Journal writing in science: Effects on comprehension, interest, and critical reflection

Kristin Wäschle, Anja Gebhardt, Eva-Maria Oberbusch, & Matthias Nückles

Department of Educational Science; University of Freiburg | Germany

Abstract: Writing-to-learn assignments typically foster deep comprehension of learning contents. Journal writing, in particular, promotes the application of learning strategies, which promote learners' comprehension, interest in a topic and ability to critically reflect on learning contents. Against this background, we conducted two longitudinal field studies. In Study 1, twenty-one students wrote learning journal entries after their biology lessons. After the intervention period, they showed better scores in comprehension, interest and critical reflection than the control class, in which students (n=25) completed other homework assignments. Mediation analyses revealed a domino effect: Journal writing improved comprehension, which led to raised interest, which resulted in superior critical reflection. Study 2 further investigated the role of learners' interest in improving critical reflection. Students in the experimental condition (n=13) received a personal-utility prompt in addition to cognitive and metacognitive prompts to support journal writing. In the control group (n=11), students only received cognitive and metacognitive prompts. The experimental group showed higher interest scores after the intervention period and a better quality of critical reflections on a bio-ethical issue than the control group. Together, these studies illustrate the potentials of journal writing for improving learners' comprehension, their interest and ability to critically reflect on complex scientific issues.

Keywords: Learning journals, comprehension, interest, critical reflection, science education

journal of WRITING RESEARCH

Wäschle, K., Gebhardt, A., Oberbusch, E.M., & Nückles, M. (2015). Journal writing in science: Effects on comprehension, interest, and critical reflection. *Journal of Writing Research, 7*(1), *41-64*. http://dx.doi.org/10.17239/jowr-2015.07.01.03

Contact: Kristin Wäschle, University of Freiburg, Department of Educational Science, Rempartstrasse11, D-79098, Freiburg | Germany - kristin.waeschle@ezw.uni-freiburg.de

The goal of science education is to help students, regardless of their individual prior knowledge, understand scientific concepts, reflect critically on scientific issues (Ford & Yore, 2012; Salomon & Perkins, 1998), and develop commitment to lifelong learning (Tunnicliffe & Ueckert, 2007). These goals are particularly important in domains where scientific knowledge accumulates and changes rapidly, such as in human biology. It is important to investigate learning assignments that promote comprehension, critical reflection and motivation. Writing-to-learn has proven itself to be a beneficial learning tool in science education (Gunel, Hand, & Prain, 2007; Reynolds, Thaiss, Katkin, & Thompson, 2012; Webb, 2010). Journal writing seems to be especially effective, even for relatively young learners (Schmidt, Maier & Nückles, 2012). However, in prior research, researchers mainly focused on the effect of journal writing on comprehension (e.g., Glogger, Schwonke, Holzäpfel, Nückles, & Renkl, 2012). The present studies demonstrate the potential of writing learning journals beyond supporting comprehension. We will argue and show that journal writing may also promote motivation and critical reflection in middle (secondary) and high school education.

1. Effects of journal writing on comprehension, motivation, and reflection

Writing assignments (e.g., writing essays or summaries) generally have positive effects on learners' comprehension of complex topics (Bangert-Drowns, Hurley, & Wilkinson, 2004; Klein, 1999). However, in order to make writing an effective learning tool, it is particularly important that writing tasks support the application of learning strategies (Nückles, Hübner, & Renkl, 2009; 2012) that allow for the integration of new learning content into the learner's already existing cognitive representations (Mayer, 2002). In science education, approaches like the Science Writing Heuristic (SWH) illustrate the positive effects of writing-to-learn activities on comprehension as compared with traditional writing tasks across different science topics and different educational levels (Chen, Hand, & McDowell, 2013; Gunel et al., 2007; Martin & Hand, 2009). Hand and colleagues suggest that this effect occurs because writing-to-learn activities promote deep-learning strategies. Following the self-regulation view in writing-to-learn (see Nückles et al., 2009), writing may serve learners as a beneficial medium for realizing the circle of self-regulated learning which consists of planning, organization and elaboration, and comprehension monitoring. Planning includes the setting of individual goals and the selection of suitable learning strategies which then are applied during the writing process. Such strategies include organizing the learning contents (organization strategies) and linking new information to the learners' prior knowledge, for example, by creating analogies or by critically reflecting on the learning contents (elaboration strategies). Monitoring helps learners detect gaps in comprehension and planning remedial strategies in order to eliminate these gaps. In this way, the writing process moves on to a new iteration of the circle of self-regulated learning (Nückles et al., 2009; Zimmerman, 2002).

In a learning journal, learners are typically asked to rework previously presented learning contents in a self-regulated way. Learners may choose important and interesting aspects of a topic, organize the content by creating an outline, elaborate on the content by linking abstract ideas to concrete experiences and by articulating personal thoughts. While writing, learners ideally plan, monitor and reflect on their learning process continuously. Accordingly, journal writing can be regarded as a promising method to encourage learners' application of cognitive and metacognitive strategies (Berthold, Nückles, & Renkl, 2007; Glogger et al., 2012; Nückles et al., 2009, 2012). However, learners often do not engage in beneficial learning activities spontaneously. Therefore, providing prompts turned out to be effective in supporting journal writing instructionally (Berthold et al., 2007; Nückles et al., 2009). Prompts are questions or hints that trigger learners' application of learning strategies which they are, in principle, capable of applying, but do not spontaneously demonstrate, or demonstrate to an unsatisfactory degree. In previous research on journal writing, students were typically prompted to engage in organization, elaboration, and comprehension monitoring (see Berthold et al., 2007; Nückles, Dümer, Hübner & Renkl, 2010). Compared to other learning tasks, such as summary writing or concept-mapping (Novak, 2010; Franzke, Kintsch, Caccamise, Johnson, & Dooley, 2005), a potential advantage of prompted journal writing might especially be that learners are encouraged to engage in a balanced application both of cognitive and metacognitive learning strategies. Combined application of cognitive and metacognitive strategies has repeatedly been proven to be a strong predictor of deep comprehension and long term retention in previous studies on journal writing (see Glogger et al., 2012; Nückles et al., 2009). Positive effects of journal writing on learning outcomes were further found in laboratory as well as in field studies in different domains including biology, mathematics and psychology (e.g., McCrindle & Christensen, 1995; see Nückles, et al., 2012 for an overview).

Besides promoting comprehension and retention of learning contents, learning journals offer ample opportunities to reflect on the personal relevance of a topic. The more the learners perceive the learning contents as personally relevant, useful, and interesting, the more they will engage in effortful and persistent learning (Schunk & Zimmerman, 2008; Wigfield, Eccles, Roeser, & Schiefele, 2008; Wolters, 2003). However, far too often, learners find it difficult to see the personal utility of the contents in science education (Assor, Kaplan, & Roth, 2002). Consequently, learners tend to show little identification with the learning contents and are weakly motivated to invest substantial effort in high quality learning processes (Kirby & Lawson, 2012) and in developing reflected ideas for solving complex science problems (Belland, Kim, & Hannafin, 2013). Journal writing could have positive effects on one's identification with the learning content.

Schmidt et al. (2012) investigated the potential of journal writing to increase learning motivation by supporting learners in reflecting on the personal relevance of scientific topics. In their study, thirteen- and fourteen-year-old high school students

received a personal-utility prompt in addition to a standard combination of cognitive and metacognitive prompts for writing a learning journal in a biology course. The students in the control condition also received cognitive and metacognitive prompts but no personal-utility prompt. Results showed that the personal-utility prompt successfully supported the students in reflecting on the personal utility of the learning contents. After a period of six weeks of journal writing (weekly entries), students in the personal-utility prompt condition reported higher degrees of interest in biology and achieved better comprehension scores compared to the students who had had no personal-utility prompt available. Thus, by reflecting on and writing about the relevance and personal utility of a topic, learners could more easily identify with the learning contents and thereby increase their interest in the subject.

Recognition of personal value as well as deep comprehension may possibly also be important facilitators for critical reflection (Rigby, Deci, Patrick, & Ryan, 1992). Critical reflection is a core goal in science education and an important aspect of scientific literacy (Ash & Clayton, 2009; Ford & Yore, 2012). Critical reflection involves providing reasons and evidence for an idea, articulating questions, contrasting ideas, challenging simplistic solutions, and making new understanding accessible for evaluation (Ash & Clayton, 2009). Through critical reflection, learners may develop a personal stance on controversial, scientific issues (Driver, Newton, & Osborne, 2000; Ford & Yore, 2012), for example, should I agree with antibiotic medication? or should I have an organ donor card? To foster learners' critical reflection, learners are typically encouraged to engage in argumentative writing. Learners are asked to provide claims and counterclaims, back them up with reasons or evidence, and integrate them into an overall final conclusion (Nussbaum & Schraw, 2007; Zohar & Nemet, 2002). Argumentation was shown to have positive effects on learners' comprehension (Wiley et al., 2009; Wiley & Voss, 1999). However, previous research shows that argumentative writing is a challenging task for many students (Nussbaum & Schraw, 2007) because writing a coherent argumentative essay that presents arguments and counterarguments in a balanced and integrated fashion is a rhetorically demanding task for novice writers. Thus, intervention studies on argumentative writing (see Nussbaum, 2008; Nussbaum & Schraw, 2007) that focused on the argumentative structure of the text yielded rather superficial improvements in the quality of argumentation. In contrast to genres like argumentative essays, learning journals specifically do not have a fixed rhetorical structure. The writing of a learning journal is a free and expressive way of writing that allows the learner to personally select which aspects of a learning episode require deeper reflection. Hence, because the rhetorical demands are low, the writing of learning journals might be especially suited for learners with comparatively little writing expertise to facilitate critical reflection on controversial scientific issues.

2. The present studies

Against this background, the aim of the present studies was to investigate the effects of journal writing on learners' comprehension of, interest in, and ability to critically reflect on scientific issues. To this end, we conducted two quasi-experimental studies in regular high school courses with biology topics. In the first study, we compared the effects of journal writing to more traditional homework writing tasks on comprehension, interest (i.e., learning motivation) and critical reflection in secondary science education. In the second study, we investigated whether the positive effects of journal writing could be further improved by directly prompting the learners to write and reflect about the personal relevance of the learning contents.

We conducted the first quasi-experimental study in a German middle school in two biology courses with 7th graders. The students in one course engaged in prompted journal writing. The students in the other course completed several other homework tasks (writing a summary, answering questions, creating a concept-map) during the intervention period. Those control students were encouraged to use the tasks for repeating and understanding the main ideas of the taught learning contents. The students in the journal writing course, in contrast, received several prompts (see Table 1) that stimulated them to employ cognitive and metacognitive strategies when writing a learning journal entry. As our previous studies in journal writing had repeatedly shown that especially a combination of cognitive and metacognitive prompts strongly fostered deep comprehension and sustained retention (see Glogger et al., 2012; Nückles et al., 2009; 2010), we expected that the students in the journal writing condition receiving both cognitive and metacognitive prompts would outperform the students in the traditional homework tasks condition with regard to deep comprehension of the learning contents. As deep comprehension can be regarded as an important prerequisite both of raised interest and critical reflection, we further predicted that the students in the learning journal condition would consider the learning contents as more interesting and, as a consequence, they would be more motivated and able to develop a reflected position on a topic-related, controversial issue than the students in the traditional homework condition. However, rather than expecting a perfect argument, we predicted that the students in the journal writing condition would be better able to consider the complex and controversial nature of the learning content, for example, by questioning a position or by formulating alternative ideas.

The aim of the second study was to investigate whether the predicted effects of the first study could be further boosted by explicitly prompting the students to reflect on the personal relevance of the contents they wrote about in their learning journal. Accordingly, we conducted another quasi-experimental study in two philosophy courses of a German high school (10th grade). All students were asked to write a learning journal over several weeks and they were prompted with a combination of cognitive and metacognitive prompts. In the experimental condition, however, the students additionally received a personal-utility prompt that asked them to write about the personal relevance of the topic. We expected that students who were asked to write

about the personal relevance of a topic would regard this topic as more valuable and interesting. Increased interest should foster their intention to acquire as much knowledge as possible and thus increase their willingness to critically reflect on a controversial topic-related issue.

3. STUDY 1

In the first quasi-experimental field study, we compared middle school students who engaged in journal writing supported by a combination of cognitive and metacognitive prompts with students who completed traditional homework writing tasks as a follow-up to their regular biology lessons.

3.1 Method

Participants and design

Forty-six middle school students (7th grade, 13-14 years old) participated in the quasi-experimental field study. They were members of two biology classes of a secondary school with about 800 students and 75 teachers in a small town in Southern Germany (about 12.000 residents). During the surveyed time span, the biology lessons were primarily about basic concepts in immunology (e.g., the functioning of white blood cells). Immunology was integrated into the broad topic of human biology and followed by sexual education. In one class, the students reflected on the learning contents of their biology lessons by writing regular learning journal entries as follow-up course work during the surveyed time span of three weeks, resulting in three entries in total (learning journal condition; n = 21; 10 girls, 11 boys; 4 participants not born in Germany). In the other class, the students did traditional homework (see material subsection) as follow-up course work (traditional homework condition; n = 25; 15 girls, 10 boys; 5 participants not born in Germany).

Given that the students belonging to a particular class were assigned to either the experimental or control condition as a whole, our design was quasi-experimental. Therefore, it is particularly important to control for pretest differences. All participating students had received biology instruction for two years with the same learning contents. Furthermore, they were comparable with regard to their gender, X^2 (N = 46) = 0.71, ns, age, $F(1, 44) = 0.03$, ns, and ethnic background, X^2 (N = 46) = 0.01, ns. The results of pretest score analyses indicated that the students were comparable regarding their prior knowledge scores, $F(1, 44) = 0.38$, ns, and their interest scores in the pretest, $F(1, 44) = 1.36$, ns. Furthermore, we kept both conditions as comparable as possible; the same teacher taught the same contents in both classes during the intervention period. He also used the same didactic methods and learning materials in both classes. Immediately and eight weeks after the intervention period, students took part in a posttest, encompassing comprehension tasks, a critical reflection task, and their interest in the topic.

Materials

As the participating 7th-grade students were inexperienced in journal writing, we provided them with a brief instruction (300 words) on how to write a learning journal. The instruction emphasized the use of learning strategies during writing. To facilitate journal writing, the instruction included two cognitive prompts stimulating elaboration and organization strategies and two metacognitive prompts stimulating comprehension monitoring and planning of remedial strategies (see Table 1). Further training was not provided because previous research with similar students showed that students at that age are able to use journal writing beneficially without intense training (e.g., Schmidt et al., 2012). The instruction emphasized that learning journals should help students better understand the topics. The students wrote 82 words per entry on average (SD = 8.50). To ensure that our treatment had the intended effects, a research assistant coded the number of statements in the learning journals that indicated the use of a learning strategy (for a detailed description of the coding system see Study 2). In a learning journal entry, on average, the students elicited M = 2.85 (SD = 1.01) organization strategies, M = 1.85 elaboration strategies (SD = 0.81), and M = 1.52 metacognitive strategies (SD = 0.85). These frequencies appear to be rather low. However, the students used the journal writing often simply to reproduce and repeat facts presented during the lessons instead of organizing and elaborating the content, and monitor their comprehension. Nevertheless, as they evidently demonstrated use of the prompted learning strategies to some extent, we concluded that the treatment was delivered as intended.

Table 1. Prompts used in the writing instructions

Prompts	
Cognitive Prompts (Organization and Elaboration)	How can you structure and summarize the contents in a meaningful way?
	Which examples can you think of that illustrate, confirm or conflict with the learning contents?
Metacognitive Prompts (Monitoring and Planning of Remedial Strategies)	Which main points do you now understand, and which haven't you understood?
	What possibilities do you have to overcome your comprehension problems?
Personal-utility Prompt	Why is the learning material personally relevant for you at present or in future out of school?

Note. Students in Study 1 and students in the standard prompts condition in Study 2 received only the cognitive and metacognitive prompts. Students in the personal-utility prompt condition in Study 2 additionally received the personal-utility prompt.

The students in the traditional homework condition had to reread the text covered by the lesson and summarize it in the first week. In the second week, they had to excerpt main ideas from the lesson and visualize them in a concept-map, draft or picture. In the third and last week, they were asked to answer questions about the topic. We assumed that these tasks would primarily trigger organization strategies because, in at least two of the tasks (i.e., summarizing and concept mapping), the students had to identify main ideas of the text and the lesson, paraphrase them (in summarizing), and organize them in a structured way (in concept-mapping). Both tasks did not explicitly encourage critical reflection or metacognition although they did not foreclose such activities either. Answering predefined questions required the students mainly to reproduce acquired knowledge. Among the traditional homework tasks, answering questions was perhaps most likely to encourage metacognitive experiences, especially if students experienced difficulties in answering a question. To ensure a meaningful use of these learning activities, students received a short introduction on summary writing (about 100 words) and concept-mapping (about 200 words).

Dependent measures

To assess the students' comprehension of immunology, we designed a comprehension test based on the guidelines for the biology curriculum of German high schools. To ensure curricular validity, the test was informally assessed by two experienced teachers of biology (teaching experience 8 and 10 years). One of the teachers taught both classes to ensure that the content of the tasks was really part of the biology lessons in which the students participated. As the guidelines in the curriculum focus on scientific literacy, for example the ability to explain biological phenomena scientifically, we designed explanation tasks that measured comprehension (e.g., "Explain the similarities and differences between an infection with bacteria or viruses" or "Explain what happens in the organism after an active immunization."). Thus, the students had to apply their acquired knowledge in order to generate explanations. A trained research assistant, who was blind to the experimental conditions, compared the students' answers with reference answers given by the teacher and counted the number of correct statements. A maximum of 18 points could be achieved when all the seven test items were answered correctly.

To assess the students' ability to critically reflect on a complex controversial issue, students were asked to write a short comment about the question, "Should people who show symptoms like coughing and fever be medicated with antibiotics? Please write a short comment and give reasons for your answer." The topic is sufficiently complex regarding its causal structure. The relevant information necessary to reflect on the topic was taught during the intervention period. In this way, we ensured that different positions were inherent and could be adopted on this issue and controversially discussed. Two research assistants, who were blind to the experimental conditions, assessed how the students reflected on their position. In a first step, they counted the number of all statements that were presented in answering the question, and justified,

or criticized a position as indicators of critical reflection. In the second step, they rated the overall argumentation quality of the students' comments on a five-point rating scale ranging from 1 (low quality) to 5 (high quality). Table 2 illustrates the levels of this rating scale. Interrater-reliability as indicated by the intra-class coefficient was very good, ICC = .92.

Table 2. Description of the quality ratings

Level	Description
1	Students neither provided any reasons to justify a position nor did they criticize a position.
2	Students provided few and low elaborated arguments to justify a position. They did not criticize a position.
3	Students provided at least one elaborated argument for one position and questioned the other position.
4	Students provided at least one elaborated argument for each position and reflected on the validity of the arguments.
5	Students provided elaborated arguments for both positions and integrated them into a reflected overall conclusion by weighing, synthesizing or refuting arguments.

To assess the students' interest in immunology, we translated and adapted items from the Intrinsic Motivation Inventory (IMI; Deci & Ryan, 2006) for the purposes of our study. The short questionnaire included five self-report items to be rated on a 5-point scale, ranging from 1 (very low degree of agreement) to 5 (very high degree of agreement) (e.g., "I enjoy discussing immunological issues very much."). High interest scores indicated that students highly valued the topics and enjoyed discussing controversial issues in immunology. The internal consistency was satisfactory, Cronbach's $\alpha = .77$.

Procedure

The whole study lasted thirteen weeks. In the first week, we asked the students in both conditions to take part in a pretest assessing their interest in and comprehension of immunology. In the 2nd, 3rd, and 4th week, students were taught in immunology two lessons per week, with the same contents, methods and materials taught by the same teacher in both conditions. The students in both conditions were asked to do their homework (journal writing or traditional tasks) once a week after the two biology lessons that were taught en bloc, resulting in three journal entries or three traditional homework tasks. The students in both conditions received a brief instruction on how to do their homework in the following weeks on a sheet of paper. To ensure that the students did their homework, the teacher collected it every week. To keep the implementation of the intervention as objective and comparable as possible across

individual students, we decided against providing feedback. The topic of immunology was not finished by a teacher-graded test. Thus, the students in both conditions had no explicit, extrinsic motivator.

In the 5th week, all students took part in the posttest, which was the same as the pretest. Between the 5th and 13th week, students were taught in biology topics other than immunology (such as sexual education). Nevertheless, the students in both classes were taught in a similar way by the same teacher, with the same methods and materials and with the same homework tasks (traditional in both classes). In the 13th week, the posttest was repeated and complemented with a critical reflection task. We ensured equal testing conditions. All tests were administered as paper-pencil tests in the regular biology lesson by a research assistant, who was blind to the experimental conditions. After the end of the study, that is, after the delayed posttest, the students in the traditional homework condition were introduced to journal writing to enable them to profit from journal writing as beneficial follow-up course work as well.

3.2 Results and Discussion

For an overview, Table 3 includes the means and standard deviations separately for the two experimental conditions. As an effect size measure of group differences, we used partial η^2 qualifying values < .06 as small effect, values in the range between .06 and .13 as medium effect, and values > .13 as large effect (see Cohen, 1988).

Table 3. Descriptive statistics of dependent variable separately for the three points of measurement and each experimental condition in Study 1

		Experimental Conditions					
		Traditional Homework			Journal Writing		
		Pretest	Posttest	Delayed Posttest	Pretest	Posttest	Delayed Posttest
Variables		*n* = 25	*n* = 25	*n* = 25	*n* = 21	*n* = 21	*n* = 20
Comprehension[a]	*M*	5.66	5.88	6.42	6.05	9.55	9.90
	SD	1.81	2.47	2.35	(2.46)	3.22	3.61
Overall reflection[b]	*M*			1.25			1.95
	SD			0.74			0.50
Reflection Quality[c]	*M*			2.38			2.86
	SD			0.92			0.79
Interest [c]	*M*	2.54	2.84	2.65	2.56	3.06	3.23
	SD	0.57	0.65	0.51	0.72	0.49	0.58

Note. a is the average number of correct answers (a maximum of 18 points was possible); b is the average number of statements coded as critical reflection; c was rated on a five point rating scale (1-5).

Comprehension

To investigate the students' increase in comprehension, we conducted a repeated measure analysis of variance with the pretest and immediate posttest as within-subjects factor and experimental condition as between-subjects factor. The results showed a main effect for time, $F(1, 44) = 31.09$, $p < .01$, partial $\eta^2 = .41$, indicating a general increase in comprehension in both conditions, and a main effect of experimental condition, $F(1, 44) = 9.34$, $p < .01$, partial $\eta^2 = .18$. However, this main effect has to be qualified by the significant interaction effect between time and experimental condition, $F(1, 44) = 24.17$, $p < .01$, partial $\eta^2 = .36$. The interaction effect revealed that the students in the learning journal condition learned significantly more during the intervention period than the students in the traditional homework condition (see Table 3).

A delayed posttest should show whether the differences between the two conditions remained stable in the weeks after the intervention period. Therefore, we conducted a second repeated measure analysis of variance with the immediate posttest and the delayed posttest as within-subjects factor and experimental condition as between-subjects factor. The results showed a significant main effect of time, $F(1, 43) = 1.94$, $p < .05$, partial $\eta^2 = .04$, indicating a small increase in the knowledge scores between the second and the third point of measurement. The main effect of experimental condition was significant as well, $F(1, 43) = 19.29$, $p < .01$, partial $\eta^2 = .31$. The interaction effect between time and experimental condition failed statistical significance, $F(1, 43) = 0.04$, ns, partial $\eta^2 = .00$. Students in both conditions improved their comprehension to some extent. However, the group differences remained stable, that is, the students in the journal writing class clearly outperformed the students in the traditional homework class both in their comprehension and also their long-term retention about two months later.

Critical reflection

To investigate whether the students in the journal writing class were better able to reflect on a complex, controversial issue in immunology than the students in the traditional homework class – even eight weeks after the intervention period –, we conducted a one-factorial analysis of variance with the number of all statements students used to question, justify or criticize a position (i.e., overall reflection, see Table 3) as dependent and the experimental conditions as independent variable. The difference between the experimental conditions failed to reach statistical significance, $F(1, 43) = 3.48$, ns, partial $\eta^2 = .08$. However, a second one-factorial analysis of variance with the quality ratings as dependent and the experimental conditions as independent variable revealed a significant main effect of experimental condition, $F(1, 43) = 13.61$, $p < .01$, partial $\eta^2 = .24$. Students in the journal writing condition produced comments with a higher quality of critical reflection as compared with the students in the traditional homework condition (see Table 3).

Interest

To detect a potential increase in students' interest during the intervention period, we conducted a repeated measure analysis of variance with the pre- and posttest interest scores as within-subjects factor and the experimental conditions as between-subjects factor. The results indicated a main effect of time, $F(1, 44) = 8.11$, $p < .01$, partial $\eta^2 = .16$. The main effect of experimental condition, $F(1, 44) = 1.00$, ns, partial $\eta^2 = .01$, as well as an interaction effect between time and experimental condition, $F(1, 44) = 0.50$, ns, partial $\eta^2 = .02$, failed to reach statistical significance. Thus, the interest of all students increased when the lesson topic was immunology (see Table 3). There was no significantly stronger increase in the journal writing condition. Again, the delayed measurement would show whether the effects on interest remained stable in the weeks after the intervention period. Therefore, we conducted a second repeated measure analysis of variance with the second and third interest scores as within-subjects factor and experimental condition as between-subjects factor. There was no significant main effect of time, $F(1, 43) = 0.00$, ns, partial $\eta^2 = .00$. However, the main effect of experimental condition, $F(1, 43) = 7.30$, $p < .05$, partial $\eta^2 = .15$, and the interaction effect between time and experimental condition $F(1, 43) = 4.78$, $p < .05$, partial $\eta^2 = .10$, were significant. While the students' interest of immunology in the traditional homework condition decreased after finishing immunology in school and dropped back to the pretest level, the students' interest in the journal writing condition further increased after the intervention period. Thus, journal writing seemed to have stronger positive long-term effects on students' interest in biological topics than follow-up course work such as summarizing, concept-mapping or question answering.

Relations between comprehension, interest and critical reflection.

As we hypothesized that journal writing would improve comprehension and improved comprehension could further increase students' interest in the topic, we conducted a mediation analysis (see Baron and Kenny, 1986) to test whether comprehension indeed mediated interest in terms of a causal variable. To test this hypothesis, we regressed, in a first step, the dependent variable (interest score in the delayed test) on the predictor (experimental condition). The results indicated a significant effect of the experimental condition on learners' interest, $t(43) = 3.59$, $ß = .48$, $p < .01$. In the second step, the potential mediator (comprehension score in the delayed test) was regressed on the predictor (experimental conditions). The results indicated a significant effect of the experimental condition on comprehension, $t(43) = 3.90$, $ß = .51$, $p < .01$. In the third step, the dependent variable (interest) was regressed simultaneously on both the predictor (experimental condition) and the mediator (comprehension). The results indicated a significant effect of comprehension on interest, $t(42) = 2.43$, $ß = .36$, $p < .05$, while the effect of experimental condition was lower compared to step 1 and failed to reach statistical significance, $t(42) = 2.01$, $ß = .30$, ns. According to Baron and Kenny, this pattern of results indicates complete mediation, that is, the students'

superior comprehension in the journal writing condition was a causal factor in raising the students' interest in immunology.

Furthermore, we had hypothesized that journal writing would contribute to improving critical reflection by improving both comprehension and interest. To test this hypothesis, we conducted a second mediation analysis with reflection quality as dependent variable, experimental condition as independent variable, and both interest and comprehension as mediators. The effect of the experimental condition on reflection quality was significant in step one, $t(43) = 3.69$, $ß = .49$, $p < .01$. The effect of the experimental condition on interest and comprehension has already been demonstrated in the previous mediation analysis. The third step of the mediation analysis showed that only interest had a significant effect on reflection quality, $t(41) = 2.43$, $ß = .37$, $p < .05$, while the effects of comprehension, $t(41) = 0.56$, $ß = .09$, ns, and experimental condition, $t(41) = 1.52$, $ß = .25$, ns, failed to reach statistical significance. According to Baron and Kenny, we can therefore assume that interest acted as a mediator between experimental condition and reflection quality. When taking the two mediation analyses together, the results provided evidence that journal writing improved comprehension, and improved comprehension lead to a raised interest in biology. The increased interest resulted in improved critical reflection while comprehension by itself did not have an additional, positive effect on critical reflection.

4. STUDY 2

The first study showed that journal writing proved to be a more effective learning-to-write activity as compared with traditional writing assignments in science education, because it promotes deep comprehension, interest in the topics and critical reflection. However, the positive effects of journal writing on students' interest occurred with a delay. Therefore, we aimed to promote students' interest explicitly by prompting them to reflect on the personal utility of the topic addressed in their learning journal. This encouragement was expected to result in an increase in the perceived value of the topic and in increased interest in the topic. Furthermore, as found in Study 1, we expected that improved interest would promote critical reflection. To this end, we compared students, who engaged in journal writing prompted by a combination of cognitive and metacognitive prompts, with students who additionally received a personal-utility prompt. This prompt asked the students to think and write about the personal utility or relevance of the topics.

4.1 Method

Participants and design

Twenty-four high school students (10^{th} grade, 16-17 years old) participated in the quasi-experimental field study. They were members of two philosophy classes of a secondary school with about 1100 students and 81 teachers in a small town in Southern Germany

(about 25.000 residents). To investigate our research question, we used a one factorial between-subject design comprised of two experimental conditions. In one class, students' journal writing was supported with the same combination of cognitive and metacognitive prompts used in Study 1 (standard prompts condition; n = 11, 6 boys, 5 girls, 1 girl not born in Germany). In the other class, students received a personal-utility prompt in addition to the combination of cognitive and metacognitive prompts (personal-utility prompt condition; n = 13, 6 boys, 7 girls, 1 boy and 1 girl not born in Germany). Given that the students belonging to a particular class as a whole were assigned to either the experimental or the control condition, our design was quasi-experimental. However, all participating students had received philosophy instruction for three years with the same learning contents. Furthermore, the students were comparable with regard to their gender, X^2 (N = 24) = 0.17, ns, age, $F(1, 22) = 0.22$, ns, ethnic background, X^2 (N = 24) = 0.22, ns, grades in philosophy, $F(1, 22) = 0.58$, ns, and interest scores before the intervention period, $F(1, 22) = 0.05$, ns. During the surveyed time span of six weeks, the students learned about "biology and ethical decisions" (e.g., pros and cons of genetic manipulation). They were taught by the same philosophy teacher with the same teaching methods and materials. At the end of each philosophy lesson (90 minutes in total), the students had 20 minutes time to reflect on the learning contents by writing a learning journal entry. Thus, the learning journals were integrated in the lessons to control for time on task to optimize the comparability of individual students. On average, one learning journal entry consisted of 118.77 words (SD = 48.50). The dependent variables encompassed students' comprehension of the biological concepts, the quality of their critical reflections and their motivation (value and interest) to engage in controversial issues. Furthermore, we analyzed frequencies of different types of learning strategies applied in the learning journals.

Instruments and measures

The writing assignment was the same as the writing assignment that was used in Study 1. Students in the personal-utility-prompt condition additionally received a personal-utility prompt that asked them to think about the personal relevance of the topic. The prompts for the two conditions are illustrated in Table 1.

To assess whether students found the topics valuable and interesting, they were asked to answer a motivation questionnaire. We used ten translated and adapted items from value and interest scale of the intrinsic motivation inventory (IMI; Deci & Ryan, 2006).Items included, for example, „The biology-related topics in Philosophy were very interesting" or "I think it is very important to talk about ethical issues in biology". The items had to be rated on a 7-point rating scale, ranging from 0 (very low degree of agreement) to 6 (very high degree of agreement). As the internal consistency was high, Cronbach's $\alpha = .82$, we computed an average intrinsic motivation score for every student for pre- and posttest.

In the comprehension test, the students answered questions about the legal grounds and standard procedures for organ donation. The topic was discussed in the philosophy

courses. Students could reach six points at maximum when their answers matched the reference answers constructed by their philosophy teacher.

To assess the quality of critical reflection, the students were asked to write a short comment about the pros and cons of organ donation. The total number of words in the students' comments was in average 102.54 words (*SD* = 28.90). To assess critical reflection, we counted the number of topic-related statements that were backed up with reasons or evidence (founded claims) and that were not backed up with reasons or evidence (unfounded claims). In addition to these rather quantitative indicators of critical reflection, we rated the comments on a five-point rating scale from 1 (low quality) to 5 (high quality) in the same way as in Study 1.

Furthermore, we aimed to identify learning strategies that could have been responsible for differences in the posttest. Therefore, we used and adapted a coding scheme developed by Nückles and colleagues (2009). Two independent raters coded the learning journals on the granular level of individual statements. They categorized individual statements as organization, elaboration or metacognition. For example, statements that highlighted the main points of the topic and their interrelations were coded as indicators of organization (e.g., students underlined important terms or highlighted them in different colors). As elaboration, we coded statements in which students associated the new content with their prior knowledge, for example, by generating examples, analogies or illustrations. We differentiated between statements addressing considerations regarding the personal relevance and importance of the topic (e.g., "It is important for me to know how medical practitioners diagnose a brain death.") (personal-utility statements) from purely cognitive elaboration (e.g., "The human immune system can be compared with the protective walls of a castle."), because the personal-utility statements included motivational aspects and were triggered with the personal-utility prompt only in one condition. We coded statements related to comprehension monitoring (e.g., "I have difficulties in understanding the differences between active and passive immunizations.") and planning of remedial strategies (e.g., "I will rework the course materials and ask the teacher if I cannot understand everything.") as metacognitive strategies. Statements that included only information given in the lessons were not coded. Inter-rater reliabilities as determined by Cohen's Kappa ranged from .88 for organization to 1.0 for personal utility.

Procedure

The whole intervention lasted eight weeks. One class was assigned to the standard prompts condition while the other class was assigned to the personal-utility prompt condition. In the first week, we asked the students in both conditions to fill out the intrinsic motivation questionnaire. Afterwards, the students received the journal writing instruction. The instruction was identical for all students except for the personal-utility prompt, which was only given to the students in the personal-utility prompt condition. In the following six weeks, the students were asked to write a learning journal entry at the end of the two philosophy lessons that were taught en bloc in a similar way in both

classes. The students wrote the journal entries in class and had 20 minutes time. The students received no feedback for the six journal entries. In the last week, all students took part in the posttest in class. The students assessed their interest with the same questionnaire used in the pretest. Furthermore, they were asked to answer the questions of the comprehension test and the critical reflection task. All tests were administered as paper-pencil tests in the regular philosophy lesson by a research assistant, who was blind to the experimental conditions. The topic "human biology and ethical decisions" was not finished by a teacher-graded test. Thus, the students had no explicit, extrinsic motivator. After the end of the study, the students in the control condition also received the instruction with the personal-utility prompt to enable them to benefit from this enhanced instruction when writing future learning journals.

4.2 Results and Discussion

For an overview, Table 4 includes the means and standard deviations separately for the two experimental conditions.

Table 4. Descriptive statistics of dependent variable separately for the experimental conditions

	Standard prompts condition n = 11		**Personal-utility prompt condition n = 13**	
	M	*SD*	*M*	*SD*
Interest in Pretest[a]	4.61	0.48	4.68	0.72
Interest in Posttest[a]	4.45	0.53	4.99	0.71
Comprehension[b]	4.36	1.21	4.85	1.07
Critical Reflection				
Unfounded Claims[c]	1.45	1.04	1.54	1.05
Founded Claims[c]	1.27	0.79	2.31	0.75
Quality Rating[d]	2.36	0.92	3.69	0.75
Statements in the learning journals				
Personal Utility[c]	0.73	0.61	1.58	0.67
Cognitive Elaboration[c]	1.00	0.68	1.41	0.70
Organization[c]	1.06	0.46	0.72	0.34
Metacognition[c]	1.11	0.86	2.38	1.94

Note. a was rated on a seven point rating scale (0-6); b is the average number of correct answers (a maximum of 6 points was possible); c is the average number of statements per entry; d was rated on a five point rating scale (1-5).

Interest

To investigate whether students in the personal-utility prompt condition considered the topics as more valuable and interesting than the students in the standard prompts

condition, we conducted a repeated measure analysis of variance with the pre- and posttest interest scores as within-subjects factor and experimental condition as between-subjects factor. The results indicated neither a significant main effect of time, $F(1, 22) = 0.69$, ns, partial $\eta^2 = .03$, nor a main effect of experimental condition, $F(1, 22) = 1.55$, ns, partial $\eta^2 = .07$. However, as predicted, the interaction effect between time and experimental condition was significant, $F(1, 22) = 6.83$, $p < .05$, partial $\eta^2 = .24$. While interest remained stable in the standard prompt condition, interest increased during the intervention period in the personal-utility prompt condition (see Table 4). Thus, by prompting writing about the personal utility of the learning contents, we could directly improve students' interest in the topics.

Comprehension

To investigate whether students in the personal-utility prompt condition reached a better comprehension level than the students in the standard prompts condition, we conducted an analysis of variance with the comprehension scores as dependent and experimental condition as independent variable. The results indicated no significant differences between the experimental conditions, $F(1, 22) = 1.08$, ns, partial $\eta^2 = .05$. However, the results showed high comprehension scores in both conditions (see Table 4).

Critical reflection

In the next step, we analyzed the students' comments about the pros and cons of organ donation. First, we investigated whether students provided a final conclusion. In the personal-utility prompt condition, 77% of the students provided a final conclusion (2 students decided against and 8 students for an organ donor card). In the standard prompts condition, 36% of the students provided a final conclusion (all for an organ donor card). The difference between the conditions was significant, X^2 (N = 24) = 4.03, $p < .05$. To investigate whether the students provided a different number of founded and unfounded claims in the experimental conditions, we conducted two analyses of variance with the number of founded and unfounded claims used to support a position as dependent variables and experimental condition as independent variable. The results indicated that the number of unfounded claims did not vary across the conditions, $F(1, 22) = 0.04$, ns, partial $\eta^2 = .00$. For founded claims, however, we found a significant advantage for the personal-utility prompt condition, $F(1, 22) = 10.48$, $p < .01$, partial $\eta^2 = .33$ (see Table 4). Finally, we investigated the quality of critical reflection by conducting a third analysis of variance with the quality ratings as dependent and experimental condition as independent variable. The results indicated a significant advantage for students in the personal-utility prompt condition, $F(1, 22) = 15.11$, $p < .01$, partial $\eta^2 = .41$. We can summarize that the students in the personal-utility prompt condition reached a higher quality in critical reflection and provided a conclusion more often than the students in the standards prompts condition.

Learning strategies in the learning journals.

To investigate whether learning strategies used by the students in the personal-utility prompt condition differed from the strategies used by the students in the standard prompts condition, we conducted analyses of variance with the types of learning strategies as dependent variables and the experimental conditions as independent variable. The results indicated that the students in the personal-utility prompt condition wrote significantly more often about the personal relevance of the topic than the students in the standard prompt condition, $F(1, 22) = 10.46$, $p < .01$, partial $\eta^2 = .32$. Thus, the intervention worked as intended. The students in the standard prompts condition used significantly more strategies to organize the content than the students in the personal-utility prompt condition, $F(1, 22) = 4.49$, $p < .05$, partial $\eta^2 = .17$ (see Table 4). We neither found significant differences for purely-cognitive elaboration strategies, $F(1, 22) = 2.09$, ns, partial $\eta^2 = .09$, nor for metacognitive strategies between the experimental conditions, $F(1, 22) = 4.07$, ns, partial $\eta^2 = .16$.

Relations between interest and critical reflection

As the implementation of a relevance prompt both improved interest and critical reflection, we were curious whether interest would mediate the effect of the relevance prompt on critical reflection. To test this mediation hypothesis, we conducted a mediation analysis with the quality ratings of reflection as dependent variable, experimental condition as independent variable and interest as mediator. The results indicated a significant effect of experimental condition (predictor) on the students' quality of critical reflection (dependent variable), $t(22) = 3.89$, $ß = .64$, $p < .01$, and a significant effect of experimental condition on interest in the posttest (mediator), $t(22) = 2.08$, $ß = .41$, $p< .05$. The results of regressing the quality of critical reflection simultaneously on both the predictor (experimental condition) and the mediator (interest in the posttest) yielded a significant effect of interest on critical reflection, $t(21) = 2.27$, $ß = .37$, $p < .05$. Furthermore, the effect of experimental condition was lower than in step 1, however it still reached statistical significance, $t(21) = 2.96$, $ß = .49$, $p < .01$. Therefore, we can assume partial mediation.

The results can be summarized as follows. First, as predicted, we found that the students in the personal-utility prompt condition used their learning journals more intensively to self-explain the relevance of the learning contents by drawing on their own experiences and by constructing personal examples. The consideration of the personal relevance of a topic during journal writing led to increased interest in the topic and improved their critical reflection. While the students provided significantly more founded claims and engaged in high quality reflections, we found no significant effect on students' comprehension. As almost all students reached high scores in the comprehension task, this suggests that a standard set cognitive and metacognitive prompts was enough to reach good results in the comprehension test. However, the

effect on interest and critical reflection could be further promoted by inviting students to think about the personal relevance of a topic.

5. General Discussion

In the present article, we reported two quasi-experimental studies examining the effects of writing learning journals on learners' comprehension, their interest in the topic about which they had to write, and their ability to critically reflect on a topic-related issue. In Study 1, we compared journal writing with a combination of several other homework tasks. In Study 2, we investigated whether the positive effects of journal writing on interest and critical reflection could be further improved by prompting learners to write about the personal relevance of the learning contents. The results can be summarized as follows.

5.1 Journal writing promotes comprehension

In study 1, students who wrote regular learning journal entries clearly outperformed students who completed several other homework tasks, including concept-mapping, summarizing and question answering, on measures of comprehension, critical reflection, and self-reported interest in biology. Previous research suggests that concept mapping (Novak, 2010) and summarizing (Franzke et al., 2005) may trigger the application of cognitive learning strategies, particularly in organizational strategies (e.g., identification and structuring of main ideas). Nevertheless, a major advantage of the prompted learning journal method can be seen in the fact that the students were prompted to apply both cognitive and metacognitive strategies when writing their journal entries. Our previous research suggests that this combined and integrated application of cognitive and metacognitive strategies best supported learning outcomes because this approach made it most likely for students to to engage in the whole cycle of self-regulated learning (Glogger et al., 2012; Nückles et al., 2009). Accordingly, in writing a learning journal, learners could deeply process the learning contents by organizing and elaborating on the new information. In monitoring their comprehension, they could detect comprehension gaps and eliminate them by applying remedial cognitive strategies. It is possible that this integrated application of both cognitive and metacognitive processes during writing led to the superior comprehension scores in the journal writing condition in Study 1. The introduction of a personal-utility prompt in Study 2 had no further enhancing effect on comprehension, although the students in the two conditions used the learning journals in partly different ways. The students in the personal-utility condition produced more statements on the personal relevance of the topic, whereas the students in the standard prompts condition used more organization strategies to structure the contents. However, the students in both conditions engaged in cognitive and metacognitive processing to a similar extent and therefore achieved comparable and, at the same time, relatively high test scores in the comprehension test.

5.2 Prompted journal writing promotes learners' identification with the learning content

After engaging in traditional learning tasks, learners typically show little identification with the learning contents, typically cannot see the personal relevance and are weakly motivated to invest substantial effort in high quality learning processes (Kirby & Lawson, 2012). Regular journal writing as follow-up course work promoted students' interest in the topic more strongly than a sequence of concept mapping, summary writing and question answering (Study 1). There are two possible reasons that may explain the improved topic-related interest: (1) Students in the learning journal condition in Study 1 reached a superior conceptual understanding of the topic. Understanding a topic in depth may in itself contribute to increased identification with and increased interest in a topic (Schunk & Zimmerman, 2008). In line with this assumption, we found that the effect of the experimental condition on interest was mediated by learners' comprehension in Study 1. Furthermore, while the effect on comprehension was measured immediately, the positive effect of journal writing on learners' interest indeed occurred with a delay and presumably as a consequence of improved comprehension. (2) Writing about the personal relevance of the learning contents could have facilitated learners' identification with contents and thereby raised their interest. Accordingly, prompting students to write about the personal relevance of the learning contents in Study 2 significantly increased the learners' interest. More generally, learners who are able to recognize or increase the personal relevance of a topic are more likely to develop a commitment to life-long learning in the domain and to engage in discourse (Salomon & Perkins, 1998). Both are imperatives to autonomy and participation in our modern knowledge-based society, especially in subject domains like biology, where scientific knowledge accumulates and changes rapidly (Tunnicliffe & Ueckert, 2007).

5.3 Journal writing promotes critical reflection

In our previous research on journal writing, we investigated writing instructions that support the application of cognitive and metacognitive strategies in order to foster self-regulated learning (Berthold et al., 2007; Nückles et al., 2009; 2010; Schmidt et al., 2012). We found positive effects on the retention of facts, on comprehension measures and, more recently, also on students' learning motivation. It is beyond controversy that an important learning goal in education is to prepare middle and high school students for discussion of controversial issues and reflected reasoning (Ford & Yore, 2012). Therefore, it is important that learners engage in critical reflection, such as articulating questions and contrasting ideas (Ash & Clayton, 2009). In learning journals, learners are encouraged to freely develop their own ideas on learning contents and to discuss the pros and cons of complex issues according to their own interests and preferences (Nückles et al., 2009). However, previous research on journal writing has paid little attention to the effects of journal writing on critical reflection as a learning outcome. In the present studies with high school students in grades seven and ten, we therefore

investigated the effects of journal writing on learners' ability to critically reflect on a controversial issue. The results of our present studies show that learners who engaged in journal writing were better able to critically reflect on a controversial issue than learners who engaged in other, also potentially beneficial homework tasks (Study 1). Although a learner's deep comprehension of the topic can be regarded as a precondition for critical reflection, a key variable to promote critical reflection is the learner's motivation, that is, their topic-related interest. Learners who regard a topic as valuable and interesting are more likely to engage in critical reflection than learners who regard a topic as irrelevant or boring (Study 1 & 2). Thus, promoting learners' identification with learning contents had beneficial effects on learners' critical reflection of the contents, which is a core learning goal in science education (Ford & Yore, 2012). Thus, in a way, the results of our studies show that journal writing facilitated a domino effect from improved comprehension to superior interest and to critical reflection.

6. Study limitations

In this article, we presented two quasi-experimental, longitudinal field studies, in which we investigated journal writing in middle and high school science education. Inasmuch as our research was motivated both by theoretical and practical issues, we consider our studies as a kind of use-inspired research (Renkl, 2013). By implementing experimental interventions in an authentic field-context, we sought to maximize ecological validity without abandoning the methodical advantages of a laboratory study (e.g., controlling for potential confounds). However, assigning whole classes to an experimental condition, instead of randomizing the participants, produces some limitations. Although we controlled for pre-existing individual differences among the students as well as for potential confounds during the intervention studies (such as teacher, teaching methods, materials, and time on task), it is possible that other uncontrolled confounds influenced the effects. Therefore, it is necessary to replicate our results using real experimental designs in which students are randomly assigned to the experimental conditions.

Another limitation of our present studies refers to the generalizability of the results to other domains. For example, the personal relevance of human biology was rather easily discernible for learners. Thus, the large effects of the personal utility prompt we obtained in Study 2 might also partially be attributed to the accessibility of the content domain. Hence, it is an open question for further research whether we would be able to obtain comparable results in other domains that are further away from the learners' realm of experience (e.g., particle acceleration in physics). Therefore, it is necessary to replicate the results in different age groups and domains.

7. Conclusion

The present article adds to the available evidence for the beneficial and sustainable effects of journal writing on students' understanding of scientific concepts, critical

reflection on scientific issues, and development of interest, the last of which is important for developing a commitment to lifelong learning. In particular, journal writing proved to be more successful in comparison with other traditional writing assignments for middle and high school students' self-regulated learning in human biology. Therefore, writing learning journals is a specific method to encourage students in writing-to-learn that turned out to be a beneficial learning tool to promote the development of science literacy. Writing learning journals encouraged learners' application of beneficial learning strategies that contribute to comprehension, interest and critical reflection. The positive effect of journal writing on learners' interest was further supported by providing a writing instruction that not only focused on cognitive and metacognitive strategies but also on motivational aspects, such as the personal relevance of a particular topic. The present article showed that the potentials of journal writing clearly reach beyond merely supporting the comprehension and retention of the learning contents. Learning journals also seem to be a promising tool for fostering learners' interest and the acquisition of complex skills such as critical reflection.

Acknowledgements

The data of Study 1 were collected by Anja Gebhardt and the data of Study 2 by Eva-Maria Oberbusch as partial fulfillment of the requirements for their teacher degree at the University of Freiburg. All data were completely reanalyzed in preparation for this paper. The authors wish to thank Wesley Dopkins for his proofreading. Kristin Wäschle is now working for the CJD (Christliches Jugenddorf) in Offenburg, Germany.

References

Ash, S. L., & Clayton, P. H. (2009). Generating, deepening, and documenting learning: The power of critical reflection for applied learning. *Journal of Applied Learning in Higher Education, 1*, 25–48.

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology*, *72*(2), 261–278. doi:10.1348/000709902158883

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74*(1), 29–58. doi:10.3102/00346543074001029

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. doi:10.1037/0022-3514.51.6.1173

Belland, B. R., Kim, C., & Hannafin, M. J. (2013).A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist*, *48*(4), 243–270. doi:10.1080/00461520.2013.838920

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*(5), 564–577. doi:10.1016/j.learninstruc.2007.09.007

Chen, Y.-C., Hand, B., & McDowell, L. (2013). The effects of writing-to-learn activities on elementary students' conceptual understanding: Learning about force and motion through writing to older peers. *Science Education*, *97*(5), 745–771. doi:10.1002/sce.21067

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, Second Edition* (Revised.). Lawrence Erlbaum Assoc Inc.

Deci, E. L., & Ryan, R. M. (2006). Intrinsic motivation inventory. Retrieved Mai 4, 2012, from http:// www.selfdeterminationtheory.org

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287–312. doi:10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A

Ford, C. L., & Yore, L. D. (2012). Toward convergence of critical thinking, metacognition, and reflection: Illustrations from natural and social sciences, teacher education, and classroom practice. In A. Zohar, & Y. J. Dori (Hrsg.), *Metacognition in science education* (p. 251–271).Springer Netherlands.

Franzke, M., Kintsch, E., Caccamise, D., Johnson, N., & Dooley, S. (2005). Summary street: Computer support for comprehension and writing. *Journal of Educational Computing Research*, *33*(1), 53–80. doi:10.2190/DH8F-QJWM-J457-FQVB

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, *104*(2), 452–468. doi:10.1037/a0026683

Gunel, M., Hand, B., & Prain, V. (2007). Writing for learning in science: A secondary analysis of six studies. *International Journal of Science and Mathematics Education*, *5*(4), 615–637. doi:10.1007/s10763-007-9082-y

Kirby, J. R., & Lawson, M. J. (2012). *Enhancing the quality of learning: Dispositions, instruction, and learning processes*. Cambridge University Press.

Klein, P. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, *11*(3), 203–270. doi:10.1023/A:1021913217147

Martin, A. M., & Hand, B. (2009).Factors affecting the implementation of argument in the elementary science classroom. A longitudinal case study. *Research in Science Education*, *39*(1), 17–38. doi:10.1007/s11165-007-9072-7

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice*, *41*(4), 226–232. doi:10.1207/s15430421tip4104_4

McCrindle, A. R., & Christensen, C. A. (1995).The impact of learning journals on metacognitive and cognitive processes and learning performance. *Learning and Instruction*, *5*(2), 167–185. doi:10.1016/0959-4752(95)00010-Z

Novak, J. D. (2010). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Taylor & Francis.

Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise-reversal effects in writing-to-learn. *Instructional Science, 38,* 237-258.

Nückles, M., Hübner, S., & Renkl, A. (2009).Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction*, *19*(3), 259–271. doi:10.1016/j.learninstruc.2008.05.002

Nückles, M., Hübner, S., & Renkl, A. (2012). Fostering self-regulated learning by journal writing: How should instructional support be designed to promote high-quality learning? In J. R. Kirby, M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 178-200). New York, NY US: Cambridge University Press. doi:10.1017/CBO9781139048224.012

Nussbaum, M. E. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology, 33, 345-359.*

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *Journal of Experimental Education*, *76*(1), 59–92.

Reigeluth, C., & Stein, R. (1983). The elaboration theory of instruction. In Reigeluth (Ed.), *Instructional Design Theories and Models*. Hillsdale, NJ: Lawrence Erlbaum.

Renkl, A. (2013). Why practice recommendations are important in use-inspired basic research and why too much caution is dysfunctional. *Educational Psychology Review*, *25*(3), 317–324. doi:10.1007/s10648-013-9236-0

Reynolds, J. A., Thaiss, C., Katkin, W., & Thompson, R. J. (2012). Writing-to-learn in undergraduate science education: A community-based, conceptually driven approach. *CBE-Life Sciences Education*, *11*(1), 17–25. doi:10.1187/cbe.11-08-0064

Rigby, C. S., Deci, E. L., Patrick, B. C., & Ryan, R. M. (1992). Beyond the intrinsic-extrinsic dichotomy: Self-determination in motivation and learning. *Motivation and Emotion*, *16*(3), 165–185. doi:10.1007/BF00991650

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education*, *23*, 1–24.

Schmidt, K., Maier, J., & Nückles, M. (2012). Writing about the personal utility of learning contents in a learning journal improves learning motivation and comprehension. *Education Research International*, *2012*. doi:10.1155/2012/319463

Schunk, D. H., & Zimmerman, B. J. (2008). *Motivation and self-regulated learning: Theory, research, and applications*. New York: Lawrence Erlbaum Associates.

Tunnicliffe, S. D., & Ueckert, C. (2007). Teaching biology - the great dilemma. *Journal of Biological Education*, *41*(2), 51–52. doi:10.1080/00219266.2007.9656061

Webb, P. (2010). Science education and literacy: Imperatives for the developed and developing world. *Science, 328*(448), doi: 10.1126/science.1182596

Wigfield, A., Eccles, J. S., Roeser, R., & Schiefele, U., (2008). Development of achievement motivation. In W. Damon & R.M. Lerner (Eds.), *Child and adolescent development: An advanced course* (pp. 933-1002). New Jersey, NJ: Wiley & Sons.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal*, *46*(4), 1060–1106. doi:10.3102/0002831209333183

Wiley, J., & Voss, J. F. (1999).Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, *91*(2), 301–311. doi:10.1037/0022-0663.91.2.301

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, *38*(4), 189–205. doi:10.1207/S15326985 EP3804_1

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. doi:10.1207/s15430421tip4102_2

Zohar, A., & Nemet, F. (2002).Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*(1), 35–62. doi:10.1002/tea.10008