STUDIA UBB PSYCHOL.-PAED., LX, 1, 2015, p. 5 - 27
(RECOMMENDED CITATION)

THE IMPACT OF INSTRUCTIONAL SUPPORT ON STUDENTS' PERFORMANCE: AN EYE TRACKING STUDY

LOREDANA MIHALCA[1]

ABSTRACT. This study examined the effects of various types of instructional support (incomplete worked-out examples - IWE; completion problems - CMP and conventional problems - CVP) on students' attention allocation and performance as they use a computer-based learning environment. Eye movement and performance data were obtained from sixty-three university students who were randomly assigned to one of the three conditions. Results indicated significant differences regarding the number and duration of fixations, not only between the three conditions, but also between these conditions across the five genetics problems students solved during training. These findings indicate that the type of instructional support differently impacted the students' allocation of visual attention, with participants in the IWE condition having significantly lower number of fixations and shorter fixation durations. In addition, the number and duration of fixations were significantly lower in the CMP condition than in the CVP condition, but only at the beginning and the end of training (i. e., problems 1 and 5). Finally, the amount of processing time as revealed by the eye tracking data was associated with the learning outcomes (e. g., training performance), indicating that visual attention is indeed related to the strategic processing during problem solving. The combined use of eye tracking data with learning measures provides a more comprehensive picture of the cognitive processes that underlie the problem solving.

Keywords: *incomplete worked-out examples, completion problems, eye tracking, attention allocation, fixation duration.*

ZUSAMMENFASSUNG. Diese Studie untersuchte die Effekte verschiedener Lernunterstützungen (ausgearbeitete Lösungsbeispiele – IWE; teilweise ausgearbeitete Problemlösungsaufgaben – CMP und konventionelle Problemlösungsaufgaben – CVP) auf die Aufmerksamkeit und Leistung während des Lernens in einer computergestützten Lernumgebung. Augenbewegungen und Leistungsdaten wurden von 63 Studierenden erhoben, die zufällig zu den drei Versuchsbedingungen zugeteilt wurden. Die Ergebnisse

[1] *Psychology & Counseling Department, Webster University Geneva, Switzerland, Email: mihalca@webster.ch*

zeigten signifikante Unterschiede hinsichtlich der Zahl und Dauer der Fixierungen, nicht nur zwischen den drei Gruppen, sondern auch zwischen den Gruppen über die fünf Genetik-Aufgaben hinweg, die die Studierenden während des Trainings lösen mussten. Diese Ergebnisse zeigen, dass die verschiedenen Lernunterstützungen die visuelle Aufmerksamkeit der Studierenden unterschiedlich beeinflussten. Darüber hinaus korrelierte die Verarbeitungszeit, die durch die Augenbewegungen erfasst wurde, mit den Lernergebnissen (der Trainingsleistung), d. h. visuelle Aufmerksamkeit ist mit der strategischen Verarbeitung während des Problemlösens verknüpft. Die Verbindung von Eye-Tracking und Leistungsmessung ergibt ein umfassenderes Bild der kognitiven Prozesse, die dem Problemlösen zu Grunde liegen.

Schlüsselwörter: *ausgearbeitete Lösungsbeispiele, teilweise ausgearbeitete Problemlösungsaufgaben, Eye-Tracking, Aufmerksamkeit, Fixationsdauer.*

1. Introduction

The benefits of presenting students with worked-out examples (i. e., step-by-step demonstrations of how to solve a problem) as compared to completion or conventional problems have been shown in many studies (e. g., Atkinson, Derry, Renkl & Wortham, 2000; Renkl & Atkinson, 2003). However, the vast majority of these studies have used product-related measures (i. e., off-line written tests) from which the benefits of worked-out examples were inferred. To our knowledge, none of these studies have tested the differential effects of instructional support type (i. e., worked-out examples, completion problems, and conventional problems) directly, for example, by using recordings of participants' eye movements (see Van Gog, Kester, Nievelstein, Giesbers & Paas, 2009). Eye movement data can provide more subtle insights into the cognitive processes of problem solving and reveal changes in these processes at an individual level that may occur over time as a result of practice (e. g., Van Gog, Paas & Van Merriënboer, 2004). Therefore, the purpose of this study was to examine students' cognitive processing while learning about genetics with different types of instructional support embedded in a computer-based learning environment (CBLE) using both eye tracking and learning measures. Specifically, this study explored the differences in students' attention allocation in learning from incomplete worked-out examples and completion problems in comparison with conventional problem solving, as well as on their performance and cognitive load.

1.1. Worked-out Examples, Incomplete Worked Examples, and Completion Problems

Problem-solving schemas (i. e., knowledge about problems and corresponding solving procedures) are more easily acquired when learning from worked-out examples rather than traditional problem solving without any support (i. e., conventional problems), at least when students have low prior knowledge in a domain (e. g., Van Gog, Kester & Paas, 2011). Worked-out examples provide students with full instructional support, which consists of a description of the problem state, the solution steps, and the final solution itself, while conventional problems provide only a description of the problem state together with the final solution and students have to complete the solution steps themselves. Numerous studies have shown that instruction using worked-out examples is more effective (i. e., higher performance) and efficient (i. e., higher performance combined with lower instructional time and/or cognitive load) than instruction consisting of conventional problems (for overviews, see Atkinson et al., 2000; Renkl, Hilbert & Schworm, 2009; Van Gog & Rummel, 2010). This is called *worked-out example effect* (see Sweller, Van Merriënboer & Paas, 1998; Sweller, Ayres & Kalyuga, 2011, for more details). The beneficial effects are explained by the fact that worked-out examples reduce unproductive search processes and thus free up the cognitive resources needed for acquiring problem-solving schemas (see Sweller et al., 1998; Van Gog et al., 2011, for reviews).

Contrary to the use of worked-out examples, when solving conventional problems students employ a means-ends analysis, which causes excessive cognitive load and ultimately hinders schema acquisition (Renkl et al., 2009; Van Gog et al., 2011). Cognitive load is defined as the burden imposed on working memory capacity by problem-solving processes (Paas &Van Merriënboer, 1994; see also Paas, Tuovinen, Tabbers & Van Gerven, 2003). Mental effort is the subjective component of cognitive load (Paas et al., 1994), which refers to the cognitive resources allocated to deal with the demands imposed by solving a given problem (Paas et al., 2003). Another subjective measure of the cognitive load refers to the ratings of task difficulty (e. g., Kalyuga, Chandler & Sweller, 1999). The perceived task difficulty is a direct result of the number of elements that need to be simultaneously processed in working memory to understand the instructional material (e. g., a given problem; see Schnotz, 2010). In general, the greater the number of elements that need to be processed simultaneously, the higher the cognitive load, which is reflected by an increase in perceived task difficulty and the amount of mental effort invested (Brünken, Plass & Pleutner, 2003).

Despite the evidence supporting the benefits of worked-out examples in comparison with conventional problems, the findings vary and are contradictory (e. g., Kalyuga, Chandler, Tuovinen & Sweller, 2001; Van Merriënboer & Sweller, 2005). For example, some studies have shown that simply providing worked-out examples is not sufficient to improve learning, because there is no guarantee that students deeply process and elaborate on the completed solutions steps (see Paas & Van Gog, 2006). A deeper processing and understanding of the worked-out examples can be fostered by making the sub-goals in a solution plan explicit through labeling or visually isolating them (see Catrambone 1995, 1996). According to the sub-goal learning model (Catrambone 1995, 1996), when the solution steps are emphasized by labeling or visually isolating them, students are more likely to self-explain how these steps are connected and contribute to the final solution, which in turn promotes successful learning (see also Renkl et al., 2009).

Another method to improve deep processing is providing so-called incomplete worked-out examples, in which one solution step is omitted. Incomplete worked-out examples have been shown to be more effective than fully worked-out examples in which all steps were presented (e. g., Renkl, 2002; Stark, 1999). Comparing incomplete worked-out examples with fully worked-out examples, Stark (1999) found that the insertion of ”blanks” into a sequence of solution steps (making the solution partially incomplete) prompted students' self-explanations, which in turn increased their performance. When faced with “blanks” in the solution steps, students engage more deeply and actively in processing the worked-out steps, because they have to complete by themselves the missing solution (e. g., Renkl, 2002).

In contrast to incomplete worked-out examples, completion problems provide a partial solution with more than just one solution step omitted, representing a transition from incomplete worked-out examples to conventional problems (Van Merriënboer & de Croock, 1992; VanMerriënboer, Schuurman, de Croock & Paas, 2002). Due to the partially provided solution, completion problems stimulate students to process more deeply the completed solution steps and enable them to acquire more complex problem-solving schemas (so-called *completion problem effect*; e. g., Van Merriënboer et al., 2005). It has been shown that completion problems decrease students' excessive cognitive load compared to conventional problems by focusing their attention on relevant aspects of the problem (i. e., the solution steps), which enables them to achieve a higher test performance (e. g., Van Merriënboeret al., 1992).

1. 2. Using eye-tracking methodology to study problem-solving processes

Eye-tracking methodology has been mostly used in text reading and scene perception for studying *where* the learners were looking (i. e., fixation locations), and *for how long* (i. e., fixation duration; for a review, see Rayner, 1998; see also Gegenfurtner, Lehtinen & Säljö, 2011). Put differently, eye-tracking methodology provides insights into the learners' allocation of visual attention to different problem elements, revealing in-depth information about problem-solving processes (e. g., Knoblich, Öllinger & Spivey, 2005). Peters (2010, p. 2) defined problem solving as "a systematic process of matching items in declarative memory with rules. This process is assumed to continue until the goal is achieved and the problem solved."

According to the eye-mind hypothesis (Just & Carpenter, 1980), which theoretically underlies eye-tracking methodology, there is a close relationship between the information a learner is looking at and the information s/he is thinking about, as well as a relationship between the time spent fixating on information (i. e., fixation duration) and the amount and difficulty of cognitive processing (see Tsai, Hou, Lai, Liu & Yang, 2012). In general, the longer the information is fixated upon, the deeper it is processed and the more difficult it is to comprehend it (e. g., Chuang& Liu, 2012).

Despite the usefulness of eye-tracking methodology, only few studies have used this tool to investigate learners' attention allocation in a problem-solving context (e. g., Hegarty, Mayer & Green, 1992; Hegarty & Just, 1993; Graesser, Lu, Olde, Cooper-Pye & Whitten, 2005), and even fewer that have explored attention allocation in a science problem-solving context (e. g., Tsai et al., 2012). For example, Hegarty and colleagues (e. g., Hegarty et al., 1992; Hegarty, Mayer & Monk, 1995) examined students' eye fixations while solving mathematics word problems. The eye-fixation and performance data obtained in Hegarty et al. 's (1995) study indicated that there were significant differences between successful and unsuccessful problem solvers regarding both the pattern of eye fixations and the pattern of errors in remembering the problem statement. More specifically, unsuccessful students paid more attention (i. e., higher percentage of fixations) to the variable names and numbers when re-reading the problem statement compared to successful students, who paid less attention to these aspects. This result suggests that unsuccessful and successful students use different comprehension strategies during the problem-solving process, that is, a direct-translation strategy (i. e., focusing more on numbers and relational terms such as *more* or *less,* and solving the problem based on these terms), and a problem-model strategy (i. e., focusing more on the variables names, and developing a mental model of the given state of the problem), respectively (Hegarty et al., 1995). Finally, compared to unsuccessful students, successful ones remembered better the given state of the problem, which indicates that they are able to form a meaningful representation of the problem.

Using eye-tracking methodology, Tai, Loehr, and Brigham (2006) investigated the differences in problem-solving behaviors of students with various expertise levels (i. e., prior knowledge) levels in three science disciplines: biology, chemistry and physics. The results indicated that the more expertise a student has in a given discipline, the fewer the fixations on the look zones (i. e., the specific areas of each assessment item such as problem statement zone, multiple choice zone, etc.) and the lower the saccades between these zones. In addition, these authors found differences in eye movements not only between students within a particular discipline, but also within students across the three science disciplines.

More recently, Tsai et al. (2012) examined students' attention allocation while solving a multiple-choice science problem. The results revealed that students paid more attention to the chosen options than the rejected ones. Successful problem solvers spent more time inspecting relevant aspects of the problem rather than the irrelevant aspects when compared to the unsuccessful students, who experienced difficulties in recognizing the relevant aspects. Furthermore, whereas successful problem solvers moved their visual attention from irrelevant to relevant aspects of the problem, for unsuccessful students the opposite was true. Although the last two mentioned studies have provided insights into students' attention allocation while solving science problems, to our knowledge no study has investigated how students allocate their visual attention to science problems when they receive different levels of instructional support (i. e., IWE, CMP, and CVP).

1. 3. The present study

The purpose of this study was to examine the influence of different types of instructional support embedded in a CBLE (i. e., incomplete worked-out examples and completion problems in comparison to conventional problems) on students' attention allocation, as well as on their performance and cognitive load when learning genetics. In addition, the current study investigated the relationship between the on-line processing of the problems (i. e., number of fixations and fixation durations) and the off-line outcomes of learning with different types of instructional support (i. e., training performance, learning time, and cognitive load during training).

Regarding learning outcomes and investments, it was expected that all types of instructional support would lead to cognitive gains (i. e., higher performance) and decreased cognitive load (i. e., perceived task difficulty and invested mental effort) from pre-to-post-test. Furthermore, we predicted that students learning with incomplete worked-out examples and completion problems would achieve a better post-test performance, and experience a lower cognitive load in post-test compared to the conventional problems condition.

Second, we expected that different patterns of visual behavior would emerge from the various indices of eye movement during learning with incomplete worked-out examples, completion and conventional problems. More specifically, we predicted significant differences between the experimental conditions in terms of eye fixations, not only across all students within each condition, but also across the five genetics problems that students had to solve during training. Put differently, it was expected that the condition differences in terms of eye fixations would be significant between the five genetics problems solved in training.

Finally, we expected a clear connection between the on-line cognitive processing revealed by eye-tracking data and off-line learning outcomes during training. More specifically, we predicted that the number and the duration of fixations would positively correlate with training performance, learning time, as well as with the perceived task difficulty and mental effort invested during training. In other words, the more often and longer a problem is fixated, the greater the performance in solving that problem, although it would be perceived as more difficult.

2. Method

2. 1. Participants

Participants in this study were sixty-three students enrolled at a small university in Germany (11 males and 52 females; age *M* = 22. 75 years, *SD* = 2. 72). The participants were randomly assigned to one of the three conditions, an incomplete worked-out examples (IWE) condition (*n*= 21), a completion problems (CMP) condition (*n*= 21), and a conventional problems (CVP) condition (*n*= 21). All participants had normal or corrected-to-normal vision and all had at least some basic knowledge of Mendel's Laws, the topic of the study. As compensation for their participation to the study, participants received either €10 or credit points toward their research experience requirement.

2. 2. Apparatus and Materials

Electronic learning environment. The learning environment consisted of a Web application written in PHP scripting language, with a MySQL database connected to it (based on Mihalca, Mengelkamp, Schnotz & Paas, 2015). The database contained 15 genetics problems addressing Mendel's laws and registered all student interactions with the system: performance, perceived difficulty, mental effort scores in all phases of the study (i. e., pre-test, training and post-test), and learning time.

Genetics problems. During the training phase, participants had to solve a total of five genetics problems on the subject of heredity according to Mendel's laws, which differed with regard to the amount of embedded support depending on the experimental condition. Whereas incomplete worked-out examples provided high support, completion and conventional problems provided low support and no support, respectively. Specifically, incomplete worked-out examples provided four out of five solution steps, and students had to complete the final step. Completion problems provided two solution steps out of five, while conventional problems did not provide any support, that is, students had to solve all the steps on their own (see Figure 1 for a screenshot of a conventional problem).

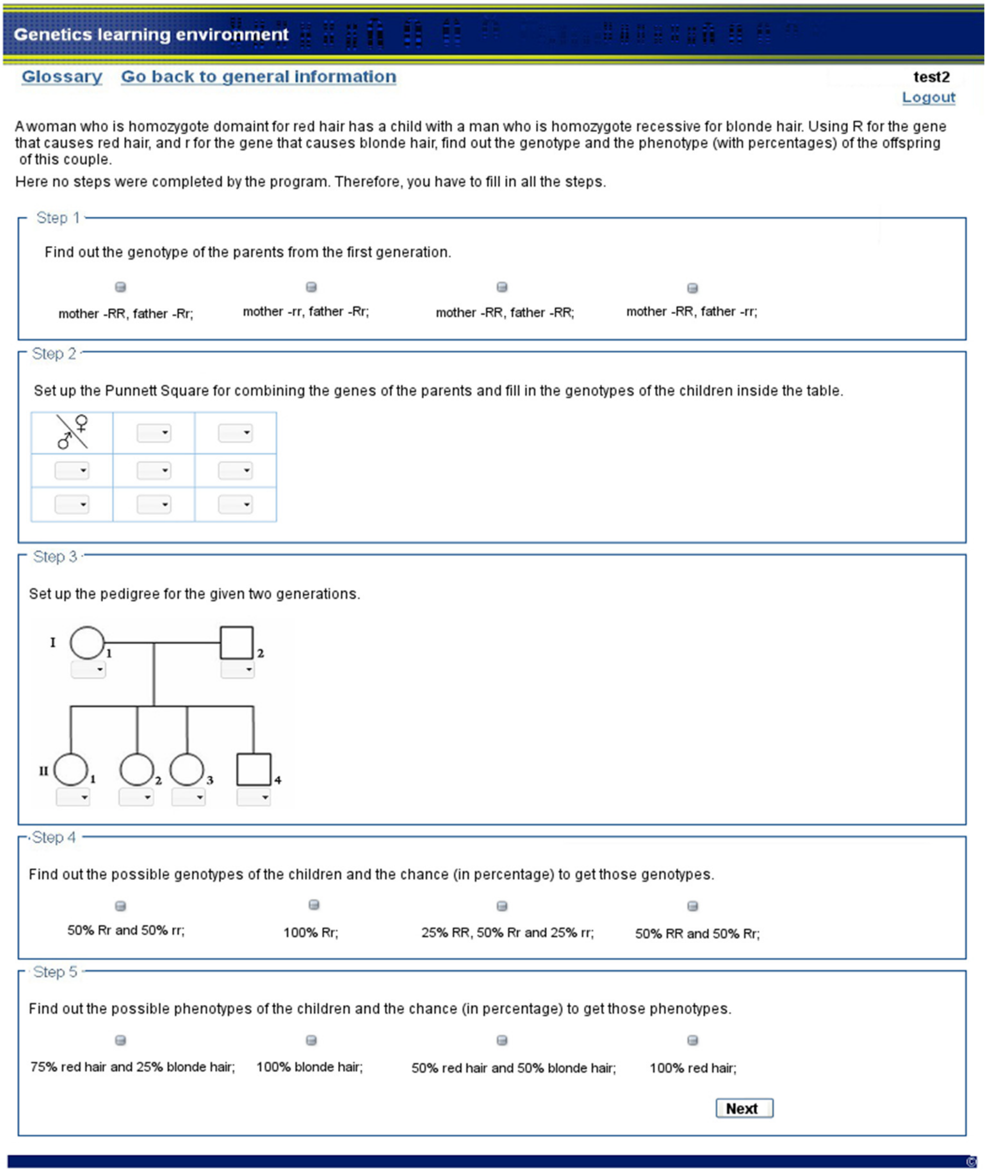


Figure 1. Screenshot of a conventional problem (CVP) in which none of the solution steps were completed by the program. It should be noted that incomplete worked-out examples (IWE) and completion problems (CMP) had the same structure as CVPs, with the difference that in the IWE condition four out of five solution steps were completed by the program and in the CMP condition two out of five steps were completed by the program.

In all conditions the genetics problems were presented in a predetermined simple-to-complex sequence from difficulty level one to difficulty level five. The five difficulty levels were defined in cooperation with two domain experts using several problem characteristics, such as the number of generations (two or three), the number of possible correct solutions (one or two), and the type of reasoning (inductive and/or deductive reasoning).

Eye tracking equipment. Participants' eye movements were recorded using a 60Hz Tobii T60XL eye tracking system, which was integrated into a 24-inch monitor with a maximum resolution of 1920 x 1200 pixels. Tobii StudioTM software was used to analyze participants' eye movements and their keyboard and mouse actions.

2. 3. Instruments

The pre- and post-test consisted of the same seven multiple-choice questions on the subject of heredity (i. e., Mendel's Laws). In order to provide the final solution for problems included in the pre-and post-test participants had to perform all sub-steps (i. e., solution steps) mentally, which was not the case for the problems included in training. The maximum score was 7 points, one point for each correct answer. The reliability (Cronbach's alpha) of the pre-test was. 62, and of the post-test was. 66.

The perceived task difficulty and invested mental effort were measured after each problem during all phases of the study (pre-test, training, and post-test) on a 5-point rating scale with values ranging from 1 (very low) to 5 (very high). These rating scales have been widely used in educational research and shown to be highly reliable (see Paas et al., 2003).

2. 4. Procedure

The experiment was run in individual sessions of approximately 90 minutes. First, participants were given general instructions explaining the procedure and introducing the topic of the study. They were asked to sign an agreement. The participants then started with the pre-test, and after that read a basic introduction. The basic introduction included the main genetics concepts needed for solving the genetics problems during training (e. g., dominant and recessive genes, homozygote and heterozygote gene pairs, genotype and phenotype). After they completed the pre-test, the participants were seated in front of the stimulus PC (65 cm away from the screen) and the eye tracking system was calibrated. During calibration, the eye tracker measured the characteristics of participants' eyes, required to accurately calculate gaze direction. To familiarize participants with the thinking aloud procedure (i. e., they had to verbalize their own thoughts while solving the genetics problems), they were given a warming-up task. As the thinking aloud data are not relevant for the hypotheses of the current study, they are not reported here. When participants had finished the warming-up task, they started with the training phase. During the training phase the eye movements of all participants were tracked. The participants had to rate the perceived difficulty of each problem and the invested mental effort before the program would let them proceed. Immediately after the training, participants performed the post-test and rated again the perceived difficulty and mental effort after each solved problem. The time spent on each part of the experiment was logged.

2. 5. Eye movement analysis

To analyze participants' eye movements, six areas of interest (AOIs) corresponding to the problem statement and each of the solution steps were defined. AOIs were defined to determine whether and for which amount of time participants were looking at each specific area during the problem solving process. First, the *total number of fixations* was determined for each participant by summing all single fixations on the six AOIs. Second, the *total fixation duration* (or the time for which participants visually inspected the problems) was computed in a similar way by summing all single fixation durations on the six AOIs. Table 1 presents the means and standard deviations for the number of fixations and fixation durationsin training for each experimental condition.

3. Results

For all analyses the level of alpha error was set to. 05. Cohen's (1992) taxonomy of effect sizes was used to classify effects as small, medium or large corresponding to values of. 01,. 06,. 14 for eta-squared, respectively. We calculated performance in pre-test and post-test as proportion correct (i. e., total number of correct items divided by number of all items).

Before testing our hypotheses we computed a MANOVA to check if there were any differences between the three conditions (i. e., IWE, CMP, and CVP) in terms of pre-test performance, perceived difficulty and mental effort invested during the pre-test. Box's M test indicated no violation of the equality of covariance matrices (p =. 293). The results of MANOVA showed no significant effect of condition using Pillai's trace, V =. 080, $F(6, 118)$ =. 824, p =. 554, η_p^2 =. 04, on any of the variables mentioned before. Therefore, there were no significant differences between the conditions prior to the training phase (see Table 2 for the means and standard deviations of the pre-test performance, perceived difficulty and mental effort in pre-test).

Table 1. Descriptive Statistics for Training Problems by Condition

	IWE (n = 21)		CMP (n = 21)		CVP (n = 21)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Training Problem 1						
1. Performance	. 64	. 26	. 61	. 31	. 60	. 22
2. Perceived difficulty	3. 31	. 77	3. 69	. 60	3. 64	. 81
3. Mental effort	3. 20	. 79	3. 23	. 63	3. 46	. 65
4. Number of fixations	301. 62	103. 10	363. 95	108. 35	390. 90	155. 41
5. Fixation duration	86. 55	39. 25	115. 01	38. 33	129. 69	48. 52
Training Problem 2						
6. Performance	. 64	. 26	. 61	. 31	. 60	. 22
7. Perceived difficulty	3. 31	. 77	3. 69	. 60	3. 64	. 81
8. Mental effort	3. 20	. 79	3. 23	. 63	3. 46	. 65
9. Number of fixations	323. 81	100. 40	634. 10	186. 70	491. 86	167. 66
10. Fixation duration	97. 37	33. 98	191. 79	46. 11	161. 61	55. 44
Training Problem 3						
11. Performance	. 64	. 26	. 61	. 31	. 60	. 22
12. Perceived difficulty	3. 31	. 77	3. 69	. 60	3. 64	. 81
13. Mental effort	3. 20	. 79	3. 23	. 63	3. 46	. 65
14. Number of fixations	268. 81	80. 74	490. 38	196. 50	335. 24	157. 74
15. Fixation duration	80. 26	30. 26	148. 73	61. 80	111. 79	45. 06
Training Problem 4						
16. Performance	. 64	. 26	. 61	. 31	. 60	. 22
17. Perceived difficulty	3. 31	. 77	3. 69	. 60	3. 64	. 81
18. Mental effort	3. 20	. 79	3. 23	. 63	3. 46	. 65
19. Number of fixations	419. 67	132. 52	850. 14	317. 12	708. 81	340. 50
20. Fixation duration	124. 55	48. 16	258. 66	113. 25	214. 49	91. 31
Training Problem 5						
21. Performance	. 64	. 26	. 61	. 31	. 60	. 22
22. Perceived difficulty	3. 31	. 77	3. 69	. 60	3. 64	. 81
23. Mental effort	3. 20	. 79	3. 23	. 63	3. 46	. 65
24. Number of fixations	435. 00	220. 07	850. 67	346. 02	992. 00	372. 85
25. Fixation duration	138. 21	80. 42	241. 95	99. 77	297. 67	90. 80

Note. IWE = incomplete worked examples, CMP = completion problems, CVP = conventional problems.

3. 1. Off-line learning outcomes

In order to examine the differential effects of conditions on learning gains (i. e., pre-to-post-test performance gain scores), a mixed ANOVA using condition as the between-subject factor, and pre-and post-test performance as the within-subject factor was conducted. The results showed that students' performance increased from pre-test (M_{total} =. 62, *SD* =. 26) to post-test (M_{total} =. 84, *SD* =. 20), and this increase was significant, $F(1, 60) = 51.13$, $p < .001$, $\eta_p^2 = .46$. However, mixed ANOVA showed no significant main effect of the condition, $F(2, 60) = .277$, $p = .759$, $\eta_p^2 = .01$, and no interaction between the increase in pre-to-post-test performance and condition, $F(2, 60) = .158$, $p = .854$, $\eta_p^2 = .01$. In other words, the increase of performance from pre-to-post-test was the same for all conditions.

In addition, a mixed ANOVA with condition as the between-subject factor and perceived difficulty from pre- to post-test as the within-subject factor showed that students' perceived difficulty decreased from pre-test (M_{total} = 3. 53, *SD* =. 75) to post-test (M_{total} = 2. 67, *SD* =. 66), and this decrease was significant, $F(1, 60) = 127.76$, $p < .001$, $\eta_p^2 = .68$. Mixed ANOVA showed no significant main effect of condition on the pre- to post-test perceived difficulty, $F(2, 60) = .871$, $p = .424$, $\eta_p^2 = .03$, and no interaction between pre- to post-test decrease in perceived difficulty and condition, $F(2, 60) = 1.91$, $p = .157$, $\eta_p^2 = .06$.

Furthermore, a mixed ANOVA showed that students' invested mental effort decreased from pre-test (M_{total} = 3. 28, *SD* =. 69) to post-test (M_{total} = 2. 95, *SD* =. 79), and this decrease was significant, $F(1, 60) = 20.50$, $p < .001$, $\eta_p^2 = .26$. However, the results showed no significant main effect of condition on the pre- to post-test mental effort, $F(2, 60) = 1.10$, $p = .339$, $\eta_p^2 = .04$, and no interaction between pre- to post-test decrease in mental effort and condition, $F(2, 60) = .753$, $p = .475$, $\eta_p^2 = .02$. Put differently, the decrease of cognitive load (i. e., perceived difficulty and mental effort) from pre- to post-test was the same for all conditions.

Finally, in order to analyze the effect of different conditions on students' performance and cognitive load after training, we calculated a MANOVA using post-test performance, perceived difficulty and mental effort invested in the post-test as dependent variables. Box's M test indicated no violation of the equality of covariance matrices ($p = .159$). The results of MANOVA showed no significant effect of condition on any of the dependent variables using Pillay's trace, $V = .061$, $F(6, 118) = .614$, $p = .718$ (see Table 2 for the means and standard deviations of the post-test performance, perceived difficulty and mental effort in post-test). Therefore, in all three conditions students obtained the same post-test performance and experienced the same level of cognitive load while completing the post-test.

Table 2. Descriptive Statistics for Pre-and-Post-test by Condition

	IWE (n = 21)		CMP (n = 21)		CVP (n = 21)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Pretest						
1. Performance	4. 48	1. 83	4. 29	2. 15	4. 19	1. 53
2. Perceived difficulty	23. 14	5. 42	25. 81	4. 18	25. 19	5. 89
3. Mental effort	22. 43	5. 53	22. 62	4. 42	23. 86	4. 53
Posttest						
4. Performance	6. 00	1. 18	6. 05	1. 36	5. 67	1. 74
5. Perceived difficulty	18. 19	5. 87	18. 38	3. 63	19. 52	4. 17
6. Mental effort	20. 19	6. 74	19. 52	4. 27	22. 29	5. 16

Note. IWE = incomplete worked examples, CMP = completion problems, CVP = conventional problems.

3. 2. On-line processing during training (i. e., eye tracking data)

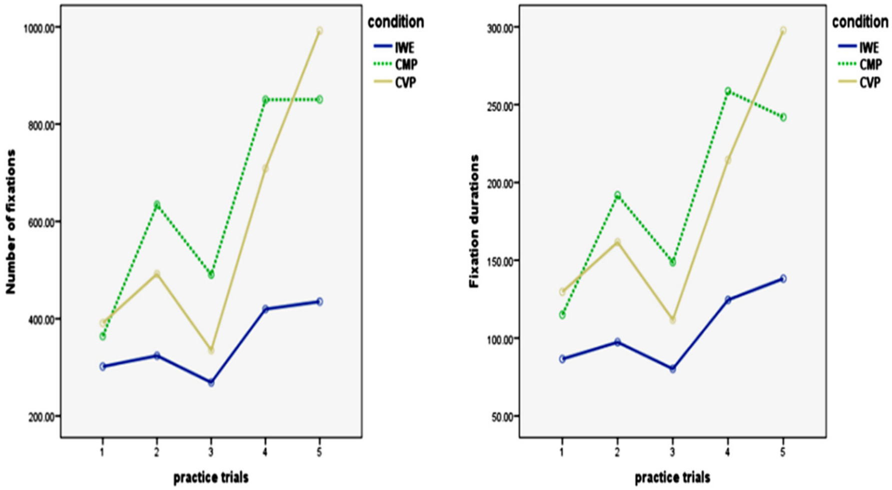
A series of mixed ANOVA with condition as the between-subject factor, and number of practice trials (i. e., problems 1 to 5 in training) as the within-subject factor were performed to examine the indices of visual attention allocation (i. e., number of fixations and fixation duration) during problem-solving process in training. Bonferroni correction was used as a post-hoc analysis to evaluate pairwise differences among practice trials, controlling for Type 1 error across tests.

The total number of fixations was analyzed using a mixed ANOVA with condition as the between-subject factor, and number of practice trials (i. e., problems 1 to 5 in training) as the within-subject factor. Mauchly's Test indicated that the assumption of sphericity had been violated ($\chi^2(9)$ = 44. 96, p<. 001), therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε =. 735). The results showed that there was a significant effect of the number of practice trials on the number of fixations, F(2. 94, 176. 31) = 70. 00, p<. 001, η_p^2 =. 54, and a significant interaction between the number of practice trials and condition, F(5. 88, 176. 31)=9. 66, p<. 001, η_p^2 =. 24. In other words, there were significant differences between the conditions in terms of number of fixations across the five genetics problems in training (see Figure 2). More specifically, the number of fixations in the CMP condition was significantly lower for the first and fifth training problems compared to the CVP condition, whereas for the other three problems (i. e., problems 2, 3 and 4) the number of fixations was higher in the CMP condition compared to the CVP condition. For the factor number of practice trials, Bonferroni corrected post-hoc tests showed that the number of fixations for problem 1 and problem 3 did not significantly differ (p = 1. 0), but the number of fixations for problem 1 was significantly lower than for all the other problems (i. e., problems 2, 4 and 5; all ps <. 05). In addition, the number of fixations for problem 5 was significantly higher compared to the number of fixations of all the other four problems (all ps <. 05).

Furthermore, a main effect of condition on the number of fixations, F(2, 60) =19. 48, p<. 001, η_p^2 =. 39, was found. Post-hoc comparisons using Games-Howell tests (a post-hoc test for heterogeneous variances between groups) indicated that participants in the IWE condition had a significantly lower number of fixations than participants in both the CMP and CVP conditions (both ps <. 001). However, the number of fixations did not differ significantly among the CMP and CVP conditions (p =. 612).

In a similar way with the number of fixations, the fixation durations were analyzed using a mixed-design ANOVA with condition as the between-subject factor and number of practice trials (i. e., problems 1 to 5 in training) as the within-subject factor. Mauchly's Test indicated that the assumption of sphericity had been violated (χ^2(9) = 57. 08, p <. 001), therefore the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε=. 722). The results showed that there was a significant effect of the number of practice trials on fixation duration, F(2. 89, 173. 23) =64. 55, p<. 001, η_p^2 =. 52, and a significant interaction between the number of practice trials and condition, F(5. 77, 173. 23)=8. 08, p<. 001, η_p^2=. 21, which indicates condition differences in terms of fixation durations across the five genetics problems (see Figure 3). Specifically, the fixation durations on first and fifth training problems were significantly shorter in the CMP condition when compared to the CVP condition, whereas for the other three problems (i. e., problems 2, 3 and 4) the fixation durations were longer in the CMP condition when compared to the CVP condition. For the factor number of practice trails, Bonferroni corrected post-hoc tests showed that the fixation durations on problem 1 and problem 3 did not significantly differ (p = 1. 0), but the fixation durations for problem 1 were significantly shorter than for all the other problems (i. e., problems 2, 4 and 5; all ps <. 001). In addition, the fixation durations on problem 5 were significantly longer compared to the fixation durations on all the other problems (all ps <. 01), except for problem 4 (p =. 177).

Finally, a main effect of condition on fixation durations, F(2, 60) =19. 85, p<. 001, η_p^2=. 40, was found. Post-hoc comparisons using Games-Howell tests indicated that participants in the IWE condition had significantly shorter fixation durations than participants in both the CMP and CVP conditions (both ps <. 001). However, the fixation durations did not differ significantly between the CMP and CVP conditions (p =. 874).



Figures 2 and 3. Graphical representation of the interaction between the type of instructional support (i. e., conditions) and the practice trials on number of fixations (Figure 2), and fixation durations (Figure 3).

3. 3. Connecting on-line processing with off-line learning in training

In order to examine whether the eye tracking measures and learning outcomes in training are related, Kendall's tau correlations were conducted (see Table 3). We choose to use Kendall's tau correlations rather than Spearman's statistic – which is the more popular of the two nonparametric correlations – because Kendall's statistics is a better estimate of the correlation in the population and therefore more accurate generalizations can be drawn (see Field, 2005, p. 181).

A positive correlation between the number and duration of fixations of each training problem and the performance and time spent on that specific problem, as well as between the number and duration of fixations of each problem and the cognitive load (i. e., perceived difficulty and mental effort) experienced while solving that specific problem was expected.

As shown in Table 3, the eye tracking parameters were positively correlated with most of the learning outcomes obtained during training. Notably, there were significant positive correlations between the number and duration of fixations on almost all training problems and the training performance on those problems, their perceived difficulty, the invested mental effort, as well as the time spent on solving the problems. However, there were no significant correlations between the number of fixations on problem 1 and problem 3 and the performance scores obtained on those two problems (τ=. 160, p =. 082, and τ=. 218, p =. 086, respectively).

Furthermore, there were no significant correlations between the number and duration of fixations on problems 4 and 5 and the mental effort invested on those two problems (all *ps*>. 05).

Table 3. Kendall's Tau Correlations Between the Measures (*N* = 63)

Training	Prob. 1 N. of fixations	Prob. 1 Fix. duration	Prob. 2 N. of fixations	Prob. 2 Fix. duration	Prob. 3 N. of fixations	Prob. 3 Fix. duration	Prob. 4 N. of fixations	Prob. 4 Fix. duration	Prob. 5 N. of fixations	Prob. 5 Fix. duration
Problem 1										
1. Performance	.16	.24*								
2. PD	.40**	.38**								
3. Mental Effort	.25*	.32**								
4. Time	.73**	.24*								
Problem 2										
1. Performance			.29**	.32**						
2. PD			.28**	.34**						
3. Mental Effort			.30**	.34**						
4. Time			.71**	.83**						
Problem 3										
1. Performance					.20*	.26**				
2. PD					.39**	.36**				
3. Mental Effort					.26**	.31**				
4. Time					.77**	.82**				
Problem 4										
1. Performance							.33**	.36**		
2. PD							.30**	.33**		
3. Mental Effort							.18	.20*		
4. Time							.76**	.80**		
Problem 5										
1. Performance									.37**	.39**
2. PD									.22*	.18
3. Mental Effort									.12	.13
4. Time									.79**	.86**

Note: PD = perceived difficulty, Prob. = problem, N = numbers, Fix. = fixations

To sum up, overall the more often students look at the problems, and the longer the amount of time they spent processing the problems, the higher the performance on those problems and the greater the perceived difficulty (and invested mental effort). Moreover, the number of fixations on each training problem was highly correlated with the total fixation durations on that specific problem (with τ values ranging from. 690 to. 808; all *ps*<. 001).

4. Discussion

Combining eye-tracking measures with performance data, this study explored the differences in attention allocation during problem-solving processes while using various types of instructional support (i. e., incomplete worked-out examples, completion and conventional problems), by

The first hypothesis that all types of instructional support would lead to cognitive gains (i. e., higher performance) and would decrease cognitive load (i. e., perceived difficulty and invested mental effort) from pre- to post-test, was confirmed by the data. More specifically, all types of instructional support significantly increased students' performance from pre- to post-test, while experiencing a lower cognitive load in post-test compared to pre-test. The combined results of improved performance and decreased cognitive load (i. e., lower "cognitive costs") suggest an increase in overall efficiency due to the training. However, the hypothesis that incomplete worked-out examples and completion problems would lead to a better performance and lower cognitive load in post-test compared to conventional problems was not supported by the results. Specifically, no differences were found between the three conditions in terms of post-test performance and cognitive load experienced in the post-test. This might be a consequence of either the type of instructional support (i. e., a sub-goals oriented support), the type of worked-out examples (i. e., incomplete worked-out examples) or the number of training problems (i. e., only five training problems) provided in the current study. Which of these explanations is more plausible is still an open question and should be addressed in future research.

The main focus of the current study was to reveal how the various types of instructional support influence students' viewing behaviors (i. e., eye fixations) and their learning performance. We hypothesized that the condition differences in terms of eye fixations would be significantly distinct across the five genetics problems solved by students in training. In line with our expectations, the results indicated significant differences in terms of number and duration of fixations not only between the three conditions, but also between these conditions across the five training problems (i. e., practice trials). More specifically, the number of fixations on incomplete worked-out examples was significantly lower than on completion and conventional problems, with no significant differences between the last two conditions. Furthermore, the number of fixations on completion problems was significantly lower for the first and fifth trials compared to conventional problems, whereas for the other training problems (i. e., trials 2, 3 and 4) the reverse was true, that is, the number of fixations was significantly higher for completion problems compared to conventional problems.

This pattern of results was also found for the fixation durations, with participants in the incomplete worked-out examples condition having significantly shorter fixation durations than participants in both the completion and conventional problems conditions. No significant differences in terms of fixation durations were found between the completion and conventional problems. Moreover, the fixation durations on first and fifth training problems were significantly shorter in the completion problems condition compared to conventional problems condition, while the fixation durations on the other training trials (i. e., problems 2, 3 and 4) were significantly longer in the completion problems condition compared to the conventional problems condition.

These findings indicate that the type of instructional support differently impacts students' allocation of visual attention. Specifically, omitting a few solution steps increased the task involvement as reflected by the higher number of fixations and the longer fixation durations in the completion problems condition compared to the conventional problems condition, but only for practice trials 2, 3 and 4. The fact that the number and duration of fixations was higher for conventional problems compared to completion problems at the beginning and the end of the training phase (i. e., problems 1 and 5) may indicate that these problems require more processing, especially in the conventional problems condition. In other words, the processing for integrating and solving the omitted solution steps of problems 1 and 5 (i. e., completion problems) required less fixation time than the processing needed to build a problem solution plan without any support (i. e., conventional problems; cf., Hegarty et al., 1992). Despite the differences in eye movement parameters (i. e., number and duration of eye fixations) between the CMP and CVP conditions across practice trials, the post-test performance attained in all three experimental conditions was the same. One possible explanation for the lack of differences between conditions in terms of post-test performance might be due to the expertise reversal effect (Kalyuga, 2007). More specifically, the differences in eye movement parameters indicate various aspects depending on students' expertise level: longer fixation durations might indicate productive involvement during problem solving for high prior knowledge students, whereas for low prior students it could suggest unproductive processing (see Schwonke, Berthold & Renkl, 2009).

Furthermore, the eye fixations analysis allows us to identify when students encounter difficulties during the problem-solving process (Chuang & Liu, 2012). In general, longer fixation durations indicate that students have encountered more difficulties in solving the problems, situation in which they do not know what to do next, and start either to stare at the problem for longer or search back and forth between the problem statement and the solution steps (cf., Knoblich et al., 2005).

Finally, it was hypothesized that there would be a positive correlation between on-line processing and off-line measures of learning in training. The results confirmed this hypothesis, indicating a positive correlation between the number and duration of fixations on training problems, on one hand, and the performance on these problems (except for trials1 and 2) and cognitive load experienced while solving these problems (with the exception of trials 4 and 5 for which no significant correlation with invested mental effort were found), on the other hand. The more often students attended to the problems and the longer the time they spent processing those problems, the greater the obtained performance, and the higher the experienced cognitive load (i. e., perceived difficulty and mental effort). The finding that the higher processing time of training problems (as revealed by eye tracking data) the better the training performance and the deeper the processing of these problems (e. g., more invested mental effort) attests that indices of visual attention are indeed strongly related to the strategic processing during problem solving. However, given the correlational nature of the data, it is not possible to establish any causal relationship between students' viewing behavior and the problem-solving processes, an aspect that should be explored by future studies.

In summary, the present study provides three important results. First, the primary findings of this study are that the type of instructional support students received affects differently the pattern of eye fixations. Second, an important advance in the present study is the fact that learning measures (i. e., training performance) were complemented by measures of on-line cognitive processing during problem solving (i. e., eye tracking data). In this way, a more comprehensive picture of the cognitive processes that underlie problem solving can be provided (cf., Van Gog, Paas, Van Merriënboer & Witte, 2005). The results suggest that learning success while using a CBLE (i. e., higher training performance) is related to the amount of time students spent looking at the training problems (e. g., fixation durations). In addition, the results indicated that students experienced higher cognitive load (e. g., exerted more mental effort) when processing the training problems for a longer time (which is revealed by both higher fixation durations and greater learning time). Third, the in-depth investigation of the way in which students interact with different types of instructional support is very useful for designing and implementing effective CBLEs. This type of research is important given that CBLEs are increasingly used in education and their characteristics may influence self-regulated learning in a different way than traditional learning materials (de Bruin & Van Gog, 2012).

Acknowledgements: This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number SOP HRD/159/1. 5/S/136077.

REFERENCES

Atkinson, R.K., Derry, S.J., Renkl, A. & Wortham, D.W. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research, 70 (2),* 181-214. DOI: 10.3102/00346543070002181

Brünken, R., Plass, J.L. & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38 (1),* 53–61. DOI: 10.1207/S15326985EP3801_7

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87 (1)*,5-17. DOI: 10.037/0022-0663.87.1.5

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22 (4)*, 1020–1031. DOI: 10.1037/0278-7393.22.4.1020

Chuang, H.-H. & Liu, H.-C. (2012). Effects of different multimedia presentations on viewers' information-processing activities measured by eye-tracking technology. *Journal of Science Education and Technology, 21*, 276-286. DOI: 10.1007/s10956-011-9316-1

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112 (1)*, 155-159. DOI: 10.1037/0033-2909.112.1.155

De Bruin, A. & Van Gog, T. (2012). Improving self-monitoring and self-regulation of learning: From cognitive psychology to the classroom. *Learning and Instruction, 22,* 245-298

Field, A. (2005). Discovering statistics using SPSS (2nd ed.). London: Sage

Gegenfurtner, A., Lehtinen, E. & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523-552. DOI 10.1007/s10648-011-9174-7

Graesser, A.C., Lu, S., Olde, B.A., Cooper-Pye, E. & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition*, *33*(7), 1235-1247

Hegarty, M. & Just, M.A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language*, *32*(6), 717-742

Hegarty, M., Mayer, R.E. & Green, C.E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology*, *84*(1), 76

Hegarty, M., Mayer, R.E. & Monk, C.A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, *87*(1), 18

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19 (4)*, 509–539. DOI: 10.1007/s10648-007-9054-3

Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13 (4),* 351–371. DOI: 10.1002/(SICI)1099-0720(199908)13:4<351::AID-ACP589>3.0.CO;2-6

Kalyuga, S., Chandler, P., Tuovinen, J. & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Education & Psychology, 93 (3)*, 579-588. DOI: 10.1037/0022-0663.93.3.579

Knoblich, G., Öllinger, M. & Spivey, M. J. (2005). Tracking the eyes to obtain insight into insight problem solving. In G. Underwood (Ed.), *Cognitive Processes in Eye Guidance* (pp. 355-75).Oxford University Press, USA

Mihalca, L., Mengelkamp, C., Schnotz, W. & Paas, F. (2015). Completion problems can reduce the illusions of understanding in a computer-based learning environment on genetics. *Contemporary Educational Psychology,* DOI: *http://dx.doi.org/DOI: 10.1016/j.cedpsych.2015.01.001*

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 (1),* 63–71. DOI: 10.1207/S15326985EP3801_8

Paas, F. & Van Merriënboer, J.J.G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: a cognitive load approach. *Journal of Educational Psychology, 86 (1)*, 122–133. DOI: 10.1037/0022-0663.86.1.122

Paas, F. & Van Gog, T. (2006). Optimising worked example instruction: different ways to increase germane cognitive load. *Learning and Instruction, 16 (2)*, 87–91. DOI: 10.1016/j.learninstruc.2006.02.004

Peters, M. (2010). Parsing Mathematical Constructs: Results from a Preliminary Eye Tracking Study. In M. Joubert (Ed.), Proceedings of the British Society for Research into Learning Mathematics, *30*(2). British Society for Research into Learning Mathematics

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372

Renkl, A. (2002). Worked-out examples: instructional explanations support learning by self-explanations. *Learning and Instruction, 12 (5)*, 529-556. DOI: 10.1016/S0959-4752(01)00030-5

Renkl, A. & Atkinson, R.K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38 (1),* 15-22. DOI: 10.1207/S15326985EP3801_3

Renkl, A., Hilbert, T. & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychological Review, 21*, 67-78. DOI: 10.1007/s10648-008-9093-4

Schnotz, W. (2010). Reanalyzing the expertise reversal effect. *Instructional Science, 38 (3),* 315-323. DOI: 10.1007/s11251-009-9104-y

Schwonke, R., Berthold, K. & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. Applied Cognitive Psychology, 23, 1227-1243

Stark, R. (1999). Lernen mit Lösungsbeispielen. Der Einfluß unvollständiger Lösungsschritte auf Beispielelaboration, Motivation und Lernerfolg [Learning by worked-out examples. The impact of incomplete solution steps on example elaboration, motivation, and learning outcomes]. Bern, Germany: Huber

Sweller, J., Van Merriënboer, J.J.G. & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10(3),* 251-296. DOI: 10.1023/A:1022193728205

Tai, R.H., Loehr, J.F. & Brigham, F.J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education*, *29*(2), 185-208

Tsai, M.J., Hou, H.T., Lai, M.L., Liu, W.Y. & Yang, F.Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. *Computers & Education*, *58*(1), 375-385

Van Gog, T., Kester, L. & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36 (3),* 212-218. DOI: 10.1016/j.cedpsych.2010.10.004

Van Gog, T., Kester, L., Nievelstein, F., Giesbers, B. & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior*, *25*(2), 325-331.

Van Gog, T., Paas, F. & Van Merriënboer, J.J.G. (2004). Process-oriented worked examples: Improving transfer performance through enhanced understanding. *Instructional Science, 32,* 83-98. DOI: 10.1023/B:TRUC.0000021810.70784.b0

Van Gog, T., Paas, F. & Merriënboer, J.J.G. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction, 18 (3)*, 211-222. DOI: 10.1016/ j.learninstruc.2007.03.003

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

Van Gog, T. & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychological Review, 22 (2)*, 155-174. DOI: 10.1007/s10648-010-9134-7

Van Merriënboer, J.J.G. & de Croock, M.B.M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research,8,* 365-394. DOI: 10.2190/MJDX-9PP4-KFMT-09PM

Van Merriënboer, J.J.G., Schuurman, J.G., de Croock, M.B.M. & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training. *Learning and Instruction, 12*, 11–37. DOI: 10.1016/S0959-4752(01)00020-2

Van Merriënboer, J.J.G. & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review, 17 (2),* 147-177. DOI: 10.1007/s10648-005-3951-0