Instr Sci (2010) 38:237–258
DOI 10.1007/s11251-009-9106-9

Expertise reversal effects in writing-to-learn

Matthias Nückles · Sandra Hübner · Sandra Dümer · Alexander Renkl

Received: 30 November 2008 / Accepted: 3 August 2009 / Published online: 27 August 2009

Abstract This article presents two longitudinal studies that investigated expertise reversal effects in journal writing. In Experiment 1, students wrote regular journal entries over a whole term. The experimental group received a combination of cognitive and metacognitive prompts. The control group received no prompts. In the first half of the term, the experimental group applied more cognitive and metacognitive strategies in their journals and showed higher learning outcomes than the control group. Towards the end of the term, the amount of cognitive and metacognitive strategies elicited by the experimental group decreased while the number of cognitive strategies applied by the control group increased. Accordingly, the experimental group lost its superiority on learning outcomes. In order to avoid these negative long-term effects of prompts, a gradual and adaptive fading-out of the prompts was introduced in the experimental group in Experiment 2 while a control group received permanent prompts. The results showed that, over the course of the term, the fading group applied increasingly more cognitive strategies while the control group applied fewer and fewer cognitive strategies. Accordingly, at the end of the term, the permanent prompts group showed substantially lower learning outcomes than the fading group. Together, these results provide evidence for an expertise reversal effect in writing-to-learn. The more the students became skilled in journal writing and internalized the desired strategies, the more the external guidance by prompts became a redundant stimulus that interfered with the students' internal tendency to apply the strategies and, thus, induced extraneous cognitive load. Accordingly, a gradual fading-out of the prompts in line with the learners' growing competencies proved to be effective in mitigating the negative side-effects of the provided instructional support.

Keywords Journal writing · Cognitive and metacognitive strategies · Expertise-reversal effect · Writing-to-learn · Prompts · Fading-procedures · Instructional design

M. Nückles (✉)
Institute of Educational Science, University of Freiburg, Rempartstraße 11, 79085 Freiburg, Germany
e-mail: matthias.nueckles@ezw.uni-freiburg.de

S. Hübner · S. Dümer · A. Renkl
Institute of Psychology, Engelbergerstraße 41, 79085 Freiburg, Germany

Introduction

Typically, lesson and lecture content "evaporates" rather quickly. After students have left the classroom or lecture auditorium, only a few continue to reflect on the learning contents they have just been confronted with. The students rarely organize learning contents in a meaningful and coherent fashion. Seldom do they come up with examples to illustrate abstract concepts. Also, few students routinely monitor their understanding for knowledge gaps and employ remedial strategies that could help close the gaps. Students' failure to apply such beneficial learning activities typically results in a lack of understanding and, thereby, also leads to poor long-term retention. The writing of learning journals (Berthold et al. 2007; McCrindle and Christensen 1995) is a learning method that may help to overcome these shortcomings. A learning journal typically represents a written explication of one's own learning processes and outcomes over an extended period of time (e.g., over a whole term or school year). Learning journals are especially appropriate for follow-up course work. They help students realize the above-mentioned learning activities.

Journal writing has been shown to be effective in improving students' learning across various educational settings and subjects (e.g., Cantrell et al. 2000; Connor-Greene 2000; McCrindle and Christensen 1995). However, there is also evidence that without appropriate instructional support, students do not apply the learning journal method in an optimal way (Nückles et al. 2004). Therefore, we have developed and experimentally tested specific prompts that support the writing of effective learning journals (e.g., Berthold et al. 2007; Nückles et al. 2009; Hübner et al. 2009). In this contribution, we present two experimental longitudinal studies that examined longer-term effects that prompts had on the application of cognitive and metacognitive learning strategies, as well as on students' learning outcomes and motivation for journal writing.

Cognitive and metacognitive strategies in writing learning journals

Typically, a learning journal is not merely a summary of a writer's learning outcomes gained in preceding learning episodes (e.g., a lecture or seminar session). Rather, it is an opportunity to apply beneficial cognitive and metacognitive learning strategies in order to deepen and expand the newly acquired knowledge. Hence, from the perspective of Cognitive Load Theory (Sweller 2005; Sweller et al. 1998), cognitive and metacognitive learning strategies are genuine activities that increase the share of germane load in working memory. Given that, generally, the capacity of working memory is limited (Baddeley 1986; Cowan 2001), germane cognitive load refers to that share of working memory resources that are devoted to the execution of beneficial learning activities, that is, strategies and processes that enable the learner to reach an intended learning goal (see Kalyuga 2007).

Cognitive learning strategies include organization and elaboration strategies. Organization strategies refer to the identification of main ideas and their interrelations, the highlighting of central concepts, and the structuring of contents (Weinstein and Mayer 1986). Thereby, "internal" links that relate relevant aspects of the new material to each other are constructed (Mayer 1984). In other words, the learning contents are organized in a meaningful way. Elaboration strategies help to construct so-called "external" links that relate the new material to the learner's prior knowledge (Mayer 1984). The generation of examples, the use of analogies, and the critical discussion of issues are commonly regarded as prototypical elaboration strategies (Weinstein and Mayer 1986). Such strategies assist the learner in going beyond the given knowledge by creating links between her or his prior knowledge and the new information (Mayer 1984). Following Mayer's (2002) selecting-organizing-integrating theory of active learning, cognitive strategies such as organization and elaboration are at the heart of meaningful learning because they enable the learner to organize learning contents into a coherent structure and integrate new information with existing knowledge, thereby enabling deep understanding and long-term retention (Barnett et al. 1981; McCrindle and Christensen 1995).

Besides cognitive strategies, the writing of learning journals is further supposed to stimulate the application of metacognitive strategies. Metacognition refers to the knowledge and awareness of one's own cognitive processes and the ability to actively control and manage those processes (Efklides and Vauras 1999; Flavell 1976). Learners may use journal writing to consciously acknowledge which aspects of the learning material they have already understood well (*positive monitoring*), or they may identify comprehension difficulties (*negative monitoring*; see Chi et al. 1989). The elicitation of metacognitive strategies during the production of a learning journal can help to prevent illusions of understanding (Chi et al. 1989; Renkl 1999) and trigger remedial cognitive strategies in order to clarify a previously identified comprehension problem.

To stimulate the use of cognitive and metacognitive strategies in writing learning journals, we developed different sets of prompts. Prompts are questions or hints that induce productive learning processes in order to overcome superficial processing (King 1992; Pressley et al. 1992). They can be conceived of as *strategy activators* (Reigeluth and Stein 1983) because they induce learning strategies that the learners, in principle, are capable of, but do not spontaneously demonstrate, or demonstrate to an unsatisfactory degree. Thus, we assumed that strategies, such as organization and elaboration strategies, would already be part of the students' strategic repertoire. Nevertheless, it would be necessary and useful to prompt students to apply these strategies when writing a learning journal as follow-up course work (Nückles et al. 2004). For example, students might not be aware that it can be productive follow-up course work to generate own examples in order self-explain a newly learned concept (lack of meta-knowledge), and they might not be familiar with applying such strategies in the context of writing a learning journal (lack of practice). In the literature on strategy development, such a strategic inadequacy is typically called a production deficiency (Brown 1978; Flavell 1978; see also Hübner et al. 2009).

Several experimental studies investigated the effects of prompts on strategy use and learning outcomes. In a study by Berthold et al. (2007), students received one of four instructions for writing a journal entry—a so-called "learning protocol"—about a video-taped lecture they had previously viewed. The instruction either included six cognitive (i.e., organizational and elaborative) prompts (e.g., "How can you best structure the learning contents in a meaningful way?"), six metacognitive prompts (e.g., "Which main points haven't I understood yet?"), a mixture of three cognitive and three metacognitive prompts, or no prompts at all (control condition). Results showed that learners who received cognitive, or cognitive *and* metacognitive prompts significantly outperformed the control group with regard to (a) the amount of cognitive and metacognitive strategies in the learning protocols, and (b) the learning outcomes on both an immediate comprehension test and a 7-days delayed retention test. Nückles et al. (2009) successfully replicated the results of Berthold et al. (2007) using an improved and expanded experimental design. They further demonstrated that prompting the application of remedial cognitive strategies in order to address comprehension problems previously identified by self-monitoring, had an added value regarding learning outcomes. Accordingly, in the Nückles et al. (2009) study, the highest learning outcomes were achieved in the experimental condition where students received prompts for (1) the organization and elaboration of learning contents, (2) the monitoring of their understanding, and (3) the application of remedial strategies in case of perceived comprehension problems.

Another experimental study by Schwonke et al. (2006) provided evidence that it might be beneficial to adapt the prompts to the individual strategic habits of the students. In their study, Schwonke and colleagues provided students either with prompts that encouraged strategies the students would spontaneously apply rather seldom, or the students received a random selection of prompts (taken from the same "prompt pool" as in the adaptation condition). Results showed that selecting the prompts in accordance with the preferred repertoire of the student was more beneficial to learning outcomes than presenting a random selection of prompts. These results suggest that it may be important to adapt the provided instructional guidance to the individual competence level of the students (see also McNeill et al. 2006; Puntambekar and Hübscher 2005).

Together, the reported experimental studies suggest that prompts can be a very effective instructional means to support the application of beneficial cognitive and metacognitive learning strategies in writing a learning journal. It should be noted that in all experiments reported above time on task was held constant. Thus, the superiority of prompting groups regarding learning outcomes did not simply result because they spent more time writing their journals as compared with the control groups. Nevertheless, in all these experimental studies, students were required to produce only a single journal entry, that is, a so-called learning protocol (see Berthold et al. 2007; Nückles et al. 2009; Schwonke et al. 2006). In real world instructional settings, however, such as university seminars or lectures, students typically do not produce just one single learning protocol or journal entry. Rather, they are required to write journal entries regularly over a longer period of time, for example, as follow-up course work over a whole term. Thus, the question arises, whether prompts will be effective in stimulating productive learning strategies not only for a short time but also in the longer term.

In the following sections, we will present two experimental longitudinal studies in which we tested the longer-term effects of different prompting procedures on strategy use, learning outcomes, and the students' motivation for journal writing. In Experiment 1, we obtained empirical evidence that prompting students permanently with the same set of prompts over a longer period of time—although initially productive—in the long run, may entail negative side-effects that can theoretically be explained in terms of an expertise reversal effect (Kalyuga 2005, 2007; Kalyuga et al. 2003). In Experiment 2, we tested whether a gradual and adaptive fading-out of the prompts might be appropriate to mitigate this expertise reversal effect.

Experiment 1: supporting journal writing by cognitive and metacognitive prompts

To investigate longer-term effects of cognitive and metacognitive prompts on strategy use and learning outcomes, we conducted an experimental longitudinal study. In this study, we were also interested in the students' motivation for writing a learning journal. Inasmuch as we conceive prompts as "strategy activators" (see Reigeluth and Stein 1983), we are curious whether the repeated presentation of the prompts would foster the internalization of the use of the prompted strategies and thus contribute to the students' motivation for journal writing. On the other hand, as prompts are an external obligation to be fulfilled by the student, it cannot be ruled out that the permanent presentation of the prompts may prove to be rather corruptive to intrinsic motivation in the long run (see Deci et al. 1994, 1999).

In our study, undergraduate students of psychology kept a learning journal as follow-up course work for an introductory course in developmental psychology. They wrote a journal entry about each weekly seminar session over the whole term. The experimental group received a combination of cognitive and metacognitive prompts that had proved to be most effective in previous research (cf. Berthold et al. 2007; Nückles et al. 2009). These students were compared with a control group that received an instruction that contained no prompts at all. We addressed the following research questions:

1. Will cognitive and metacognitive prompts foster the application of cognitive and metacognitive strategies in the learning journals not only in the short term but also in the longer term?
2. Will cognitive and metacognitive prompts foster learning outcomes not only in the short term but also in the longer term?
3. How will the prompts affect the students' motivation for journal writing in the longer term?

Method

Sample and design

Fifty first semester students of Psychology (34 females, 16 males, mean age: 21.74 years) participated in the study. They were randomly assigned to two parallel introductory courses in Developmental Psychology. The courses lasted 4 months. Within this time, 14 seminar sessions were held. Except the first and the last session, the students were required to write a learning journal entry after each session as follow-up course work. Hence, the students had to write twelve journal entries in total. In both courses, the same contents were taught by the same lecturer (i.e., the first author). The students completed these courses as regular part of their undergraduate studies in Psychology at the University of Freiburg. They received 20 Euro for their participation in the testing sessions. To investigate the long-term effects of prompts, we used a control-group design. The participants in one course (i.e., the experimental group, $n = 25$) received prompts for writing their journal entries, whereas the participants in another course (i.e., the control group, $n = 25$) received no prompts. Dependent variables encompassed measures of the learning strategies elicited in the learning journals, the students' learning outcomes as well as measures of the students' motivation for writing the learning journal.

Materials, codings, and instruments

Instructions for writing learning journals (*experimental variation*). In both conditions, a brief general instruction on writing a learning journal was given. The participants in the experimental condition additionally received six prompts, that is, three cognitive and three metacognitive prompts. The cognitive prompts were intended to stimulate organization strategies ("Which were the main points of today's seminar session in your opinion?") and elaboration strategies ("What examples can you think of that illustrate, confirm, or conflict with the learning contents?"—"Which cross-links can you construct between today's seminar session and the previous sessions?"). We further applied two types of metacognitive prompts: Monitoring prompts were meant to elicit monitoring strategies ("Which main points haven't I understood yet?"—"Which main points have I already understood well?"). A planning-of-remedial-strategies prompt encouraged the students to consider ways of regulating their learning process by applying remedial cognitive strategies ("What possibilities do I now have to overcome my comprehension problem?").

Analysis of the learning journals. Two independent raters, who were blind to the experimental conditions, scored the amount of cognitive and metacognitive learning strategies in the journal entries by using 6-point rating scales ranging from 1 (= *dimension not present*) to 6 (= *dimension clearly present*). A score of 6 on the rating scale of cognitive strategies could be achieved if the journal entries were highly organized (e.g., by identifying main points and arranging them in an ordered sequence, such as "first …, second …, third …") and highly elaborated (e.g., by providing own examples to illustrate abstracts concepts: "A good example of Piaget's notion of infantile egocentrism is when my little son shows me something in his picture book and disregards that I cannot see what he sees from my perspective."). A score of 6 on the rating scale of metacognitive strategies could be achieved if the journal entries included a high amount of monitoring (e.g., by specifying which contents were not yet understood: "I have not yet understood the exact difference between Piaget and Carey."), and planning of remedial strategies (e.g., attempts to solve the perceived gap in one's knowledge: "I will try to call to mind the presentation where different approaches to conceptual change were explained…"). Inter-rater reliability as determined by the intra-class coefficient was very good (ICC = .81 for the rating of cognitive strategies, ICC = .84 for the rating of metacognitive strategies).

Learning outcomes. We used two comprehension tests to assess learning outcomes, one of which was administered after the first half of the term and the other at the end of the term. Each test consisted of six open-end questions regarding the topics that had so far been discussed in the preceding seminar sessions. In order to answer these questions, the learners had to apply their knowledge, for example, by using theoretical concepts to explain self-generated examples (e.g., "Please provide a moral justification of a solution for the 'Heinz dilemma' which is on the conventional level according to Kohlberg's theory of moral development!"), or by dealing with the material in a critical manner (e.g., "Please discuss the challenges and rewards of Piaget's theory of cognitive stages critically!"). To score the level of comprehension in the answers we used the SOLO-Taxonomy ("Structure of Observed Learning Outcome") by Biggs and Collis (1982). Following the SOLO-Taxonomy, each answer was differentiated into six levels of knowledge ranging from 1 (= *no central points, low level of understanding, incoherent*) to 6 (= *all central points, high level of understanding, very coherent*). Inter-rater reliability as determined by the intra-class coefficient was very high (ICC = .96).

Motivation for writing learning journals. The students' motivation for writing learning journals was assessed after the first half of the term and again at the end of the term. We were interested to what extent the students enjoyed writing a learning journal, evaluated this type of follow-up coursework as valuable and useful, and how competent they perceived themselves in doing this. To measure these motivational factors, the students received subscales of the intrinsic motivation inventory which we adapted to the domain of journal writing (IMI; cf. Deci et al. 1994). More specifically, we administered modified versions of the subscale *interest/enjoyment* (e.g., "I enjoyed doing this activity very much"), of the subscale *effort/importance* (e.g., "I put a lot of effort into this"), and of the subscale *perceived competence* (e.g., "I think I am pretty good at this activity"). The students responded to these items on a 7-point rating scale ranging from 1 (= *not at all true*) to 7 (= *very true*). The reliability of the scales was good (Cronbach's α = .74 to .86).

Procedure

The students were asked to write a journal entry after each weekly seminar session. The required minimum text length was one page. For writing their weekly journal entry, the students logged on a web server. They downloaded a prepared file in Rich Text Format which included the instructions for writing the journal entry. Thus, it was guaranteed that the students had the instructions available while writing. The students were asked to note the exact time when they started writing a journal entry and when they had completed it. These data allowed us to roughly estimate the time students spent on writing journal entries. After completing a journal entry, the students uploaded it on the web server. Students who failed to upload their journal entry in time got a friendly reminder via email by the experimenter. That way it was ensured that the journals entries were written regularly and the number of missing journal entries was kept low. The students' learning outcomes and their motivation for journal writing were assessed twice: Once after the first half of the term and once again at the end of the term. These assessments took place in extra sessions at the Institute of Psychology. As part of these sessions, the students completed a comprehension test (learning outcomes) and the motivation questionnaire. Each session lasted about 1.5 h.

Results

Table 1 shows the mean scores and standard deviations for both experimental groups of the time spent on writing journal entries, the learning strategy measures, the comprehension tests, and the motivation scales. The mean scores are determined separately for the two measurement times, that is, after the first half of the term and at the end of the term. In case of the time spent on writing journal entries, mean scores were obtained by averaging the individual amounts of time indicated by the students for the six journal entries written in the first half of the term, and by averaging the amounts of time for the six journal entries produced until the end of the term. Similarly, mean scores for the learning strategy ratings were obtained by averaging the ratings for the six journal entries which a student had produced in the first half of the term, and by averaging the following six entries the student had written until the end of the term. An alpha level of .05 was used for all statistical tests. As an effect size measure, we used η^2-qualifying values $< .06$ as a weak effect, values in the range between .06 and .13 as a medium effect, and values $> .13$ as a large effect (see Cohen 1988).

Analysis of the learning journals

To check whether the amount of time spent by the students on writing their learning journals was comparable across experimental groups, we firstly conducted a mixed repeated measures analysis of variance with measurement time (first half of the term vs. end of the term) as a within-subjects factor and experimental condition (prompts vs. no prompts) as a between-subjects factor. Neither the main effect of experimental condition, $F(1, 48) = 0.28$, *ns*, nor the main effect of measurement time, $F(1, 48) = 1.92$, *ns*, were significant. The interaction between experimental condition and measurement time was also not significant, $F(1, 48) = 0.23$, *ns*. From this analysis it can be concluded that the time students spent on writing journal entries neither differed significantly between experimental groups nor between the first and second half of the term. Hence, potential

Table 1 Means and standard deviations (in parentheses) of the analysis of the learning journals, learning outcomes and motivation for journal writing in Experiment 1

Time	Groups	Time spent on writing a journal entry (minutes)	Cognitive strategies[a]	Meta-cognitive strategies[a]	Learning outcomes[b]	Interest/ enjoyment[c]	Effort/ importance[c]	Perceived competence[c]
1st Half of term	Prompts	74.91 (30.33)	4.00 (0.76)	2.68 (1.18)	3.95 (0.97)	3.96 (0.89)	4.43 (1.20)	3.84 (0.72)
	No prompts	68.80 (38.65)	3.70 (0.74)	2.19 (0.91)	3.55 (0.88)	4.01 (1.15)	4.44 (1.06)	3.75 (1.17)
2nd Half/end of term	Prompts	69.21 (27.70)	3.78 (0.63)	1.63 (0.71)	3.51 (0.79)	3.10 (1.00)	3.38 (1.09)	3.41 (0.89)
	No prompts	66.05 (31.59)	3.90 (0.82)	1.68 (0.84)	3.56 (0.85)	3.89 (1.31)	4.01 (1.17)	3.90 (1.42)

Means and SDs are separately plotted for measurement times and experimental groups

[a] 6-Point rating scale ranging from 1 (= dimension not present) to 6 (= dimension clearly present)

[b] 6-Point rating scale ranging from 1 (= no central points, low level of understanding, incoherent) to 6 (= all central points, high level of understanding, very coherent)

[c] 7-Point rating scale ranging from 1 (= not at all true) to 7 (= very true)

differences between experimental groups regarding learning outcomes cannot simply be attributed to differences in time spent on writing journal entries.

In the next step, we analyzed the extent to which cognitive and metacognitive strategies were present in the learning journals. For this purpose, we compared the mean ratings of the six journal entries written in the first half of the term with the mean ratings of the six journals entries written in the second half of the term. Averaging across single journal entries had two major advantages. Firstly, the single journal entries produced by a student may vary considerably. Thus, using average scores provided us with a more reliable measure of the learning strategies which a student elicited in her or his learning journal. Secondly, we avoided an unnecessary loss of data because if we had analyzed the learning strategy measures with repeated measures analyses of variance, students who had one or more missing journal entries would have been completely excluded as cases from the data analysis. Accordingly, the mean ratings for cognitive and metacognitive strategies were subjected to separate mixed repeated measures analyses of variance with measurement time (first half of the term vs. end of the term) as a within-subjects factor and experimental condition (prompts vs. no prompts) as a between-subjects factor.

In the first analysis of variance, the mean ratings of cognitive strategies were treated as the dependent variable. Neither the main effect of experimental condition, $F(1, 48) = 0.20$, *ns*, nor the main effect of measurement time, $F(1, 48) = 0.34$, *ns*, were significant. However, the ANOVA revealed a significant interaction effect between experimental condition and measurement time, $F(1, 48) = 9.68$, $p < .01$, $\eta^2 = .17$ (large effect). Figure 1 (left graph) shows the interaction between experimental condition and measurement time with regard to the presence of cognitive learning strategies in the students' learning journals. As the left diagram in Fig. 1 shows, students who regularly received prompts for writing their learning journal, elicited more cognitive strategies in their learning journal in the first half of the term as compared with the journal entries produced in the second half. In contrast, students who received only a general instruction that contained no specific prompts applied less cognitive strategies in their journal entries written in the first half of the term than in the second half of the term.

Similar results were obtained with regard to the presence of metacognitive strategies in the learning journals. The main effect of the experimental condition was not significant, $F(1, 48) = 0.88$, *ns*. However, a significant main effect of measurement time was found, $F(1, 48) = 48.07$, $p < .001$, $\eta^2 = .50$ (large effect). Again, there was a significant interaction effect between experimental condition and measurement time, $F(1, 48) = 5.78$, $p < .05$, $\eta^2 = .11$ (medium effect). As Fig. 1 (right diagram) shows, the extent to which

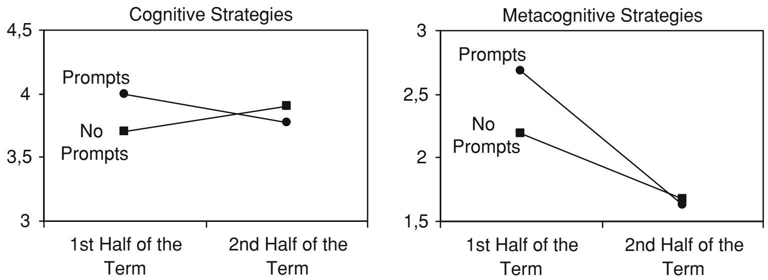

Fig. 1 Interaction between experimental condition and measurement time in Experiment 1. *Left diagram* interaction effect on cognitive strategies elicited in the learning journals. *Right diagram* interaction effect on metacognitive strategies

the students in both conditions elicited metacognitive strategies in their learning journals clearly decreased in the second half of the term. However, the significant interaction effect indicates that the students in the prompting condition were more strongly affected by this decrease because they started on a higher level and applied more metacognitive strategies in their journals entries written in the first half of the term than the students in the control condition.

Learning outcomes

We analyzed the students' learning outcomes by experimental condition. As one comprehension test had been administered after the first half of the term and the other at the end of the term, the tests covered different topics. Although we carefully sought to construct items of similar difficulty, it cannot be ruled out that the tests nevertheless differed with regard to their difficulty. Consequently, a general main effect of measurement time should not be interpreted as an increase or decrease in learning outcomes because it may be due to test difficulty. However, it makes sense to interpret potential interaction effects because a significant interaction would indicate different trends within the experimental conditions.

In this analysis of variance, the main effect of experimental condition was not significant, $F(1, 48) = 0.59$, *ns*. However, the main effect of measurement time was significant, $F(1, 48) = 4.10$, $p < .05$, $\eta^2 = .08$ (medium effect). The ANOVA further revealed a significant interaction effect between experimental condition and measurement time, $F(1, 48) = 4.29$, $p < .05$, $\eta^2 = .08$ (medium effect). As Fig. 2 (left diagram) shows, the main effect of measurement time has to be qualified by the significant interaction effect: In the first comprehension test after half of the term, the students in the prompting condition achieved clearly higher test scores than the students in the control condition. However, when learning outcomes were measured again at the end of the term, the prompting condition group did not perform any better than the control group without prompts.

Motivation in writing learning journals

To investigate the effects of prompts on the students' motivation for writing a learning journal, we analyzed the students' mean ratings for interest/enjoyment, effort/importance and perceived competence. The results fit in the overall pattern reported so far. For interest/enjoyment, the main effect of experimental condition was not significant, $F(1, 48) = 2.13$, *ns*. However, a significant main effect of measurement time resulted, $F(1, 48) = 14.47$,

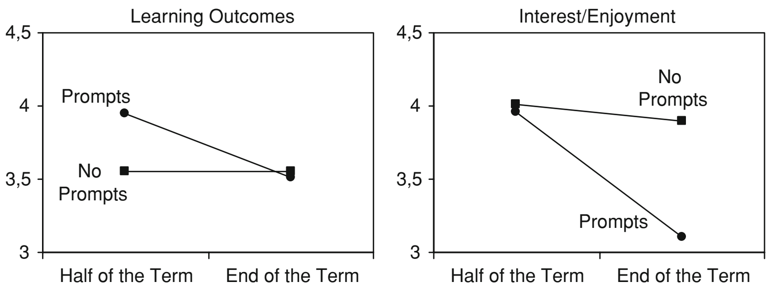

Fig. 2 Interaction between experimental condition and measurement time in Experiment 1. *Left diagram* interaction effect on learning outcomes. *Right diagram* interaction effect on the students' interest/enjoyment in writing learning journals

$p < .001$, $\eta^2 = .24$ (large effect), which should be qualified by the significant interaction between experimental condition and measurement time, $F(1, 48) = 8.36$, $p < .01$, $\eta^2 = .15$ (large effect). As Fig. 2 (right diagram) shows, the students' enjoyment of writing a learning journal decreased over the course of the term. However, this decrease was evidently much more marked for the prompting condition as compared with the control condition.

Similar results occurred for the students' invested effort: The main effect of experimental condition was not significant, $F(1, 48) = 1.21$, *ns*. Yet, a significant main effect of measurement time, $F(1, 48) = 30.61$, $p < .001$, $\eta^2 = .39$ (large effect), and a significant interaction between experimental condition and measurement time, $F(1, 48) = 5.37$, $p < .05$, $\eta^2 = .11$ (medium effect), were obtained. Thus, the students generally put less effort in writing their learning journal towards the end of the term. However, this decrease was substantially stronger for the students in the prompting condition than for the students in the control condition.

In the final analysis, we tested whether the students' perceptions regarding their competence in writing a learning journal changed over the course of the term. In this ANOVA, the main effects of measurement time and experimental condition were not significant, $F(1, 48) = 1.28$ and $F < 1$, respectively. However, consistent with the previous analyses, the interaction between experimental condition and measurement time reached statistical significance, $F(1, 48) = 5.62$, $p < .05$, $\eta^2 = .11$ (medium effect). The students in the experimental condition felt more competent in journal writing in the beginning of the term than in the end of the term. In contrast, the perceived competence of the students in the control condition increased towards the end of the term.

Discussion

The results of the present study demonstrate the pitfalls of prompting procedures in writing-to-learn. In the beginning of the term, the prompts successfully activated cognitive and metacognitive strategies that resulted in superior learning outcomes. However, in the longer term, this picture changed. Students who received prompts applied fewer strategies than in their initial journal entries. Their learning outcomes, interest in journal writing, and their invested effort also decreased.

How can these results be explained? It is possible, that the more the students became familiar with the learning journal method and adopted the desired strategies, the more the external guidance by prompts became dispensable and interfered with the students' internal tendency to apply the strategies by themselves. These results may be explained from the perspective of Cognitive Load Theory (Sweller 2005; Sweller et al. 1998). Generally, in terms of Cognitive Load Theory, germane cognitive load refers to those working memory resources that are invested in the execution of useful activities—such as the application of cognitive and metacognitive learning strategies—that help the learner to reach the intended learning goal (see Kalyuga 2007). In contrast, extraneous cognitive load refers to those cognitive resources that are devoted to unproductive activities irrelevant to learning goals (in the context of journal writing, e.g., concentrating on spelling and grammar instead of identifying comprehension difficulties). Extraneous cognitive load typically results from design-related factors such as poor presentation design or inadequate instructional support (see Kalyuga 2007). Against this theoretical background, the results of the present study may be interpreted as follows: In the beginning of the term, when the students were largely unfamiliar with the learning journal method, the prompts might have effectively supported them in applying beneficial cognitive and metacognitive strategies when writing a journal entry. Thus, the instructional support provided by prompts might have increased germane (useful) cognitive load and, at the same time, reduced extraneous (wasteful) cognitive load because the complex task of producing a journal entry was facilitated by the prompts (e.g., by reducing the need for effortful search processes). However, the more the students became skilled in journal writing and knew themselves how to apply the prompted strategies, the more the external guidance by prompts turned into a redundant stimulus that interfered with the students' internal tendency to apply the strategies. Thus, at some point in the term, the prompts probably ceased to function as strategy *activators* (Reigeluth and Stein 1983) but rather functioned as strategy "inhibitors". Instead of contributing to increasing germane cognitive load, the prompts might have imposed unnecessary extraneous cognitive load. Together, these results provide evidence for an expertise reversal effect in writing-to-learn (Kalyuga et al. 2003; Kalyuga 2005, 2007): Instructional aids which effectively off-load working-memory and facilitate learning for beginners may produce reverse effects when offered to advanced learners with higher levels of prior knowledge or skills.

Following Collins et al. (1989) theory of Cognitive Apprenticeship (see also Collins 2006), a gradual and adaptive fading of the prompts might offer a possible solution to the pitfalls of over-prompting. According to this instructional approach, the prompts could gradually be faded out with increasing individual competence in applying cognitive and metacognitive strategies (see McNeill et al. 2006; Puntambekar and Hübscher 2005). For example, as soon as a student shows sufficient elaboration strategies in his/her learning journal, elaboration prompts could be removed from the instruction for the subsequent journal entries. The following study tested empirically such an adaptive fading procedure.

Experiment 2: adaptive fading-out of prompts

The following rationale for sequencing the fading procedure was used in this study (see Fig. 3 for a graphical overview of the sequence). In the beginning of the term, the students in the fading group of Experiment 2 received the same combination of cognitive and metacognitive prompts which had previously been used in Experiment 1. In order to provide ample opportunity to get familiar with the affordance of writing a learning journal with the help of the prompts, the students were offered the complete set of prompts for writing the first four journal entries in the term. The fading-out of prompts started from the 5th journal entry on. Every fading-step was based on an analysis of the previous two journal entries. That is, for each of the six prompts, we determined to what extent each student had been able to realize the prompted strategy in these journal entries by using 6-point rating scales ranging from 1 (= *dimension not present*) to 6 (= *dimension clearly present*). If the mean rating of the two journal entries regarding a prompted strategy

	No Fading				Fading 1		Fading 2		Fading 3		Fading 4	
Journal Entry	1	2	3	4	5	6	7	8	9	10	11	12
Basis for Fading					Ratings of Entries 3 and 4		Ratings of Entries 5 and 6		Ratings of Entries 7 and 8		Ratings of Entries 9 and 10	

Fig. 3 Fading sequence in Experiment 2

(e.g., mean rating for generated examples to illustrate abstracts concepts) reached the threshold value of 4.5, the corresponding prompt was removed from the instruction for the subsequent journal entries to be written in the remainder of the term. The criterion of 4.5 as the threshold value was determined on practical grounds by inspecting the learning journals collected in Experiment 1. As most students in that study were able to reach such a level of strategy use in their journals after some time of practice, this criterion seemed to be viable enough for the purposes of this study.

Sample and design

Sixty-two first semester students of Psychology (50 females, 12 males, mean age: 22.75 years) participated in the study. They were randomly assigned three parallel introductory courses in Developmental Psychology. The students completed these courses as regular part of their undergraduate studies in Psychology at the University of Freiburg. The courses lasted 4 months with weekly seminar sessions. As in the previous study, the students had to write twelve journal entries within this time. In all three courses the same contents were taught. However, because of organizational reasons, it was not possible that all three courses were taught by the same lecturer. One course was taught by a different lecturer. As this lecturer did not demand from her students journal entries as follow-up course work to the weekly sessions, these students were treated as a baseline group ($n = 17$). Since a different lecturer may constitute a confounding variable, the results of this group should be interpreted with some reservation.

To investigate the effects of the fading procedure, we drew upon the two courses which were taught by the same lecturer (i.e., a member of our research group). The adaptive fading-out of prompts was realized in one of the courses (the faded prompts group, $n = 20$). In the other course (the permanent prompts group, $n = 25$), the six prompts were presented permanently, that is, the students received the same six cognitive and metacognitive prompts for each weekly journal entry over the whole term. Hence, a one-factorial design was used comprising three groups with repeated measurements on several dependent variables. As in Experiment 1, dependent variables encompassed measures of the learning strategies elicited in the learning journals, the students' learning outcomes, as well as measures of the students' motivation for writing the learning journal.

Materials, instruments, codings, and procedure

The instructions for writing the learning journal, the instruments used to measure learning outcomes and the students' motivation for writing the learning journal were the same as in Experiment 1. The analysis of the learning journals was conducted as described in the corresponding section of Experiment 1. The same procedure was used except that the instruction for writing the learning journal gradually changed over time in the fading group, because the prompts were faded out individually for each student according to his/her level of competence.

Results

Table 2 shows the mean scores and standard deviations on the learning strategy measures, the comprehension tests, and the motivation scales. The mean scores are plotted separately for the different measurement times and experimental groups. The mean scores for learning

Table 2 Means and standard deviations (in parentheses) of the analysis of the learning journals, learning outcomes and motivation for journal writing in Experiment 2

Time	Groups	Cognitive strategies[a]	Meta-cognitive strategies[a]	Time	Learning outcomes[b]	Interest/ enjoyment[c]	Effort/ importance[c]	Perceived competence[c]
Before fading	Faded prompts	3.51 (0.83)	2.32 (0.87)	Half of the term	3.61 (0.68)	3.67 (1.41)	3.85 (1.32)	3.61 (1.22)
	Permanent prompts	3.53 (0.52)	1.90 (0.69)		3.56 (0.78)	4.00 (1.00)	4.43 (0.90)	3.69 (1.41)
	Baseline group	–	–		2.99 (1.15)	–	–	–
During fading	Faded prompts	3.86 (0.60)	1.93 (0.73)	End of the term	3.31 (0.79)	3.05 (1.31)	3.27 (1.91)	3.45 (1.63)
	Permanent prompts	3.34 (0.59)	1.70 (0.56)		2.76 (0.86)	3.59 (0.94)	4.24 (0.93)	3.83 (1.37)
	Baseline group	–	–		2.95 (0.91)	–	–	–

Means and SDs are separately plotted for measurement times and experimental groups

[a] 6-Point rating scale ranging from 1 (= dimension not present) to 6 (= dimension clearly present)

[b] 6-Point rating scale ranging from 1 (= no central points, low level of understanding, incoherent) to 6 (= all central points, high level of understanding, very coherent)

[c] 7-Point rating scale ranging from 1 (= not at all true) to 7 (= very true)

strategy ratings were obtained by averaging the ratings of the four journal entries which a student had produced before the fading procedure started (see Table 2, "before fading"), and, similarly, by averaging the following eight entries the student had written until the end of the term after the fading procedure had started (see Table 2, "during fading"). Scores for learning outcomes and the motivation for writing a learning journal were collected twice during the term, after half of the term and at the end of the term (see Table 2, fifth column).

Analysis of the learning journals

To analyze the extent to which cognitive and metacognitive strategies were present in the learning journals, we subjected the mean ratings for cognitive and metacognitive strategies to separate mixed repeated measures analyses of variance with measurement time (before fading vs. during fading) as a within-subjects factor and experimental condition (faded prompts vs. permanent prompts) as a between-subjects factor.

In the ANOVA with the mean ratings of cognitive strategies as the dependent variable, neither the main effect of experimental condition, $F(1, 43) = 2.20$, *ns*, $\eta^2 = .05$ (small effect) nor the main effect of measurement time, $F(1, 43) = 0.83$, *ns*, $\eta^2 = .02$ (small effect) were significant. However, the ANOVA revealed a significant interaction effect between experimental condition and measurement time, $F(1, 43) = 9.19$, $p < .01$, $\eta^2 = .18$ (large effect). Figure 4 (left graph) shows the interaction between experimental condition and measurement time with regard to the presence of cognitive strategies in the students' learning journals in Experiment 2. As the left diagram in Fig. 4 shows, the students in the faded prompts group started almost exactly from the same level of cognitive strategies in the learning journals as the permanent prompts group. However, after the fading procedure begun, the amount of cognitive strategies in the faded prompts group clearly increased while, at the same time, the amount of cognitive strategies elicited by the permanent prompts group decreased.

The ANOVA with the mean ratings of metacognitive strategies as the dependent variable showed a different pattern of results. The main effect of experimental condition just failed to reach statistical significance, $F(1, 43) = 3.18$, $p < .10$, $\eta^2 = .07$ (small effect). The main effect of measurement time, $F(1, 43) = 7.18$, $p < .01$, $\eta^2 = .14$, (large effect) proved to be significant. In contrast to the previous analysis, the interaction between experimental condition and measurement time was not significant, $F(1, 43) = 0.71$, *ns*, $\eta^2 = .02$ (small effect). As the right diagram in Fig. 4 shows, the amount of metacognitive

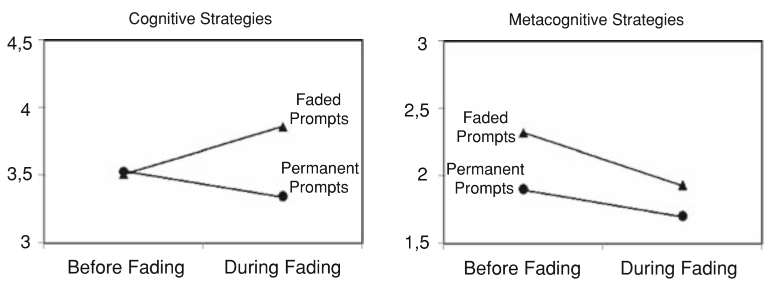

Fig. 4 Interaction between experimental condition and measurement time in Experiment 2. *Left diagram* interaction effect on cognitive strategies elicited in the learning journals. *Right diagram* interaction effect on metacognitive strategies

strategies in the learning journals decreased over the term regardless of whether the prompts were adaptively faded-out or presented permanently. Thus, compared with the extent of cognitive strategies in the learning journals, the fading procedure apparently was not powerful enough to prevent the decline of metacognitive strategies over the course of the term.

Learning outcomes

As in Experiment 1, two comprehension tests with open-ended questions were administered in additional sessions, one taking place after the first half of the term and the other at the end of the term. A mixed analysis of variance was conducted with measurement time as a within-subjects factor (half of the term vs. end of the term) and experimental condition (faded prompts, permanent prompts, baseline group) as a between-subjects factor. Unfortunately, we had quite a considerable loss of data in this analysis, because 7 students in the faded prompts group and five students in the permanent prompts group failed to participate in the second comprehension test at the end of the term. Thus, the ANOVA is based on 13 cases in the faded prompts group, 20 cases in the permanent prompts group, and 17 cases in the baseline group.

In this analysis, the main effect of experimental condition was not significant, $F(2, 47) = 1.57$, *ns*, $\eta^2 = .06$ (small effect). There was a significant main effect of measurement time, $F(1, 47) = 8.58$, $p < .01$, $\eta^2 = .15$ (large effect), but—as in Experiment 1—this effect should not be interpreted because it cannot be ruled out that the two comprehension tests differed with regard to their difficulty. Nevertheless, it is possible to interpret the significant interaction effect between measurement time and experimental condition, $F(2, 47) = 3.34$, $p < .05$, $\eta^2 = .12$ (medium effect), because this interaction indicates that apparently there were different trends within the experimental conditions. As Fig. 5 illustrates, both the faded prompts and the permanent prompts groups achieved higher learning outcomes than the baseline group in the first comprehension test after half of the term, $F(1, 47) = 4.81$, $p < .05$, $\eta^2 = .09$ (medium effect). However, in the second comprehension test at the end of the term, the students in the permanent prompts group evidently lost their superiority over the baseline group. This trend is statistically supported by a significant interaction effect that results when only the permanent prompts group and the baseline group are compared in another mixed ANOVA with repeated measures (i.e., first vs. second comprehension test), $F(1, 35) = 6.39$, $p < .05$, $\eta^2 = .15$ (large effect). On

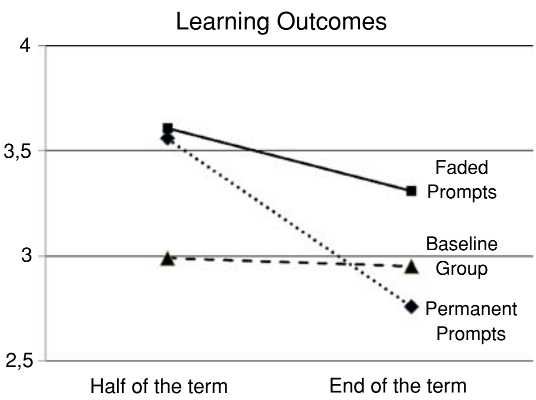

Fig. 5 Interaction between experimental condition and measurement time in Experiment 2: effect on learning outcomes

the other hand, when the faded prompts group is contrasted with the baseline group, the interaction between experimental condition and measurement time is not significant, $F(1, 28) = 0.42$, *ns*, $\eta^2 = .02$ (small effect). The main effect of experimental condition, however, just fails to reach statistical significance, $F(1, 28) = 3.12$, $p = .09$, $\eta^2 = .10$ (medium effect). Apparently, the adaptive fading-out of the prompts helped the students in the faded prompts group to better maintain their level of performance as compared with the students in the permanent prompts condition.

Motivation in writing learning journals

Besides learning outcomes, we further assessed the students' motivation for writing a learning journal. Against the background of the results of Experiment 1, we were especially interested whether the gradual and adaptive fading-out of the prompts would help to mitigate the negative long-term effects of the prompts on the students' interest/enjoyment, effort/importance, and perceived competence in journal writing. However, the results of the present study turned out to be quite disappointing in this regard. The analysis of variance for interest/enjoyment yielded a significant main effect of measurement time, $F(1, 31) = 14.43$, $p < .001$, $\eta^2 = .32$ (large effect), showing that the students' interest in writing a learning journal decreased both in the fading and permanent prompts conditions over the course of the term. Neither the main effect of experimental condition, $F(1, 31) = 1.29$, *ns*, $\eta^2 = .04$ (small effect), nor the interaction of measurement time and experimental condition, $F(1, 31) = 0.59$, *ns*, $\eta^2 = .02$ (small effect), were statistically significant. A similar pattern occurred for the analysis of the students' effort/importance ratings. The main effects of measurement time and experimental condition just failed to reach statistical significance, $F(1, 31) = 4.08$, *ns*, $\eta^2 = .12$ (medium effect) and $F(1, 31) = 3.72$, *ns*, $\eta^2 = .11$ (medium effect), respectively. The interaction of measurement time and experimental condition was not significant, $F(1, 31) = 1.05$, *ns*, $\eta^2 = .03$ (small effect). In the final ANOVA with the students' perceived competence ratings as the dependent measure, neither the main effects of measurement time, $F(1, 31) = 0.001$, *ns*, $\eta^2 = .0005$ (small effect), and experimental condition, $F(1, 31) = 0.21$, *ns*, $\eta^2 = .007$ (small effect), nor the interaction of both factors were significant, $F(1, 31) = 0.95$, *ns*, $\eta^2 = .03$ (small effect). Overall, this pattern of results suggests that the students' motivation for writing a learning journal decreased in both experimental conditions of Experiment 2 regardless of whether the prompts were presented permanently or faded-out. We will discuss the implications in the following section.

General discussion

We consider journal writing as a productive way of doing follow-up coursework in a self-guided fashion (Nückles et al. 2009). In our instructional approach, students typically are required to write a journal entry about the contents of a lecture or seminar session they have previously attended. In the lecture or seminar session, the lecturer, or sometimes peer students present new information in didactic ways that take into account the limitations of human working memory (e.g., by presenting students with visualizations and worked-examples of complex concepts). Back home, students then are expected to perform follow-up course in a more self-guided fashion: By writing a journal entry, students should in particular attempt to organize the previously presented information into a coherent whole and to integrate it into their prior knowledge (Mayer 2002). Externalizing one's thoughts in a learning journal may further facilitate taking a metacognitive stance towards one's own learning and thinking processes, thereby supporting the monitoring of comprehension and evaluation of learning outcomes (Nückles et al. 2009). In a nutshell, we conceive of journal writing as a valuable complement to more teacher-guided course work that supports the acquisition and consolidation of deep and sustained knowledge.

The aim of the present longitudinal studies was to investigate longer-term effects of prompts as a method to support the writing of productive learning journals. Previous experimental research suggests that prompts may be very effective in stimulating cognitive and metacognitive strategies in writing a journal entry (Berthold et al. 2007; Nückles et al. 2009; Hübner et al. 2009; Schwonke et al. 2006). However, usually students are not required to merely produce a single journal entry as in those studies. In school and academic educational settings, journal writing is typically introduced as a regular follow-up course work activity, for example, writing a journal entry after each weekly seminar session (McCrindle and Christensen 1995; Nückles et al. 2004). Thus, the question arises how the provision of prompts as strategy activators (Reigeluth and Stein 1983) would influence the application of learning strategies, the students' learning outcomes, and also their motivation in keeping a learning journal over a longer period of time.

The results of the presented two studies confirm the results of previous experimental research inasmuch as short-term effects are concerned. In Experiment 1, prompts had positive effects on strategy use and learning outcomes in the beginning of the term. Students who received cognitive and metacognitive prompts elicited a higher degree of cognitive and metacognitive strategies in their initial journal entries than the students in the control condition. Accordingly, the prompted students outperformed the control students in the comprehension test after half of the term. In the longer term, however, this picture changed. Students who received prompts applied significantly fewer strategies than in their initial journal entries. Their learning outcomes, their invested effort and interest in journal writing clearly decreased. The control students' writing, in contrast, developed more positively over the course of the term. They elicited more cognitive strategies in their journal entries written in the second half of the term than in their entries produced in the first half. Also, their writing motivation evidently suffered less than the motivation of the students in the experimental group.

From the perspective of Cognitive Load Theory, these results can be explained in terms of an expertise reversal effect (Kalyuga 2005, 2007; Kalyuga et al. 2003). Accordingly, the prompts might have effectively supported the experimental group's students in applying cognitive and metacognitive strategies as long as they were unfamiliar with the learning journal method. Thus, in the beginning of the term, the prompts might have increased germane cognitive load in as much as strategy use was raised. The prompts successfully functioned as strategy activators (see Reigeluth and Stein 1983). They facilitated the task for the students to enact beneficial strategies, such as organization and elaboration strategies, which following Mayer's (2002) selecting-organizing-integrating theory of active learning are at the heart of meaningful learning. It can be assumed that the undergraduate students in our sample, in principle, already knew these strategies, but—without instructional support—they would not apply, or apply them to an unsatisfactory degree in the context of journal writing as follow-up course work. Extraneous cognitive load was reduced because the external guidance by prompts prevented the inexperienced students from unproductive search processes and uncertainties involved in writing journals entries. However, the more the students in our study became proficient in applying the desired strategies in the context of journal writing, the more the external guidance by prompts became redundant and therefore contributed more to unnecessary extraneous cognitive load rather than to germane load. Accordingly, the experimental students' strategy use, their learning outcomes and motivation for writing the journal decreased towards the end of term.

Experiment 2 investigated whether a gradual and adaptive fading-out of the prompts would be an appropriate means to mitigate the expertise reversal effect observed in Experiment 1. According to the theory of Cognitive Apprenticeship (see Collins 2006; Collins et al. 1989) this procedure should work because a particular prompt would be available to the student as long as s/he needs it to successfully enact the corresponding strategy. But as soon as the student has internalized the strategy and mastered it to a sufficient degree in her/his journal entries, the prompt would be removed. Hence, in terms of Cognitive Load Theory (Sweller 2005), the instructional guidance in the form of prompts would be provided as long as it serves to increase germane load and to reduce extraneous load. However, as soon as this proportion tends to reverse, the guidance would be removed. The results of Experiment 2 are partially consistent with this theoretical assumption. The adaptive fading-out of prompts raised the extent to which the students applied cognitive strategies in their learning journals over the course of the term. In contrast, presenting the prompts permanently resulted in a substantial decrease of cognitive strategies. In line with these results, the students in the faded prompts condition achieved high learning outcomes in both half-of-the-term and end-of-the-term comprehension tests, whereas the permanent prompts group fell below the level of the baseline group in the second comprehension test. On the other hand, the fading procedure proved to be rather ineffective in preventing the decrease of metacognitive strategies over the course of the term. Similarly, the students' motivation for journal writing decreased from the first to the second half of the term regardless of whether the prompts were presented permanently or gradually and adaptively faded-out.

Why was the fading procedure only partially successful? With regard to the decrease of metacognitive strategies, it is noteworthy that such a decrease occurred in all the journal-writing conditions of Experiment 1 and Experiment 2. Thus, it may generally be difficult for students to maintain a high level of metacognitive reflection and comprehension monitoring over a longer period of time. In other words, metacognitive reflection possibly is a germane-load activity that students "naturally" tend to minimize because they may find it laborious and little rewarding to continuously question their own understanding over a longer period of time.

The fact that the fading-out of the prompts did not affect the students' motivation for journal writing positively is more difficult to explain. Despite the small and also unequal samples in Experiment 2 (only 13 cases in the faded prompts condition vs. 20 cases in the permanent prompts condition), a significant decrease of the students' motivation for journal writing was observed regardless of whether the prompts were faded out or presented permanently (main effects of measurement time with regard to interest/enjoyment and effort/importance). This result is consistent with Experiment 1. In that study, the prompted students' writing motivation decreased sharply whereas in the "free-writing" control condition, students' writing motivation either decreased more slowly or even increased over the term. Thus, it is possible that structuring the writing of learning journals by prompts, irrespective of whether the prompts are faded out or not, has negative longer-term side-effects on the students' writing motivation. Providing students with meta-strategic knowledge about the prompted strategies, that is, knowledge about how, when, and why to use cognitive strategies might offer a possible solution to this problem (see Hübner et al. 2009; Schraw 1998; Zohar and Peled 2008). In the literature on cognitive training, such an approach is typically called “informed training” (Paris et al. 1983) or “informed prompting” (Hübner et al. 2009).

Further limitations of the present studies

Another important question that needs to be addressed refers to the extent by which the prompted cognitive and metacognitive strategies may transfer to other unrelated areas. In other words, are the students in our experimental studies now better learners as a consequence of being taught cognitive and meta-cognitive strategies in all areas they subsequently face? To be honest, it is currently impossible to provide a satisfactory answer to this question. We certainly do hope that, in the long run, students will adopt the writing of learning journals in order to deepen and reflect on learning contents in other domains than the one investigated here. However, up to now, we do not have any systematic data on this issue. Further research is needed to examine whether and to what extent students internalize the learning journal method and adopt it as a personal way of deepening and reflecting about learning contents.

What are the broader theoretical implications of the present research?

Negative side-effects of instructional support methods, which are similar to the expertise reversal effect in writing-to-learn observed in Experiment 1, have also been reported in other domains and learning settings (e.g., see also other contributions to this special issue). For example, in the domain of computer-supported collaborative learning, several authors have recently discussed the problem of “over-scripting” in relation to computer-supported cooperation scripts (cf. Dillenbourg and Jerman 2007; Weinberger et al. 2005). In the case of cooperation scripts, the danger of over-scripting is particularly likely if the script makes very concrete and detailed prescriptions of how to behave. As long as the learners are pretty unfamiliar with the activities stipulated by the script, they may find the prescriptions helpful. Accordingly, the script might enable them to enact beneficial learning activities, such as argumentation skills, which they either would not realize spontaneously or realize to an unsatisfactory degree. However, the more the learners become proficient in the scripted activities, the more the prescriptions may limit the learners’ autonomy and latitude. As a result, the learner’s motivation to enact the activities stipulated by the script may be corrupted. Thus, over-scripting effects in CSCL are often explained by referring to motivational concepts. Similarly, Deci and Ryan discussed the detrimental effects of external regulation on intrinsic motivation in their theory of self-determination (see Deci et al. 1999; Ryan and Deci 2000).

Against this background, the present research suggests that “over-prompting” or “over-scripting” effects should ideally be examined from both cognitive and motivational perspectives. Experiment 1 showed that a prompting procedure which had successfully been implemented in previous experimental studies (see Berthold et al. 2007; Nückles et al. 2009) yielded negative longer-term effects on both cognitive and motivational variables. In Experiment 2, the negative longer-term effects on cognitive strategies and learning outcomes were successfully mitigated. The detrimental effects on the students’ motivation, however, persisted. Thus, in order to fully understand the negative side-effects of instructional aids as those described above, theoretical explanations in terms of Cognitive Load Theory (Sweller 2005) should be systematically related to motivational theories, such as self-determination theory (see Ryan and Deci 2000). This would allow for a comprehensive and detailed analysis of how cognitive effects of instructional aids on working memory capacity are associated with (or rather dissociated from) motivational effects such as perceived autonomy and competence. It is up to further research to explore these issues.

Acknowledgments This research was supported by the Deutsche Forschungsgemeinschaft (DFG; [German Research Foundation]) with a project grant awarded to Matthias Nückles (contract NU 129/2-1). We would like to thank Bettina Brockhaus and Julian Roelle for their assistance in coding the learning journals.

References

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Barnett, J. E., DiVesta, F. J., & Rogozinski, J. T. (1981). What is learned in note taking? *Journal of Educational Psychology, 73*, 181–192.

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564–577.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Brown, A. L. (1978). Knowing when and how to remember: A problem of metecognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1, pp. 225–253). Hillsdale, NJ: Erlbaum.

Cantrell, R. J., Fusaro, J. A., & Dougherty, E. A. (2000). Exploring the effectiveness of journal writing on learning social studies: A comparative study. *Reading Psychology, 21*, 1–11.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collins, A. (2006). Cognitive apprenticeship. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 47–60). Cambridge: Cambridge University Press.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction* (pp. 453–494). Hillsdale, NJ: Erlbaum.

Connor-Greene, P. A. (2000). Making connections: Evaluating the effectiveness of journal writing in enhancing student learning. *Teaching of Psychology, 27*, 44–46.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114.

Deci, E. L., Eghari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*, 119–142.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

Dillenbourg, P., & Jerman, P. (2007). Designing integrative scripts. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning. Cognitive, computational and educational perspectives* (pp. 259–289). Berlin: Springer.

Efklides, A., & Vauras, M. (Eds.). (1999). Metacognitive experiences in their role in cognition (special issue). *European Journal of Psychology of Education, 14*, 455–584.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum.

Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural/process theories of complex human behaviour* (pp. 213–245). Alphen aan den Rijn, The Netherlands: Sijthoff and Noordhoff.

Hübner, S., Nückles, M., & Renkl, A. (2009). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*. doi:10.1016/j.learninstruc.2008.12.001.

Kalyuga, S. (2005). Prior knowledge principle. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 325–337). New York: Cambridge University Press.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509–539.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31.

King, A. (1992). Comparison of self-questioning, summarizing, and note-taking-review as strategies for learning from lectures. *American Educational Research Journal, 29*, 303–323.

Mayer, R. E. (1984). Aids to text comprehension. *Educational Psychologist, 19*, 30–42.

Mayer, R. E. (2002). *The promise of educational psychology: Vol. 2. Teaching for meaningful learning*. Upper Saddle River, NJ: Prentice Hall.

McCrindle, A., & Christensen, C. (1995). The impact of learning journals on metacognitive and cognitive processes and learning performances. *Learning and Instruction, 5*, 167–185.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*, 153–191.

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

Nückles, M., Schwonke, R., Berthold, K., & Renkl, A. (2004). The use of public learning diaries in blended learning. *Journal of Educational Media, 29*, 49–66.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293–316.

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*, 91–109.

Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40*, 1–12.

Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 335–382). Hillsdale, NJ: Lawrence Erlbaum Associates.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education, 14*, 477–488.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*, 113–125.

Schwonke, R., Hauser, S., Nückles, M., & Renkl, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior, 22*, 77–92.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Sweller, J., van Merriënboer, J. G., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Weinberger, A., Reiserer, M., Ertl, B., Fischer, F., & Mandl, H. (2005). Facilitating collaborative knowledge construction in computer-mediated learning environments with cooperation scripts. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), *Barriers, biases and opportunities of communication and cooperation with computers* (pp. 15–37). New York: Springer.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In C. M. Wittrock (Ed.), *Handbook of research in teaching* (pp. 315–327). New York: Macmillan Publishing Company.

Zohar, A., & Peled, B. (2008). The effects of explicit teaching of metastrategic knowledge on low- and high-achieving students. *Learning and Instruction, 18*, 337–353.