $M_{\text{no}} = 6.52, SD = 1.91; M_{\text{monitoring}} = 7.30, SD = 1.41$ ),  $F(1,56) = 3.05, p = .043$  (non-directional hypothesis:  $p = .086$ ).

The instruction to monitor did not affect performance on the Simple puzzle, which was solved completely and correctly by almost all participants (max. score of 12:  $M_{\text{no}} = 11.81, SD = 0.75; M_{\text{monitoring}} = 11.85, SD = 0.77$ ). An ANOVA showed that it did, however, affect performance on the Complex puzzle, which was significantly lower for the Monitoring condition (max. score of 51:  $M_{\text{no}} = 4.64, SD = 4.51; M_{\text{monitoring}} = 2.07, SD = 2.63$ ),  $F(1,56) = 6.75, p = .006$  (non-directional hypothesis:  $p = .012$ ). Not surprisingly considering these findings on performance and mental effort measures, efficiency on the Complex puzzle, was lower for the Monitoring condition ( $M = -0.42, SD = 0.89$ ) than the No Monitoring condition ( $M = 0.36, SD = 1.35$ ),  $F(1,56) = 6.49, p = .007$  (non-directional hypothesis:  $p = .014$ ), whereas there was no difference in efficiency between the conditions ( $M_{\text{no}} = 0.04, SD = 1.29; M_{\text{monitoring}} = -0.04, SD = 1.19$ ) on the Simple puzzle,  $F(1,56) < 1, ns$ .

### Discussion

The findings show that, in line with our hypotheses and the results from the pilot study, monitoring indeed resulted in significantly lower performance, a trend towards higher cognitive load and significantly lower efficiency on complex tasks (i.e., high in intrinsic load), but not on simple tasks (i.e., low in intrinsic load). Because most meaningful learning tasks are high in intrinsic load (Sweller et al, 1998), these results may provide at least a partial explanation for why self-regulated learning is often ineffective, especially for novices. This explanation may be only partial, as there are other factors that can lead to inaccurate self-assessment (which might in turn lead to selection of inappropriate learning tasks), such as biases to which people are prone when assessing themselves (for a review, see Bjork, 1999) and a lack of knowledge of performance criteria and standards (i.e., knowledge of what aspects of performance to assess and what constitutes good, average, or poor performance on

those aspects; Dunning, Johnson, Erlinger, & Kruger, 2003). Nonetheless, these findings show that cognitive load may be an important factor to take into account in theories of self-regulated learning.

This cognitive load perspective can also explain why self-regulated learning seems to be more effective for high prior knowledge learners (e.g., Azevedo et al., 2008). Tasks that impose a high intrinsic load for novice learners are lower in intrinsic load for advanced learners because of their prior knowledge. That is, as mentioned in the introduction, high prior knowledge learners have combined different information elements present in the task in a cognitive schemata stored in long-term memory that can be retrieved and handled as a single element in working memory, which decreases the intrinsic load of the task (Sweller et al., 1998). Therefore, high prior knowledge learners may have enough cognitive capacity available for performing the learning task and monitoring their performance simultaneously. However, in effective learning trajectories, advanced learners are unlikely to work on the same tasks as novices, because they would ideally work on new tasks, not on tasks they can already perform quite well (unless the goal would be to automate their task performance). Thus, in effective learning trajectories, tasks are likely to always be quite high in intrinsic load for all learners.

An important question for future research is how this cognitive load perspective can aid the development of instructional support for self-regulated learning. One option might be to temporarily reduce the need for concurrent performance monitoring, for example by recording learners' task performance which they can then use for self-assessment afterwards (cf. Kostons, Van Gog, & Paas, in press). Another option that could maintain positive effects on motivation or involvement without negative effects on learning outcomes, might be to provide novices with only partial control initially (Corbalan et al., 2006), and to gradually increase that control with learners' increasing expertise.

In future studies, it would be interesting to include tasks of various intermediate levels of complexity. On our simple tasks, a ceiling effect occurred, which may have masked a potential effect of monitoring. It would also be interesting to include measures of performance on the secondary task (i.e., monitoring), because a limitation of this study is that we did not investigate effects on the quality of monitoring. For tasks that were high in intrinsic load, monitoring had a negative effect on performance on the primary task. It would be interesting to investigate in future studies how performance on the secondary task is affected in order to find out whether the negative effect on primary task performance occurs because more cognitive resources are allocated to monitoring in order to perform that task well, or whether performance on both tasks suffers as a consequence of the need to divide cognitive resources between them.

## References

- Azevedo, R., & Cromley, J. G., (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*, 523-535.
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. C. (2008). Why is externally-regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*(1), 45-72.
- Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: MIT Press.
- Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*, 53-61.
- Camp, G., Paas, F., Rikers, R. M. J. P., & Van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort, and mental efficiency. *Computers in Human Behavior, 17*, 575-595.
- Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2006). Towards a personalized task selection model with shared instructional control. *Instructional Science, 34*, 399-422.
- Dunning, D., Johnson, K., Erlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*, 83-87.
- Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558-568.
- Kanfer, R., & Ackerman, P. (1989). Motivation and cognitive abilities: An integrative/apptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-690.

- Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology, 23*, 1256-1265.
- Niemiec, R. P., Sikorski, C., & Walberg, H. J. (1996). Learner-control effects: A review of reviews and a meta-analysis. *Journal of Educational Computing Research, 15*, 157-174.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology, 84*, 429-434.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71.
- Paas, F., & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*, 737-743.
- Salden, R. J. C. M., Paas, F., & Van Merriënboer, J. J. G. (2006). Personalised adaptive task selection in air traffic control: Effects on training efficiency and transfer. *Learning and Instruction, 16*, 350-362.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.
- Van den Boom, G., Paas, F., & Van Merriënboer, J. J. G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction, 17*, 532-548.
- Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16-26.

- Van Gog, T., & Paas, F. (2009). Effects of concurrent performance monitoring on cognitive load as a function of task complexity. In N. Taatgen, & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1605-1608). Austin, TX: Cognitive Science Society.
- Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*, 237-244.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*, 3-17.



Footnote

<sup>1</sup> Note that the term monitoring (or monitoring accuracy) is used by some authors to refer to learners' judgments about their learning made after engaging in the task, which is what we would call self-assessment (accuracy) here.