Learning and Instruction 22 (2012) 281–289

Contents lists available at SciVerse ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Is adaptation to task complexity really beneficial for performance?

Stephanie Pieschl [a,*], Elmar Stahl [b], Tom Murray [c], Rainer Bromme [a]

[a] University of Muenster, Institute of Psychology, Fliednerstr. 21, 48149 Muenster, Germany
[b] University of Education, Kunzenweg 21, 79117 Freiburg, Germany
[c] University of Massachusetts, Amherst, MA and Perspegrity Solutions, 252 Chesterfield Rd, Westhampton, MA, USA

ARTICLE INFO

Article history:
Received 28 January 2011
Received in revised form
27 August 2011
Accepted 30 August 2011

Keywords:
Self-regulated learning
Metacognition
Adaptation
Hypertext
Performance

ABSTRACT

Theories of self-regulated learning assume that learners flexibly adapt their learning process to external task demands and that this is positively related to performance. In this study, university students ($n = 119$) solved three tasks that greatly differed in complexity. Their learning processes were captured in detail by task-specific questionnaires and computer-generated log files. Results indicate that students adapted almost all learning processes significantly to task complexity. For example, students accessed more hypertext pages for complex tasks than for simple tasks. However, this kind of adaptation was not consistently related to performance. For variables capturing learners' self-regulation, such as the number of accessed hypertext pages, more pronounced adaptation was significantly and positively related to performance even when learners' general processing depth was statistically controlled. Results were less consistent for variables capturing learners' self-monitoring, such as their judged task complexity.

1. Introduction

Adaptation is a central mechanism proposed by most theories of self-regulated learning (Pintrich, 2004; Winne & Hadwin, 1998; Zimmerman, 2002). Based on processes of self-monitoring and self-control (Nelson & Narens, 1994), skillful self-regulated learners are assumed to adapt to a multitude of internal and external cues to optimize their learning processes (Hadwin, Winne, Stockley, Nesbit, & Woszczyna, 2001). Accordingly, adaptation is also hypothesized to be associated with superior performance. Given the pervasive nature and importance of these basic assumptions, the lack of studies explicitly investigating these issues is surprising, especially studies conducted in relatively authentic learning settings. The current study will explore these questions focusing on learners' adaptation to task complexity as an illustrative example.

1.1. Metacognitive self-monitoring, self-regulation, and adaptation

Models of self-regulated learning conceptualize adaptation on different levels of granularity. Fine-grained adaptation could refer to adapting study strategies within the enactment stage of learning. Large-grain adaptation, on the other hand, could have a forward reaching nature affecting learners' more permanent trait-like approaches to learning, such as their learning style. In this paper, we focus exclusively on adaptation on a moderate level of granularity that is consistent with one specific kind of adaptation suggested by the COPES model of studying (Winne & Hadwin, 1998): Adaptation that addresses coordinating activities across several stages of studying and results in large scale adjustments of learning. One example of this kind of adaptation is learners' between-task adaptation to task complexity. In the current study we investigate this phenomenon by confronting learners with three tasks of differing complexity and by investigating their between-task changes in their learning processes.

Independent of the assumed level of granularity, all conceptualizations of adaptation propose that it is based on learners' self-monitoring and learners' self-regulation. Metacognitive self-monitoring implies bottom-up information processing; learners monitor object-level information such as their own knowledge, their cognitive operations, their learning processes, or their products of learning (Nelson & Narens, 1994). The COPES model posits that during monitoring learners compare these perceptions with their internal standards for learning and that this comparison results in evaluations (Winne & Hadwin, 1998). Learners' internal standards can be influenced by monitoring processes: Within the first stages of learning, learners perceive potential constraints, available resources, and the given goal of the task by monitoring external conditions such as task complexity. As a result, learners

* Corresponding author. Tel.: +49 251 8331386; fax: +49 251 8339105.
E-mail addresses: pieschl@uni-muenster.de (S. Pieschl), elmar.stahl@ph-freiburg.de (E. Stahl), tmurray@cs.umass.edu (T. Murray), bromme@uni-muenster.de (R. Bromme).

0959-4752/$ – see front matter
doi:10.1016/j.learninstruc.2011.08.005

generate a comprehensive task perception and subsequently select or generate idiosyncratic goals which translate into internal standards for subsequent metacognitive self-monitoring. In the current study, we use learners' answers on task-specific questionnaires as indicators of monitoring.

Metacognitive self-control or self-regulation, on the other hand, implies that information flows top-down. Based on metacognitive monitoring and corresponding evaluations, learners actively regulate object-level activities such as their cognitive operations or learning processes (Nelson & Narens, 1994). The COPES model (Winne & Hadwin, 1998) suggests that if a significant discrepancy is detected between learners' internal standards and their actual learning process or product due to metacognitive monitoring, then metacognitive control is exerted. According to the model, metacognitive self-control can take two forms: With toggling, a learner turns cognitive operations on or off; editing, on the other hand, implies a change of cognitive operations. In the current study, we use learners' task-specific actions within a hypertext learning environment as indicators of their self-regulation.

1.2. Task complexity according to Bloom's revised taxonomy

We use Bloom's revised taxonomy of educational objectives (Anderson et al., 2001) as a conceptual framework for task complexity. This taxonomy proposes six levels of cognitive processes of ascending complexity: (1) remember, (2) understand, (3) apply, (4) analyze, (5) evaluate, and (6) create. Note that this kind of task complexity is often confused with task difficulty even though these dimensions are not necessarily related. In the current study we focus exclusively on task complexity. Given that this study is the first exploratory study that assesses between-task adaptation to task complexity, we intentionally selected only two types of tasks that required participants to either remember or to evaluate information. We deliberately excluded the most complex tasks requiring the creation of innovative solutions, because we doubted that laypersons in our study would be able to solve such tasks.

According to Bloom's revised taxonomy, remembering involves "retrieving relevant knowledge from long-term memory" (Mayer, 2002, p. 228). Corresponding simple tasks can either be recognition tasks for which learners must determine whether presented material is consistent with the information stored in long-term memory (e.g., multiple-choice questions) or free recall tasks (e.g., in an open-answer format). On the other hand, evaluating involves "making judgments based on criteria and standards" (Mayer, 2002, p. 230). These criteria could be externally provided or self-generated; most often they involve effectiveness or efficiency. Corresponding complex tasks either involve checking, detecting inconsistencies, or critiquing or judging based on a criterion. Most often these tasks have an open-answer format.

To avoid two potential problems, we sequenced our three target tasks as follows: a simple remember task in a multiple-choice format (Task A), a complex evaluation task in an open-answer format (Task B), and simple remember task in an open-answer format (Task C). First, one could argue that learners automatically enhance their processing depth as they become more familiar with the learning setting, with the content matter, and with our hypertext learning environment. To anticipate this issue, we implemented a task sequence that required participants to increase their processing depth (from Task A to Task B), as well as the reverse process of decreasing their processing depth (from Task B to Task C). Second, one could argue that task complexity and task format were confounded and that learners adapted to the task format instead of to task complexity. To anticipate this issue, we designed the simple and complex tasks using the same format (i.e., Task B and Task C have the same open-answer format). For reasons of ecological validity, we also included a multiple-choice remember task (Task A); we predicted no significant learning process differences between the simple remember Tasks A and C.

1.3. Empirical evidence of between-task adaptation to task complexity

Empirical studies often do not make a clear distinction between task complexity and task difficulty. Moreover, operational definitions of task complexity vary, often by only calculating the number of steps involved to solve a task and not taking into account the complexity of the required cognitive operations (Anderson et al., 2001). Despite these shortcomings, some findings are noteworthy.

Empirical studies from diverse contexts demonstrate that learners, in fact, adapt their learning process to task complexity in line with theoretical assumptions. For example, learners plan superficial strategies for simple tasks, and they plan deep elaborative strategies for complex tasks (Stahl, Pieschl, & Bromme, 2006). Learners also access fewer hypertext pages to solve simple tasks than to solve complex tasks (Pieschl, Bromme, Porsch, & Stahl, 2008), request less information to solve simple decision tasks than to solve complex decision tasks (Klayman, 1985), display less and shorter episodes of socially shared metacognition regarding easy problems than regarding difficult problems (Iiskala, Vauras, Lehtinen, & Salonen, 2011), show less metacognition while studying easy texts than while studying difficult texts (Veenman & Beishuizen, 2004), and acquire more factual and conceptual knowledge when solving complex tasks than when solving simple tasks (Gall, 2006).

Notwithstanding, evidence suggests that these adaptations may be insufficient for the learner. That is, learners might not execute sufficiently deep learning strategies regarding very complex tasks (Bromme, Pieschl, & Stahl, 2010). While learners display good metacognition (Boekaerts & Rozendaal, 2010; Veenman & Elshout, 1999) and good learning processes (Klayman, 1985; Winne & Jamieson-Noel, 2003) for relatively simple tasks, for more complex tasks, learners display inadequate metacognitive skills (Boekaerts & Rozendaal, 2010; Veenman & Elshout, 1999), enact fewer learning strategies (Winne & Jamieson-Noel, 2003), and do not request enough information (Klayman, 1985).

Additionally, only relations between the execution of learning strategies or general processing depth and performance have been demonstrated empirically (e.g., Gall, 2006), whereas no consistently positive relation between learners' adaptation and their performance has been reported to date (Pieschl et al., 2008).

1.4. Research questions

In the current study, we confronted learners with three consecutive tasks of different complexity: A simple remember Task A, a complex evaluation Task B, and another simple remember Task C. They were asked to solve these tasks with the help of a hypertext learning environment. To capture their between-task adaptation to task complexity, we investigated their learning processes in detail. We administered task-specific post-hoc questionnaires as indicators of learners' monitoring and we collected task-specific log files about their actions in the hypertext learning environment as indicators of their self-regulation. We investigated the following two research questions:

Do learners adapt their learning to task complexity (Research Question 1)? We hypothesized that learners will demonstrate significant between-task adaptation to task complexity by systematically varying their learning processes between tasks. More specifically, we assume that they would demonstrate significantly enhanced processing depth for the complex Task B in comparison with the simple Tasks A and C. We predicted this effect for indicators of their monitoring (questionnaire) as well as for indicators of their self-regulation (log files). To test this effect, we employed two methods: First, we tested whether there are significant effects of the within-subject variable task complexity in repeated-measure analyses for the learning process variables. Second, we computed adaptation scores that represent the magnitude of between-task differences in values for each learning process variable and determined the significance of these scores.

Is adaptation to task complexity beneficial for task performance (Research Question 2)? We hypothesized that more pronounced between-task adaptation to task complexity is positively related to superior performance. In this study, we operationalized students' performance in two ways: Correctness indicates how many of the three target tasks were solved correctly, whereas the performance score incorporates qualitative scores regarding the open answers of tasks B and C. To test the relation between learners' adaptation and these two scores, we employed two methods: First, we correlated students' adaptation scores of each learning process variable (see research question 1) with the correctness and performance scores. Second, as a more rigorous test, we additionally controlled for learners' general processing depth by computing partial correlation coefficients.

Table 1
Tasks A, B, and C; correct answers in italics.

Title	Task
Task A	Which steps are not parts of the mtDNA analysis? ○ Determining the exact sequence of the hypervariable regions. ○ Visual/microscopic analysis of the material. ○ *Determining a band pattern with gel electrophoresis.* ○ Extraction and purification of mtDNA. ○ Multiplication of mtDNA via PCR.
Task B	Imagine that you study biology. Your professor also handles consultation about genetic fingerprinting. In this role he often receives requests about the suitability of DNA analysis methods for paternity testing. He reports that many laypersons ask for Y-STR or STR analysis. As part of your scientific term paper it is your task to discuss these two DNA analysis methods in writing regarding their suitability for paternity testing and regarding the certainty and informative value of their results. One result of this discussion should be the recommendation of one method. *Correct recommendation: STR analysis.*
Task C	Imagine your family is into genealogy. For this purpose they had mtDNA profiles of all family members made. The molecular biology institute that did these analyses also sent a report. Within this report, the experts refer to "matches" between many family members. You don't know what this technical term means. Therefore, you investigate this issue. Subsequently, you explain the term "match" to your family as follows: *Correct answer: "Match" refers to the same mtDNA sequence.*

2. Method

2.1. Participants

Participants of this study were 129 university students. For the purpose of this paper, we only analyzed the data of those students who answered all three target tasks ($n = 119$; 33 males, 86 females). These students had a mean age of 23.91 years ($SD = 4.51$) and, on average, studied in the 5th semester ($M = 5.08$, $SD = 3.24$) psychology ($n = 47$), other humanities ($n = 34$), sciences ($n = 17$), or other majors ($n = 21$). Students from different majors did not differ significantly in age, semester, (self-rated) prior knowledge, and computer and internet use. Therefore, all subsequent results will be reported for the whole sample. These students were representative in their knowledge about the topic of our study, genetic fingerprinting, and in their computer and internet use which is relevant because we used a hypertext learning environment. More specifically, they rated their prior domain knowledge in molecular biology to be low ($M = 2.16$, $SD = .88$, on a 5-point scale from 1 = very low to 5 = very high), and this was confirmed by their results of a short knowledge test about molecular biology ($M = 3.09$, $SD = 1.75$; with scores between 0 and 8 items correct). Furthermore, students used computers ($M = 17{:}37$ h/week, $SD = 14{:}91$) and the internet ($M = 12{:}87$ h/week, $SD = 12{:}02$) extensively.

2.2. Materials

2.2.1. Three target tasks and measures of performance

On average, students completed $M = 5.87$ ($SD = 1.86$) tasks. In our analyses, we focused exclusively on the three consecutive target tasks. As can be seen in Table 1, Task A was a simple remember task with a multiple-choice format, task B was a complex evaluation task in an open-answer format, and Task C was another simple remember task in an open-answer format.

We analyzed two indicators of performance on different levels of granularity: Students' answers on all three target tasks were scored as either correct or incorrect (0–1 point). For example, in Task B the method of STR analysis is the only method suitable for paternity testing and therefore this recommendation was categorized as correct (see Table 1). These three scores were summed for a score of overall correctness. Additionally, we evaluated all open answers with qualitative task-specific rubrics. Answers regarding Task B received a maximum of 13 points for mentioning and describing different DNA analysis methods (0–4 points), for putting forth pro and contra arguments for these methods (0–6 points), for drawing a comprehensive conclusion (0–2 points), and for giving the correct recommendation (0–1 points). Open answers regarding Task C received a maximum of 5 points for a thorough description (0–2 points), for putting forth critique (0–2 points), and for giving the correct definition (0–1 points). These two qualitative scores, and the correctness of Task A, were summed for an overall performance score.

2.2.2. Hypertext about genetic fingerprinting and log file variables

Tasks were solved with a hypertext about genetic fingerprinting that was created with MetaLinks (Murray, 2003) and comprised 106 pages of content. This content was mainly structured in three chapters describing different DNA analysis methods (mtDNA analysis, STR analysis, and Y-STR analysis; see Fig. 1). All pages contained written text and most included pictures or tables (see Fig. 2). These hypertext nodes were primarily linked in a hierarchical structure offering learning material on different levels of complexity. Hierarchical navigation in this hypertext structure was facilitated by a family tree metaphor: More detailed information could be accessed via the "child" command, simpler content could be accessed via the "parent" command, and subsequent or previous pages on the same hierarchical level could be accessed via the commands "next sibling" or "previous sibling." Additionally, users could jump from any hypertext page to any other hypertext page by using a number of advanced navigational features: For example, users could search for any term and directly access any page by clicking on the result list, or users could use the hypertexts' table of content (TOC) to access any page directly.

Time-stamped log files of all student actions were automatically collected as indicators of their task-specific self-regulation of their learning process. We analyzed students' *Time for Task Completion* (TTC) and their *Number of Accessed Nodes* (NAN; see "hypertext pages accessed by this student" in Fig. 1) as indices of their processing depth with the caveat that we could not clearly distinguish

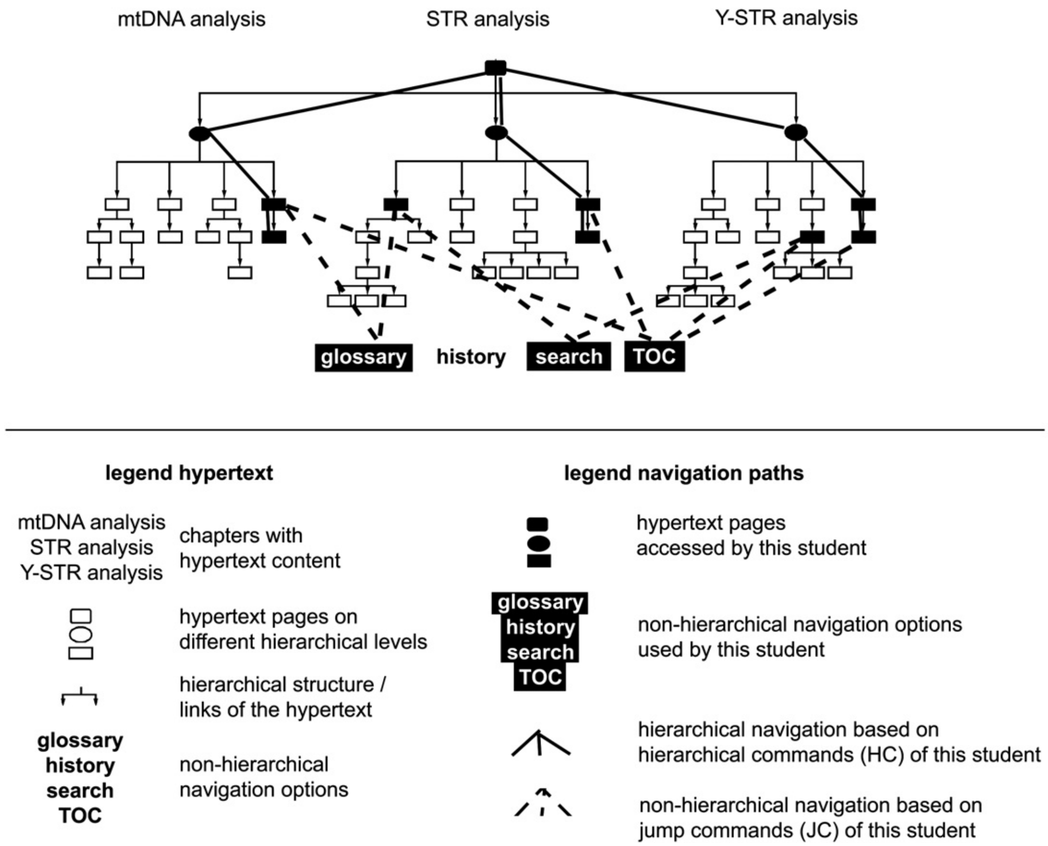


Fig. 1. Visualization of the hierarchical structure of the hypertext learning environment on genetic fingerprinting and visualization of hierarchical (HC) and jump command (JC) navigation of a fictitious student.

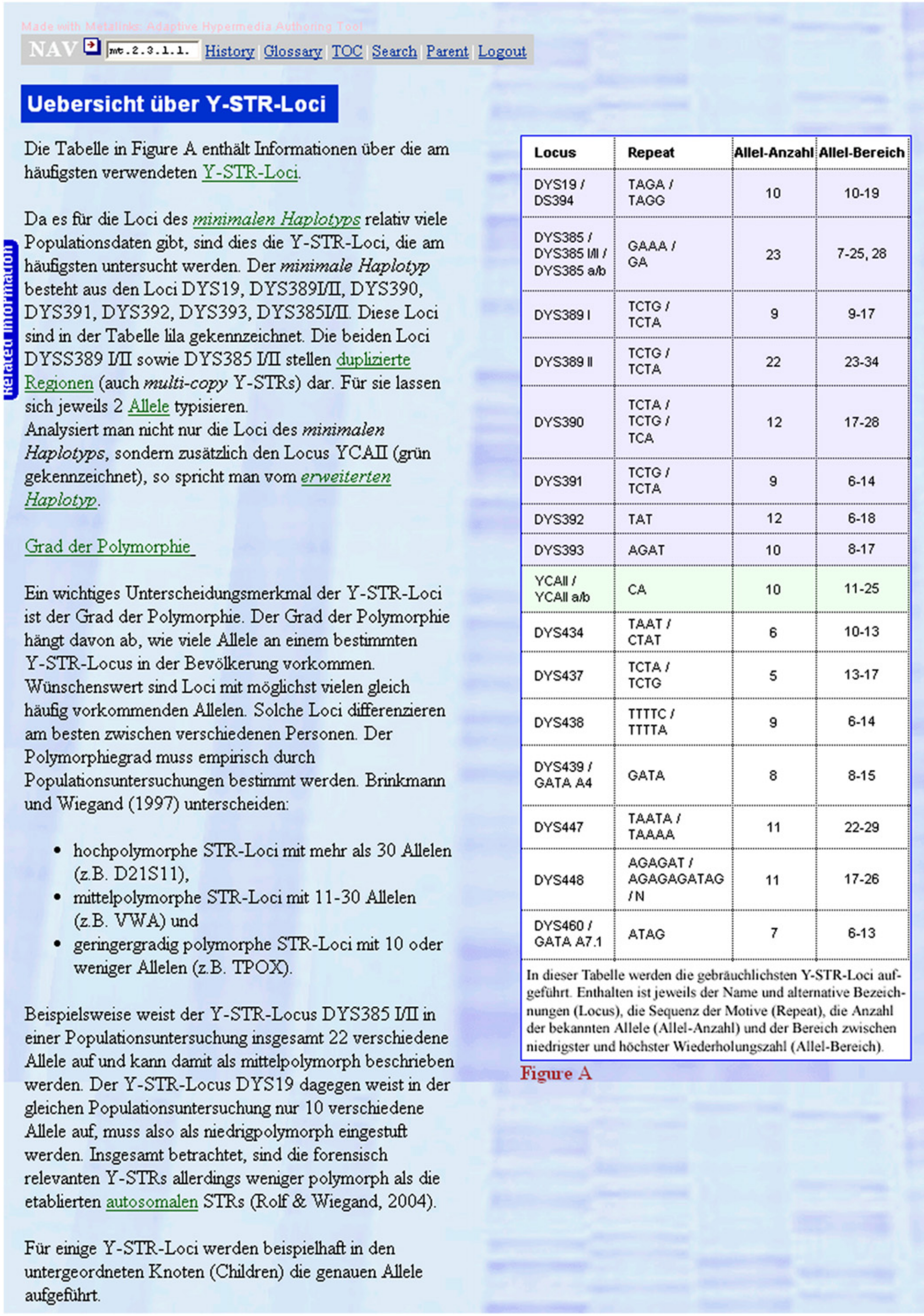
between processing depth and being "lost in hyperspace" or potential other time-consuming phenomena. We analyzed their use of *Hierarchical Commands* (HC; see Fig. 1 top half) as an indicator of how much students follow the given structure of the hypertext, which might also be closely related to their processing depth given that the use of such hierarchical commands often indicates the use of more detailed information. Additionally, we analyzed their use of *Jump Commands* (JC; see Fig. 1 bottom half) as an indicator of how much students purposefully select content, a variable that might be indicative of their orientation within the hypertext.

2.2.3. Task-specific questionnaire

After the task solution phase, students answered task-specific questionnaires (TSQ) about each of the three target Tasks A, B, and C. The TSQ consists of 36 items of different formats and primarily captures students' monitoring. For the purpose of this paper, we will exclusively report the results of three sub-scales. All of the corresponding items had 7-point response scales from 1 = low to 7 = high. *Judged Task Complexity* measured students' subjective judgments about the complexity of the target tasks with 9 items (e.g., "This task was simple (= 1) – complex (= 7)"; Cronbach's $\alpha = .81–.86$ for Tasks A – C). This scale indicates students' monitoring of task complexity and most likely represents their corresponding internal standards. *Processing Efficiency* measured, for example, students' subjective opinion about their own efficiency, their task satisfaction, and their confidence in their own task solution with 5 items (e.g., "My strategies were inefficient (=1) – efficient (= 7)"; Cronbach's $\alpha = .77–.85$ for Tasks A – C). This scale indicates students' self-monitoring and evaluation of their own learning process. *Depth of Processing* measured students' judgments about goals of deep processing and task-appropriate strategies for the target tasks with 10 items (e.g., "The strategy of integrating information across multiple hypertext pages was unimportant (= 1) – important (= 7) for this task"; Cronbach's $\alpha = .69–.87$ for Tasks A – C). This scale also indicates students' monitoring of task complexity and most likely represents their corresponding internal standards.

2.3. Procedure

Students were recruited by a posting at a German university and received 18 € or a certificate of participation as compensation (students of psychology are required to participate in empirical studies at German universities). Data was collected in group sessions with 2–6 students that lasted about 2.5 h. In the first part of this session, students completed a questionnaire about their prior domain knowledge in molecular biology. In the main part of the study, students watched a short standardized video explaining navigation in the hypertext learning environment about genetic fingerprinting. After the video, students solved tasks of different complexity as they interacted with this hypertext. During this task solution process, computer-generated log files were automatically collected. At first, they received a booklet with five tasks that contained the three consecutive target Tasks A, B, and C. Booklets with additional tasks were handed out upon request. This part of the study was self-paced to enhance ecological validity. Students were instructed to solve these tasks "as well as possible." The task solution phase lasted 1 h. Afterward, all students answered the TSQ about the target Tasks A, B, and C. At the end of the study, students



Made with MetaLinks: Adaptive Hypermedia Authoring Tool

NAV mt.2.3.1.1. History Glossary TOC Search Parent Logout

Related Information

Uebersicht über Y-STR-Loci

Die Tabelle in Figure A enthält Informationen über die am häufigsten verwendeten Y-STR-Loci.

Da es für die Loci des *minimalen Haplotyps* relativ viele Populationsdaten gibt, sind dies die Y-STR-Loci, die am häufigsten untersucht werden. Der *minimale Haplotyp* besteht aus den Loci DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385I/II. Diese Loci sind in der Tabelle lila gekennzeichnet. Die beiden Loci DYSS389 I/II sowie DYS385 I/II stellen duplizierte Regionen (auch *multi-copy* Y-STRs) dar. Für sie lassen sich jeweils 2 Allele typisieren.
Analysiert man nicht nur die Loci des *minimalen Haplotyps*, sondern zusätzlich den Locus YCAII (grün gekennzeichnet), so spricht man vom *erweiterten Haplotyp*.

Grad der Polymorphie

Ein wichtiges Unterscheidungsmerkmal der Y-STR-Loci ist der Grad der Polymorphie. Der Grad der Polymorphie hängt davon ab, wie viele Allele an einem bestimmten Y-STR-Locus in der Bevölkerung vorkommen. Wünschenswert sind Loci mit möglichst vielen gleich häufig vorkommenden Allelen. Solche Loci differenzieren am besten zwischen verschiedenen Personen. Der Polymorphiegrad muss empirisch durch Populationsuntersuchungen bestimmt werden. Brinkmann und Wiegand (1997) unterscheiden:

- hochpolymorphe STR-Loci mit mehr als 30 Allelen (z.B. D21S11),
- mittelpolymorphe STR-Loci mit 11-30 Allelen (z.B. VWA) und
- geringergradig polymorphe STR-Loci mit 10 oder weniger Allelen (z.B. TPOX).

Beispielsweise weist der Y-STR-Locus DYS385 I/II in einer Populationsuntersuchung insgesamt 22 verschiedene Allele auf und kann damit als mittelpolymorph beschrieben werden. Der Y-STR-Locus DYS19 dagegen weist in der gleichen Populationsuntersuchung nur 10 verschiedene Allele auf, muss also als niedrigpolymorph eingestuft werden. Insgesamt betrachtet, sind die forensisch relevanten Y-STRs allerdings weniger polymorph als die etablierten autosomalen STRs (Rolf & Wiegand, 2004).

Für einige Y-STR-Loci werden beispielhaft in den untergeordneten Knoten (Children) die genauen Allele aufgeführt.

Locus	Repeat	Allel-Anzahl	Allel-Bereich
DYS19 / DS394	TAGA / TAGG	10	10-19
DYS385 / DYS385 I/II / DYS385 a/b	GAAA / GA	23	7-25, 28
DYS389 I	TCTG / TCTA	9	9-17
DYS389 II	TCTG / TCTA	22	23-34
DYS390	TCTA / TCTG / TCA	12	17-28
DYS391	TCTG / TCTA	9	6-14
DYS392	TAT	12	6-18
DYS393	AGAT	10	8-17
YCAII / YCAII a/b	CA	10	11-25
DYS434	TAAT / CTAT	6	10-13
DYS437	TCTA / TCTG	5	13-17
DYS438	TTTTC / TTTTA	9	6-14
DYS439 / GATA A4	GATA	8	8-15
DYS447	TAATA / TAAAA	11	22-29
DYS448	AGAGAT / AGAGAGATAG / N	11	17-26
DYS460 / GATA A7.1	ATAG	7	6-13

In dieser Tabelle werden die gebräuchlichsten Y-STR-Loci aufgeführt. Enthalten ist jeweils der Name und alternative Bezeichnungen (Locus), die Sequenz der Motive (Repeat), die Anzahl der bekannten Allele (Allel-Anzahl) und der Bereich zwischen niedrigster und höchster Wiederholungszahl (Allel-Bereich).

Figure A

Fig. 2. Part of a page of the hypertext learning environment on genetic fingerprinting with moderately complex content about Y-STR analysis. The navigation bar can be seen in the top of each page; in this sample page further links to subordinate pages are cut off.

completed another battery of questionnaires, for example about their demographics.

3. Results

3.1. Do learners adapt their learning to task complexity?

To answer the first research question (Research Question 1) we employed two methods. First, we tested whether there are significant effects of task complexity by computing repeated-measure MANOVAs across the three target Tasks A, B, and C, one for the three TSQ scales and one for the four log file variables. We hypothesized significant effects of the within-subject variable task complexity and we predicted significant pairwise differences between Tasks A and B and between Tasks B and C, but no significant difference between Tasks A and C. For Task B, students should report greater *Judged Task Complexity*, report less *Processing Efficiency*, report greater *Depth of Processing*, take more time (TTC) and

Table 2
Descriptives of variables by Tasks and pairwise comparisons.

Variable	Task A	Task B	Task C	Pairwise Comparisons	Sum Score	Adaptation Score[b]
Judged Task Complexity[a]	3.40 (1.01)	5.19 (.78)	3.09 (.96)	A < B***, B > C***, A > C**	11.68 (1.76)	3.89 (2.24)***
Processing Efficiency[a]	5.30 (1.18)	4.36 (1.08)	5.22 (1.14)	A > B***, B < C***, A = C ns.	14.89 (2.56)	.33 (2.40)
Depth of Processing[a]	3.29 (1.10)	4.73 (.93)	3.05 (1.42)	A < B***, B > C***, A > C*	11.06 (2.82)	3.12 (1.97)***
TTC	05:18 (02:47)	25:28 (08:37)	07:37 (04:56)	A < B***, B > C***, A < C***	38:18 (08:45)	37:35 (19:41)***
NAN	6.76 (5.71)	24.75 (12.58)	7.37 (6.62)	A < B***, B > C***, A = C ns.	39.03 (17.80)	35.29 (24.61)***
HC	4.45 (4.64)	14.99 (8.98)	3.84 (5.01)	A < B***, B > C***, A = C ns.	23.28 (13.75)	21.48 (17.14)***
JC	.64 (.97)	1.96 (2.38)	1.08 (1.32)	A < B***, B > C***, A < C**	3.72 (3.60)	2.21 (4.26)***
correctness	.88 (.32)	.50 (.50)	.98 (.13)	–	2.37 (.70)	–
performance score	.88 (.32)	7.05 (2.14)	2.41 (1.06)	–	10.34 (2.65)	–

Values in columns – except "pairwise comparisons" – represent means; standard deviations are shown in brackets. The sum score was computed as follows: A + B + C. The adaptation score was computed as follows: (B – A) + (B – C). Abbreviations: TTC = *Time for Task Completion*; NAN = *Number of Accessed Nodes*; HC = (number of) *Hierarchical Commands*; JC = (number of) *Jump Commands*; ns. = not significant.
*$p < .010$, **$p < .005$, ***$p < .001$.
[a] Rated on scales from 1 = low to 7 = high.
[b] We tested if these scores systematically differed from 0.

access more hypertext pages (NAN), and use more *Hierarchical* (HC) and *Jump Commands* (JC) than for the simple Tasks A and C. The task-specific descriptive statistics of all variables are presented in Table 2. The correlations between all variables are presented in the upper half of Table 3.

The MANOVA results for the TSQ scales indicated a significant multivariate main effect of the repeated-measure factor task complexity ($F(6, 470) = 40.76$, $p < .001$, $\eta_p^2 = .34$) that was replicated on all sub-scales: We found significant effects of task complexity for *Judged Task Complexity* ($F(2, 236) = 201.13$, $p < .001$, $\eta_p^2 = .63$), *Processing Efficiency* ($F(2, 236) = 38.17$, $p < .001$, $\eta_p^2 = .24$), and *Depth of Processing* ($F(2, 236) = 138.02$, $p < .001$, $\eta_p^2 = .54$). The results of the more detailed pairwise comparisons are provided in Table 2.

The MANOVA results for the log file variables indicated a significant multivariate main effect of the repeated-measure factor task complexity ($F(8, 456) = 38.62$, $p < .001$, $\eta_p^2 = .40$) that was replicated on all sub-scales. Because the assumption of sphericity was violated we applied the Greenhouse–Geisser correction. We found significant effects of task complexity for TTC ($F(1.40, 160.72) = 344.11$, $p < .001$, $\eta_p^2 = .75$), NAN ($F(1.53, 175.54) = 184.77$, $p < .001$, $\eta_p^2 = .62$), HC ($F(1.55, 177.82) = 139.46$, $p < .001$, $\eta_p^2 = .55$), and JC ($F(1.65, 189.23) = 25.63$, $p < .001$, $\eta_p^2 = .18$). The results of the more detailed pairwise comparisons are displayed in Table 2.

Table 3
Correlations between adaptation scores (top) and with performance scores (bottom).

	1	2	3	4	5	6	7
Adaptation Scores							
1 *Judged Task Complexity*	–	–.19*	.55**	.43**	.25**	.28**	.12
2 *Processing Efficiency*		–	–.02	–.11	–.12	–.14	–.19*
3 *Depth of Processing*			–	.42**	.25**	.27**	.09
4 TTC				–	.56**	.47**	.20*
5 NAN					–	.87**	.14
6 HC						–	.01
7 JC							–
Performance Scores							
correctness	.18*	.12	.12	.15	.32**	.32**	.06
correctness partial[a]	.14	.01	.18	.11	.32***	.34***	.07
performance score	.22*	–.01	.17	.28**	.28**	.30**	–.01
performance score partial[a]	.18	–.12	.25**	.07	.20*	.24*	–.05

Abbreviations: TTC = *Time for Task Completion*; NAN = *Number of Accessed Nodes*; HC = (number of) *Hierarchical Commands*; JC = (number of) *Jump Commands*.
*$p < .010$, **$p < .005$, ***$p < .001$.
[a] We computed partial correlation coefficients controlling for the effects of the respective sum scores, for example for the correlation between TTC and the performance score we partialed out the TTC sum score.

Second, we computed adaptation scores for each learning process variable. These scores indicate the magnitude of the predicted between-task differences and are computed as follows: adaptation score = ((B – A) + (B – C)) (with the letters representing the task-specific values). Therefore, high positive values indicate markedly higher values on the complex target Task B than on the simple target Tasks A and C. The descriptive statistics of these adaptation scores are reported in Table 2. We determined significance by testing the sample means of these scores against zero, because zero would indicate no between-task adaptation to task complexity. Except for *Processing Efficiency* ($t(119) = 1.52$, ns.), significant values were found for adaptation scores on all process variables, *Judged Task Complexity* ($t(118) = 18.93$, $p < .001$), *Depth of Processing* ($t(118) = 17.31$, $p < .001$), TTC ($t(115) = 20.57$, $p < .001$), NAN ($t(115) = 15.45$, $p < .001$), HC ($t(115) = 13.50$, $p < .001$), and JC ($t(115) = 5.58$, $p < .001$).

3.2. Is adaptation to task complexity beneficial for task performance?

To answer the second research question (Research Question 2) we also employed two methods: First, we correlated students' adaptation scores for all learning process variables with the correctness and performance scores. The descriptives of these scores are provided in Table 2. On average, students answered $M = 2.37$ ($SD = .70$) tasks correctly out of a maximum of 3, and they achieved an average performance score of $M = 10.34$ ($SD = 2.65$) out of a maximum of 19. Overall, correctness and performance scores were significantly related ($r = .48$, $p < .001$).

Note that in the upper part of Table 3 the correlations between all learning process variables are displayed, whereas the lower part of Table 3 presents the correlations with the correctness and performance scores. These results indicate that adaptation scores regarding *Judged Task Complexity*, NAN, and HC are significantly and positively correlated with students' correctness. Furthermore, the adaptation scores regarding *Judged Task Complexity*, TTC, NAN, and HC are significantly and positively correlated with students' performance score. All significant correlations are of small to moderate effect size.

Second, as a more rigorous test, we verified these results by additionally controlling for learners' general processing depth regarding the respective variable. For this analysis we computed partial correlation coefficients. For each learning process variable we controlled for students' general processing depth regarding that specific variable, measured by the corresponding sum score. All sum scores are also reported in Table 2 and were computed as follows: sum score = (A + B + C) (with the letters representing the task-specific values).

The results presented in the lower part of Table 3 show that after controlling for the respective general processing depth, the adaptation scores regarding NAN and HC are significantly and positively correlated with students' correctness. Furthermore, after controlling for the respective general processing depth, the adaptation scores regarding *Depth of Processing*, NAN, and HC are significantly and positively correlated with students' performance score. All significant correlations are of small to moderate effect size.

4. Discussion

4.1. Do learners adapt to task complexity?

Students demonstrated significant between-task adaptation to task complexity regarding nearly all learning process variables. We found significant effects of the factor task complexity in all repeated-measure analyses and almost all adaptation scores were statistically significant. Therefore, these findings confirm our first research question (Research Question 1).

Regarding the questionnaire scales, the results were most consistent for the scales *Judged Task Complexity* and *Depth of Processing*: The repeated-measure analyses showed significant effects of task complexity, pairwise comparisons showed that students had significantly higher values regarding the complex evaluation Task B compared to each of the simple remember Tasks A and C, and the mean adaptation scores were significant. We conclude that students successfully monitored task complexity and presumably translated these perceptions into adequate internal standards for further self-monitoring and self-regulation of their learning process. Students also displayed significantly higher values on the multiple-choice remember Task A than on the open-answer remember Task C − an effect we did not predict. We can only speculate that the content of Task A might have appeared more complex to participants. To answer Task A correctly, participants needed to read about the process of conducting an mtDNA analysis in the lab, a highly technical content. To answer Task C correctly, participants read about matching of mtDNA sequences. Although this content is also technical, it was framed in an everyday experience. The results regarding the third scale, *Processing Efficiency*, were less consistent. The repeated-measure analysis showed a significant effect of task complexity, pairwise comparisons showed that students had significantly lower values regarding the complex evaluation Task B compared to each of the simple remember Tasks A and C, but the mean adaptation scores were non-significant. Conservatively, we conclude that students did not consistently adapt their *Processing Efficiency* to task complexity, probably because this scale is not systematically related to processing depth (see upper part of Table 3).

Regarding the log file variables almost all results are consistent for all variables, namely students' *Time for Task Completion* (TTC), *Number of Accessed Nodes* (NAN), number of *Hierarchical Commands* (HC) and number of *Jump Commands* (JC): The repeated-measure analyses showed significant effects of task complexity, pairwise comparisons showed that students had significantly higher values regarding the complex evaluation Task B compared to each of the simple remember Tasks A and C, and the mean adaptation scores were also significant. We conclude that students successfully self-regulated their learning process, presumably based on their perception of task complexity and their self-monitoring of their own learning process. For TTC and JC, students also displayed significantly lower values on the multiple-choice remember Task A than on the open-answer remember task C − an effect we did not predict. At first glance, these effects seem to contradict those of the questionnaire scales that indicated that students perceived Task A to be more complex. However, we speculate that these behavioral effects are not due to the tasks' content but due to the task format. A hand-written recalled answer (Task C: open-answer, free recall) might have taken students longer (TTC) than to recognize and check an option in a multiple-choice task (Task A, recognition). Due to the growing familiarity with the hypertext and the task order, students might have also become more courageous to use advanced navigational features such as *Jump Commands* (JC) later in the study (Task C) in comparison with earlier in the study (Task A).

Even though prior empirical research about solving tasks of differing complexity exists, the results of this study can contribute to the literature. These findings replicate results from strictly controlled experimental research (Luwel, Verschaffel, Onghena, & De Corte, 2003) within a more ecologically valid learning setting and they expand results from the preparatory planning stages of self-regulated learning (Stahl et al., 2006) to the whole learning process. On a more fundamental level, this study tested one of the most basic assumptions of models of self-regulated learning, namely learners' adaptation based on their monitoring and regulation.

4.2. Is adaptation to task complexity beneficial for task performance?

Students' between-task adaptation to task complexity was not consistently and positively associated with their performance. We found significant (partial) correlations between adaptation scores of selected learning process variables and performance. However, for other variables these correlations were inconsistent, and for another group of variables we found no significant associations with performance. Therefore, these findings only partly confirm our second research question (Research Question 2).

Regarding the variables of *Processing Efficiency* and *Jump Commands* (JC), we found no significant correlations between the corresponding adaptation scores and performance. We conclude that between-task adaptation of these self-monitoring and self-regulating variables is unrelated to learners' performance. We assume that this lack of correlation is due to the facts that *Processing Efficiency* is probably not systematically related to processing depth and that the use of *Jump Commands* (JC) might instead represent a growing familiarity with the hypertext navigation (see research question 1).

For the variables of *Judged Task Complexity*, *Depth of Processing* and *Time for Task Completion* (TTC), we found consistently positive but inconsistently significant correlations between the corresponding adaptation scores and performance that seemed to depend on the rigor of the hypothesis testing (correlations vs. partial correlations) and on the performance measure (correctness vs. performance score). We conclude that between-task adaptation of these variables might be beneficial for performance. We assume that these variables are closely related to self-monitoring and self-regulation of processing depth but that these variables might be strongly influenced by other variables such as general processing depth. For example, if the more rigorous partial correlation coefficients are considered, only the adaptation score of *Depth of Processing* would be significantly related to a higher performance score.

We found consistently positive and significant correlations between the corresponding adaptation scores and performance for the variables *Number of Accessed Nodes* (NAN) and the use of *Hierarchical Commands* (HC). We conclude that between-task adaptation of these self-regulation variables is definitely beneficial for performance. The overt action of accessing more information within the hypertext learning environment and of using more hierarchical navigation for more complex tasks than for simpler tasks manifests in superior performance.

Most models of self-regulated learning propose adaptation fueled by metacognitive self-monitoring and self-regulation as a central mechanism of learning that, in turn, should be associated with superior performance. Given the pervasive nature and importance of these basic assumptions, the lack of empirical research regarding these issues is surprising. This study, to our knowledge, is a first exploration of the complex relationship between adaptation and performance in a moderately authentic learning setting. Our results show that a relation between adaptation and performance is not a desirable ideal that is only attainable by the most skillful self-regulated learners but that it is a real-life phenomenon that can be found in authentic settings and with average learners, at least regarding variables that directly capture self-regulation.

4.3. Limitations and implications

The generalizability of our findings is limited in several ways. First, we do not know if similar effects would have been found with other groups of learners (vs. university students), for learning in one's own field of expertise (vs. in a new field), for solving tasks related to other categories of Bloom's revised taxonomy (vs. remember and evaluate) or with different formats (vs. multiple-choice and open-answer), for solving the tasks in a different order (vs. simple − complex − simple), and with adaptation to other external demands (vs. task complexity). For example, Hadwin et al. (2001) found that students significantly adapted their learning tactics to different learning contexts, indicating that not only task complexity is a relevant external demand for adaptation. Second, the COPES model posits that learners derive idiosyncratic goals and plans in the first stage of studying that might not match objective task demands. Therefore, we do not know exactly how students perceived task complexity, how they interpreted the instruction to solve the tasks as well as possible, and which goals and plans they adopted. Related to this issue, students may not have perceived all their options within the complex hypertext learning environment but might have felt hindered by the experimental setting, for example, by the time constraints or by the limited content and options within the learning environment. Because of these issues, students might have enacted less between-task adaptation than they were capable of or than they would have enacted in their natural learning environment. Third, we deliberately decided to create at least a moderately authentic setting without invasive measures to capture students' online cognitive processes. To compensate, we administered post-hoc questionnaires where students' answers might have been influenced by the whole experience. That is, they might not have been able to answer the task-specific questionnaires (TSQ) for Tasks A, B, and C independently. Therefore, the validity of the questionnaire data can be doubted. To counteract this problem, we collected online behavioral data via log files and systematically compared our results from questionnaire and log file data.

Despite these limitations, several theoretical as well as practical implications of the results of this study are noteworthy. On a theoretical level, these results imply that the hypothesized all-pervasive mechanism of adaptation can be found in relatively authentic learning settings, but that learners do not adapt all aspects of their learning processes equally. We assume that for the illustrative case of between-task adaptation to task complexity all learning process variables related to processing depth are systematically adapted − based on accurate self-monitoring and skillful self-regulation − while other variables such as *Processing Efficiency* in this case are not systematically adapted. A similar conclusion can be drawn regarding the relevance of adaptation to performance: We only detected this relationship for few variables representing processing depth.

These results can be generalized to educational practice, but the conclusions are less straightforward as they might appear at first glance. One could conclude that students simply needed to strongly adapt their processing depth to task complexity, which would, for example, imply that they needed to spend significant effort and cognitive resources on complex tasks. Notwithstanding, this type of processing is clearly advocated: Modern instructional design theories revolve around learning with authentic and complex real-life tasks and advocate variable practice to achieve long-term retention and transfer (Van Merrienboer, Kester, & Paas, 2006). In this context, a certain level of task complexity might constitute a beneficial and "desirable difficulty" for learning (Bjork & Linn, 2006). This line of argumentation is also supported by empirical research where students solving tasks from multiple content areas and tasks that required most effort performed best on a transfer test (Eisenberger, 1992). Therefore, we recommend that schools and universities should use a variety of tasks, especially more complex ones (see also Anderson et al., 2001; Lodewyk & Winne, 2005; Perry, Philips, & Dowler, 2004).

We would like to point out, however, a potential caveat of this approach. Authentic and complex real-life tasks might slow down the learning process and might even exceed learners' capacity. Within a cognitive load approach, for example, a complex task's intrinsic load even under ideal instructional conditions (based on an interaction between task complexity and the learners' expertise) might exceed the learners' cognitive capacity and thus result in cognitive overload (Van Merrienboer et al., 2006). Even the concept of "desirable difficulty" implies that task complexity should not exceed a "desirable" level (Bjork & Linn, 2006). In this study, for example, there is some tentative evidence that the complex evaluation Task B might have exceeded the capacity of our layperson participants. Specifically, students showed significantly less *Processing Efficiency* for this task than for the simple tasks and their performance showed that only 50% of them were able to provide the correct recommendation (see Table 2). We have no final solution for this problem but encourage educators to heed the following advice, especially when encountering complex tasks: Students might profit from scaffolding with regard to adequately monitoring task demands, adequately self-monitoring their learning processes, and adequately self-regulating their corresponding learning tactics and strategies (Kramarski & Michalsky, 2010; Nückles, Hübner, & Renkl, 2009). Moreover, educators should clearly communicate task demands to all students. Given that students still might interpret these task demands differently, monitoring students' task understanding on a very specific level is necessary, as well as making sure their understanding matches the educators' understanding. Asking students about their task-specific goals and about the concrete tactics and strategies they plan to execute might also be helpful.

Acknowledgments

This study was funded by the German Research Foundation (DFG). We thank our further collaborator within this project, Torsten Porsch, as well as our student research assistants, Leoni Pytlik and Janina Wolters.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Bjork, R. A., & Linn, M. C. (2006). The science of learning and the learning of science: introducing desirable difficulties. *The APS Observer, 19*(3), Retrieved from http://www.psychologicalscience.org/observer/getArticle.cfm?id=1952.

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*, 372–382. doi:10.1016/j.learninstruc.2009.03.002.

Bromme, R., Pieschl, S., & Stahl, E. (2010). Epistemological beliefs are standards for adaptive learning: a functional theory about epistemological beliefs and metacognition. *Metacognition and Learning, 5*(1), 7–26. doi:10.1007/s11409-009-9053-5.

Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99*(2), 248–267. doi:10.1037/0033-295X.99.2.248.

Gall, J. (2006). Orienting tasks and their impact on learning and attitudes in the use of hypertext. *Journal of Educational Multimedia and Hypermedia, 15*(1), 5–29, Retrieved from http://www.editlib.org/?fuseaction=Reader.ViewFullText&paper_id=5992.

Hadwin, A. F., Winne, P. H., Stockley, D. B., Nesbit, J. C., & Woszczyna, C. (2001). Context moderates students' self-reports about how they study. *Journal of Educational Psychology, 93*(3), 477–487. doi:10.1037/0022-0663.93.3.477.

Iiskala, T., Vauras, M., Lehtinen, E., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction, 21*, 379–393. doi:10.1016/j.learninstruc.2010.05.002.

Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior and Human Decision Processes, 35*, 179–201. doi:10.1016/0749-5978(85)90034-2.

Kramarski, B., & Michalsky, T. (2010). Preparing preservice teachers for self-regulated learning in the context of technological pedagogical content knowledge. *Learning and Instruction, 20*, 434–447. doi:10.1016/j.learninstruc.2009.05.003.

Lodewyk, K. R., & Winne, P. H. (2005). Relations among the structure of learning tasks, achievement, and changes in self-efficacy in secondary students. *Journal of Educational Psychology, 97*(1), 3–12. doi:10.1037/0022-0663.97.1.3.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003). Analysing the adaptiveness of strategy choices using the choice/no-choice method: the case of numerosity judgments. *European Journal of Cognitive Psychology, 15*(4), 511–537. doi:10.1080/09541440244000292.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice, 41*(1), 226–232, Retrieved from http://www.jstor.org/stable/1477407.

Murray, T. (2003). MetaLinks: authoring and affordances for conceptual and narrative flow in adaptive hyperbooks. *International Journal of Artificial Intelligence in Education, 13*, 2–4, Retrieved from http://hal.archives-ouvertes.fr/docs/00/19/73/17/PDF/Murray03.pdf.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition, knowing about knowing* (pp. 1–25). Cambridge: MIT Press.

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction, 19*, 259–271. doi:10.1016/j.learninstruc.2008.05.002.

Perry, N. E., Phillips, L., & Dowler, J. (2004). Examining features of tasks and their potential to promote self-regulated learning. *Teachers College Record, 106*(9), 1854–1878. doi:10.1111/j.1467-9620.2004.00408.x.

Pieschl, S., Bromme, R., Porsch, T., & Stahl, E. (2008). Epistemological sensitization causes deeper elaboration during self-regulated learning. International perspectives in the learning sciences: cre8ting a learning world. In. *Proceedings of the eighth International Conference for the Learning Sciences – ICLS 2008, Vol. 2* (pp. 2-213–2-220). London: Lulu Enterprises, Retrieved from http://portal.acm.org/citation.cfm?id=1599898.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407, Retrieved from http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=967479dc-1816-4601-a662-78af243f234d%40sessionmgr111&vid=43&hid=127.

Stahl, E., Pieschl, S., & Bromme, R. (2006). Task complexity, epistemological beliefs and metacognitive calibration: an exploratory study. *Journal of Educational Computing Research, 35*(4), 319–338. doi:10.2190/1266-0413-387K-7J51.

Van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20*, 343–352. doi:10.1002/acp.1250.

Veenman, M., & Beishuizen, J. J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraints. *Learning and Instruction, 14*, 621–640. doi:10.1016/j.learninstruc.2004.09.004.

Veenman, M., & Elshout, J. J. (1999). Changes in the relationship between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education, XIV*(4), 509–523, Retrieved from http://www.springerlink.com/content/c3l17r053r6mx267/.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates.

Winne, P. H., & Jamieson-Noel, D. (2003). Self-regulating studying by objectives for learning: students' reports compared to a model. *Contemporary Educational Psychology, 28*, 259–276. doi:10.1016/S0361-476X(02)00041-3.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory Into Practice, 41*(2), 64–70, Retrieved from http://www.jstor.org/stable/1477457.