Effects of Instruction in Writing-to-Learn on Low-Achieving Adolescents in Biology and Mathematics Classes

Van Dijk, A.; Van Gelderen, A.; Kuiken, F.

DOI
10.3390/educsci12090595

Publication date
2022

Document Version
Final published version

Published in
Education Sciences

Citation for published version (APA):
Van Dijk, A., Van Gelderen, A., & Kuiken, F. (2022). Effects of Instruction in Writing-to-Learn on Low-Achieving Adolescents in Biology and Mathematics Classes. *Education Sciences*, *12*(9), Article 595. https://doi.org/10.3390/educsci12090595

education sciences

Article

Effects of Instruction in Writing-to-Learn on Low-Achieving Adolescents in Biology and Mathematics Classes

Aartje Van Dijk [1,*], Amos Van Gelderen [1] and Folkert Kuiken [2]

1 Research Centre Urban Talent, Rotterdam University for Applied Sciences, 3015 CX Rotterdam, The Netherlands
2 Amsterdam Center for Language and Communication, University of Amsterdam, 1012 ZA Amsterdam, The Netherlands
* Correspondence: a.c.van.dijk@hr.nl

Abstract: This study investigates the effects of instruction in genre writing with planning and revising activities (GWPR) on learning. This type of instruction appeared to be successful in promoting learning in several types of education. However, there are few studies on the effects on low achievers. Therefore, two studies were conducted with low-achieving students, each comprising a quasi-experimental study and a small-scale think-aloud study, both of which were embedded in regular education for low-achieving adolescents. The first study took place in biology classes (grade 7, three lessons); the second study was in mathematics classes (grade 10, six lessons). The researchers co-created writing-to-learn tasks with the teachers. The results showed positive effects on learning in mathematics classes as compared with the control group, but not in biology classes. The think-aloud study in the experimental mathematics class condition provided evidence of the learning by writing process. In the experimental biology class condition, such evidence was barely present. The results suggest that the experimental intervention in biology classes was too short for the students to grasp the essentials of learning by writing. This paper also discusses suggestions for further research and pedagogical implications.

Keywords: instructing writing-to-learn; learning by writing process; teaching cognitive strategies; low achievers

check for updates

Citation: Van Dijk, A.; Van Gelderen, A.; Kuiken, F. Effects of Instruction in Writing-to-Learn on Low-Achieving Adolescents in Biology and Mathematics Classes. *Educ. Sci.* **2022**, *12*, 595. https://doi.org/10.3390/educsci12090595

Academic Editor: Lisa Bendixen

Received: 26 February 2022
Accepted: 23 July 2022
Published: 31 August 2022

1. Introduction

The present study investigates the effects of instruction in writing-to-learn in two disciplines. Writing-to-learn (or: learning by writing) means that students carry out writing tasks intended to stimulate reflection on their knowledge about a disciplinary topic, which may lead to new insight and topic knowledge [1–3]. Topic knowledge refers to basic factual knowledge, such as names and dates. By insight, we mean the ability to relate new concepts to students' prior knowledge.

Over the years, writing-to-learn in educational contexts has been studied frequently. Researchers observed that teaching often consisted of mere transmission of knowledge and rarely stimulated students to arrive at new insights [4–6]. Therefore, writing-to-learn was suggested as a promising addition to traditional teaching.

At first, researchers held varying views on how to stimulate the process of writing-to-learn and research showed mixed results [7]. For instance, the "Writing Across the Curriculum" (WAC) movement suggested that using elements of good writing, such as attention to text organization and conciseness, leads to students' insight in subject matter [8]. Bazerman [9], however, observed that each discipline entails a specific way of reasoning, which means a disciplinary approach of writing-to-learn is required.

Recent research has shown that writing can lead to positive effects on insight and topic knowledge [3,10–13]. Studies are directed at students in all levels of education (from primary to higher education). However, few studies have been specifically aimed at writing-to-learn interventions for low-achieving adolescents. Although many studies may have included low achievers, results have rarely been reported specifically for that group. Therefore, it is not yet known whether low achievers can be supported by writing in the same way as other students. As Rivard [14] states, it may be that low achievers need a different type of writing-to-learn instruction compared to high achievers. Therefore, we set out to bolster the scant research specifically directed at low-achieving students in secondary education.

We investigate whether instruction in writing-to-learn can be beneficial for low-achieving students. On the one hand, there may be reasons why it is not. For instance, low-achieving students often have difficulty with the "basics" of the writing process (such as formulating correct sentences and spelling), and therefore have too little cognitive room to focus on text contents and reflect on disciplinary topics. In addition, low achievers may be less motivated to write in comparison to students with higher achievement levels, which may profoundly inhibit the conditions for writing-to-learn.

On the other hand, there is reason to believe that instruction in writing-to-learn is effective for low-achieving students as well because writing facilitates reflection on one's own thoughts by externalizing these thoughts in written text [15]. Low-achieving adolescents may profit from this externalizing of their own ideas in ways similar to those of other students. We investigate whether instruction in specific writing-to-learn tasks in two disciplines (biology and mathematics) improves low-achieving adolescents' topic knowledge and insight and whether we can identify indicators of the process of learning by writing in their task approach.

1.1. Research into Instruction in Writing-to-Learn

In recent research, explicit instruction is regarded as conditional for stimulating writing-to-learn. Bangert-Drowns et al. [16] were the first to highlight this point-of-view. They conducted a meta-analysis based on 48 studies in primary, secondary and higher education, concluding cautiously that instruction in cognitive and metacognitive strategies may lead to learning by writing. Cognitive writing strategies are understood as organizing strategies, such as goal setting, selecting, and structuring contents. Metacognitive writing strategies are understood as strategies for monitoring task performance and evaluating texts (reviewing and revising). Bangert-Drowns et al. [16] assumed that such instruction stimulates reflection on writing products, which may lead to new insight and topic knowledge.

More recent reviews of research into writing-to-learn confirmed that instruction in cognitive and metacognitive strategies may stimulate learning by writing [3,11–13]. In their meta-analysis of 56 (quasi) experiments, Graham et al. [11] found positive effects of explicit instruction in cognitive and metacognitive strategies for writing to learn in 82% of (quasi) experiments in science, social studies, and mathematics in grades 1–12. However, the researchers could not determine which elements of instruction were most effective in promoting learning.

A systematic review of 43 studies by Miller et al. [12] took place in regular courses in grades 6–12 in science, social studies, and mathematics. The researchers concluded that three types of instruction resulted in positive effects on learning. The first type was explicit instruction of cognitive strategies. The second was inquiry-based instruction stimulating students to discover how to use cognitive and metacognitive strategies without guidance by the teacher. The third was instruction of metacognitive strategies for students' reflection on their learning. Overall, in 46.5% of reviewed studies, instruction consisting of one of these types had positive effects on learning.

Hand et al. [13] reviewed 81 theses (41 quantitative and 40 qualitative) on the use of the Science Writing Heuristic (SWH) in secondary education. The SWH is a tool for instructing cognitive strategies for planning (organizing and generating) a science text. The researchers found that its use in the quantitative studies resulted in the growth of insight regardless of grade or cultural background. Additionally, a qualitative study on elements maximizing effects on learning showed that a determining factor is the amount of time available. This appeared to be true for students when engaged in using the SWH, as well as for teachers gaining experience with it.

Van Dijk et al. [3] reviewed 43 studies in primary (from grade 5), secondary, and higher education in all disciplines. The authors distinguished four types of instruction, three of which were based upon hypotheses about the process of writing-to-learn proposed by Klein [7], whereas the fourth type emerged from a considerable amount of reviewed studies. The first type is Forward Search, which consists of metacognitive strategies for revising the contents of a draft. The second is Genre Writing, which provides students with a model of the assigned genre, sometimes supplemented by cognitive and/or metacognitive strategies for planning and revision. The third is Backward Search, which means including instruction of cognitive strategies for planning and metacognitive strategies for revising a draft. The fourth is Planning Only, which is instructing cognitive strategies for planning. Results showed positive effects on insight and topic knowledge for all four types. In total, positive effects were found in 66% of the experimental studies.

Apart from the cognitive and metacognitive strategies involved in writing-to-learn that emerged from the above research syntheses, there are also some important conditions mentioned in the literature. Newell [17] considered genre knowledge conditional for writing-to-learn. If writers are not familiar with conceptual relations characteristic of an academic genre (for instance, causality in expositions), writing a text in that genre becomes extremely difficult. In addition, Klein and Kirkpatrick [18], and Klein and Samuels [19] concluded that genre knowledge is conditional for writing-to-learn. In the review by Van Dijk et al. [3], Genre Writing appeared as one of the main types of instruction. Genre Writing refers to teachers' explanation of characteristic ways conceptual relations are accomplished in the given genre. Another important condition for writing-to-learn concerns the audience writers are targeting. Prain [20] suggested that instructing students to write for a lay audience requires them to reflect more deeply on how to reformulate the academic language from their textbooks into everyday language. This reflection and reformulation may lead to new insight and topic knowledge.

The reviews by Graham et al. [11] and Miller et al. [12] reported small numbers of studies investigating writing-to-learn by students of various abilities. Graham et al. [11] noted that in two of 56 experiments, the participants were low-achieving adolescents, with both showing positive effects. Miller et al. [12] mentioned one study leading to positive effects. Hand et al. [13] mentioned two studies not showing significant differences between low achievers and high achievers on a post-test when compared to each other. In the review by Van Dijk et al. [3], no studies were encountered focusing exclusively on low-achieving students.

1.2. Low- and High-Achieving Students in Research on Writing-to-Learn

Indications of differences in the usefulness of writing-to-learn activities for higher- and lower-achieving students can be found in only a few studies. Rivard [14] investigated which specific language activities were beneficial for learning for 154 low- and high-achieving adolescents (eighth grade) in science class. The study compared three experimental groups to a control group, each group comprising high- and low-achieving adolescents. One group was instructed to carry out a task by means of peer discussion, another by means of writing preceded by peer discussion, and the third by writing only. The three groups were compared to a control group that performed restricted writing tasks (such as matching exercises, true-or-false questions, or fill-in-the-blanks). Low-achieving students appeared to acquire more topic knowledge after performing writing activities preceded by peer discussion about content selection and organization of their text. High achievers benefited the most from writing activities only.

Akkus, Gunel, and Hand [21] compared 322 experimental with 270 control students in science education, each condition comprising of (seventh to eleventh grade) high- and low-achieving students. Experimental students were instructed to perform inquiry activities by means of a checklist and to discuss these with peers before writing, just as

Rivard [14] did. Control students listened to the teacher's explanation of subject matter before writing. The researchers found that experimental low-achieving adolescents showed larger effects on insight and topic knowledge than control students. The explanation by Akkus et al. [21] is that peer discussion taking place in everyday language rather than academic language, which is usual in class, may support low achievers in understanding subject matter. High-achieving students gained equally high scores in both conditions; thus, the type of instruction did not matter.

Faber, Morris, and Lieberman [22] noticed that low-achieving students appeared to struggle when instructed to write notes for understanding their textbook, whereas high-achieving students did not. Therefore, they investigated whether a preparatory training on note taking for reading textbook sections led to positive effects on low-achieving students' learning. The study compared 112 experimental ninth grade students receiving note-taking training before studying their textbooks, with 90 control students receiving the business-as-usual program of the subject World Cultures. Both groups of students consisted of high- as well as low-achieving students. The activities preparing experimental students on note taking comprised the training of reading and note-taking skills. Low- as well as high-achieving experimental students showed more insight and topic knowledge into subject matter than control students.

The three studies above support the idea that low-achieving students may benefit from instruction using writing to learn tasks. Peer discussion for planning (selecting and organizing) or a preparatory activity on note taking appeared to lead to positive effects on low-achieving students' acquisition of topic knowledge Rivard [14], or on topic knowledge as well as insight in different disciplines. Therefore, it makes sense to further investigate the effects of instruction in writing-to-learn for low-achieving adolescents.

1.3. Supporting Low-Achieving Students for Learning

In our study, we combined the elements from the literature that appeared promising for stimulating writing-to-learn for low-achieving adolescents. We called the result Genre Writing instruction combined with Planning and Revising (GWPR). The elements are the following: genre knowledge, cognitive and metacognitive strategies, peer discussion, and writing for a lay audience. GWPR instruction provides low-achieving students with support at three phases of the writing process: before planning, during planning, and during reviewing.

First, GWPR provides genre knowledge in a preparatory activity. The genre is explained by means of a model text written in the given genre. Genre knowledge is defined as knowledge of the genre's rhetorical goal and prevalent conceptual relations between text elements needed to arrive at this goal [23]. For instance, the rhetorical goal of the genre explanation is clarifying, and a conceptual relation to arrive at this goal is 'condition' (e.g., in mathematics, 'If you know how many data you have, you can determine the size of your table').

Students need genre knowledge for understanding their textbooks and for relating concepts in their own texts. Therefore, the model text should illustrate how conceptual relations can be linguistically realized, for instance by means of 'because' or 'therefore' for relating a statement to an argument, or by means of 'first'...', then ... ', 'finally...' for relating steps in an instruction. By reflecting on composing a conceptual relation, for instance an argumentative relation, students may acquire new insights into the meaning of the conceptual relation [17–19]. The model text supports students in writing a text in a given genre. Additionally, the model text is directed at a lay audience, demonstrating how to formulate the conceptual relations in a non-academic fashion for readers with little prior knowledge of the subject.

Second, GWPR contains the cognitive strategies selecting and organizing needed for planning how to write the text. For selecting contents, students use their memory and their textbooks. For organizing contents, they decide how to relate selected contents by means of the previously demonstrated conceptual relations. Students discuss their planning in pairs, implying that they evaluate the appropriateness of each other's selection of contents for the intended lay audience as well as the comprehensibility of the conceptual relations. Peer discussion on planning supports them in reflecting on how to relate selected concepts and to organize these in a text. Students' reflections may result in new insights [1]. Additionally, students' discussion about the accessibility of their texts to their audience may lead to new insights [20]. The model text supports students in directing their texts at a lay audience.

Third, after students have written their drafts individually, they review their peers' drafts by reflecting on the description of conceptual relations and comprehensibility to the lay audience. Next, peers reflect on each other's feedback and ask for clarification, if needed. Then, students revise their drafts individually using the received feedback requiring them to reflect on their original ideas, which may lead to new insights [1].

Finally, writing to a lay audience is inherent in GWPR in the planning and revising phases. It entails that writers cannot use the academic language from their textbooks for an audience that has little prior knowledge about the topic. Instead, they must use everyday language to make their text comprehensible.

1.4. The Process of Writing-to-Learn

Only a few studies examined the process of writing-to-learn by using think aloud studies or keystroke logging for analyzing students' thinking. One study took place in elementary education [1], and four were conducted in higher education [2,24–26]). Low achievers were not involved in these studies. In our study, we decided to add to the existing evidence using empirical data for analyzing low-achieving adolescents' writing-to-learn processes.

Two theories about the nature of the writing-to-learn process were proposed, one by Bereiter and Scardamalia [1] and one by Galbraith [2]. Bereiter and Scardamalia [1] assume that the writing process of experienced writers is a recursive process between contents and rhetorical goals for planning, reviewing, and revising. Writers seek to reconcile contents and rhetorical goals, and therefore adjust their text several times on rhetorical and content aspects, resulting in knowledge transformation. By performing such activities, writers acquire new insights. According to this theory, learning by writing takes place in a cyclic process entailing reflection on content and rhetorical goals interactively.

Galbraith [2] proposed the dual process theory that distinguishes between a knowledge retrieval process and a knowledge constituting process]. Writers use their knowledge retrieval system for retrieving content while considering rhetorical goals. While writing, they use their knowledge constituting system to make connections between concepts, some of which may be connections the writer was not aware of previously (implicit knowledge). These new connections lead to new insights. The constitution of connections is a cyclic process in which writers alternately revise their text and refer to their knowledge constituting system. The two theories have in common that learning by writing takes place in a cyclic process entailing reflection on and adaptation of content and rhetorical goals.

Bereiter and Scardamalia [1] tested their theory in an empirical study with sixth grade students. They considered reflection as indicating knowledge transformation. In a think aloud study comparing six experimental with six control students, they found evidence that experimental students progressed more in transforming their knowledge during the planning of their texts in comparison to the control group.

Recently, Baaijen and Galbraith [24] used keystroke logging for measuring 78 university students' writing processes as described in the dual process theory. Students' revision of text, which entailed reflection, appeared to be related to increased insight.

In a think aloud study with 56 university students aimed at identifying writing-to-learn processes by conducting an exploratory think-aloud study, Klein [25] confirmed the outcomes by Bereiter and Scardamalia [1] that reflection on planning activities, such as goal setting, organizing, and generating, appeared to promote insight. Additionally, Klein [25] reported that reflection on reviewing and revising also resulted in insight. Van Dijk et al. [26] found that experimental biology and mathematics teacher students receiving

GWPR instruction showed more reflection on planning as well as on reviewing and revising than a control group. Thus, Klein [25] and Van Dijk et al. [26] appear to find support for both theories. Klein [25] concluded that positive effects of writing on learning might result from both knowledge transformation and knowledge constitution.

We investigate the writing-to-learn process of low-achieving students in secondary education to find out whether reflection on planning, reviewing, and revising are involved, just as they are for other student populations. In doing so, we intend to show how GWPR instruction contributes to low achievers' learning.

1.5. The Present Studies in Biology and Mathematics Education

The present studies investigate whether GWPR instruction leads to increased insight and topic knowledge of low-achieving adolescents. Additionally, the studies aim to identify indicators of the writing-to-learn process. The studies were carried out in the context of Dutch pre-vocational education. This educational track is aimed at students with the 30% lowest scores on an academic aptitude test (language and mathematics) administered before the start of secondary education (in grade 6). Therefore, these students are characterized as low achieving.

Pre-vocational education appeared to be a relevant context because writing-to-learn can be an appropriate way of learning for students in this type of education, providing them with opportunities for more active understanding and participation in academic discourse by writing. In addition, the teachers involved were fourth (last) year apprentice teachers, who had participated in a previous writing-to-learn experiment as students. Their class had served as experimental condition in a study on effects of GWPR instruction in teacher education [26]. Therefore, we could take advantage of the apprentices' experience and knowledge about writing-to-learn in designing the interventions and adapting them to the target group of students.

The studies were situated in two subject areas: biology and mathematics. This allows us to compare results from two widely different subject areas, in regard to the role that writing plays in regular educational practice. In biology, writing is a regularly used activity (for instance, for executing tasks from the textbook), whereas in mathematics, writing rarely takes place [27].

We report two studies, each comprising a quasi-experiment, one in biology (Study 1) and one in mathematics class (Study 2). GWPR was applied in both studies to be able to compare effects between both disciplines. Following recommendations by Hand et al. [13], the interventions were embedded in regular lesson series that were part of the curricula for biology and mathematics. The studies were carried out in co-creation with the teachers for organizing, composing, and embedding the intervention in the regular courses.

Additionally, think-aloud studies were included in Study 1 and Study 2. The combination of a quasi-experiment and a think-aloud study offers the opportunity to investigate whether experimental students not only acquire more knowledge and insight but also shows specific indicators of learning in their writing process. The process of learning while writing was analyzed by comparing samples from the control and the experimental group.

We formulated two research questions:

1. Does GWPR instruction lead to more insight and topic knowledge of low-achieving adolescents in the context of biology and mathematics when compared to business-as-usual lessons?
2. Does GWPR instruction in the above contexts lead to observable differences in processes of writing to learn between experimental and control students?

2. Study 1: Writing-to-Learn for Low Achieving Students in Biology Class

2.1. Materials and Methods

2.1.1. Participants

The study took place in May and June 2014. Four seventh grade classes of two pre-vocational schools in two cities participated. Classes within a school were randomly assigned to an experimental and a control condition. From the start, 102 students were present in these classes (50 control and 52 experimental). A number of 96 students (46 control and 50 experimental students) completed a (required) prior knowledge and vocabulary test at the start of the experiment. Six students were absent. Two other students were absent during the administration of the post-test, resulting in a sample of 94 students. Due to technical problems in one school during the data collection, students' answers on the post-test were incomplete. Therefore, we were not able to analyze the data of all 94 participants. This resulted in a final sample of 75 participants (32 control and 43 experimental). Table 1 presents the age, gender, and mother tongue of the final sample.

Table 1. Characteristics of the participants.

	Experimental Group (N = 43)	Control Group (N = 32)
Age	M: 12.8 (SD: 0.74)	M: 12.8 (SD: 0.61)
Gender	Female: 21	Female: 15
Mother tongue	Dutch: 29	Dutch: 22

For answering the second research question about the process of writing-to-learn, 10 students were randomly selected from the sample of 75 students, five students belonging to the control group and five to the experimental condition. Table 2 shows characteristics of these students.

Table 2. Characteristics of the participants of the think-aloud study.

	Experimental Group (N = 5)	Control Group (N = 5)
Age	M: 12.4 (SD: 0.55)	M: 12.6 (SD: 0.55)
Gender	Female: 4	Female: 3
Mother tongue	Dutch: 3	Dutch: 2

Two biology teacher students completing their final internship in two schools were involved as teachers in the study. In one school, the teacher student instructed the experimental condition while an experienced colleague instructed the control group. In the other school, the teacher student taught both the experimental and control groups.

The teacher students were fourth (final) year students at the teacher education institute of a University of Applied Sciences. In the third year of their study, they had participated in an experiment [26] directed at the effects of GWPR instruction in biology teacher education. Therefore, they were familiar with our approach of instruction in writing-to-learn and motivated to adapt this approach to the teaching context of pre-vocational education.

2.1.2. Design

We used a quasi-experimental, post-test only design, with a business-as-usual control group and an experimental group. The dependent variables were insight and topic knowledge into biology subject matter taught. Prior knowledge of biology and vocabulary knowledge were used as covariates. The control group received regular lessons (without writing tasks) while the experimental group received GWPR lessons comprising writing tasks aimed at writing-to-learn.

For analyzing the process of writing-to-learn, a think-aloud multiple case study was carried out with five experimental and five control students, randomly selected from the final sample. A post-test writing task was used for the think-aloud procedure (see Instruments). Utterances were coded and systematically analyzed to investigate differences between the writing processes of experimental and control students.

2.1.3. Treatment

The experimental intervention took place in three 45-min lessons. The lessons were part of a series of eight lessons aimed at insight and topic knowledge into the musculoskeletal system. Preceding the experiments, the first author observed regular lessons in both schools to determine their structure in preparation for composing the intervention (more specifically, how to insert the writing-to-learn tasks). The lesson structure appeared to be identical in both schools.

The first author cooperated with the two teacher students in designing three writing-to-learn tasks, accompanying model texts and a teacher's manual. First, they determined which lesson series from the textbook (both teachers used the same) was most suitable for embedding the writing tasks. Then, they discussed which parts of the lessons to replace with writing tasks. Finally, the first author proposed topics for the three writing tasks and the model texts, which were evaluated by the teachers on how these were related to the main aims and contents of the lesson series.

It was decided to focus on the conceptual relations 'purpose and means', 'comparing', and cause and effect' because these relations appeared to be frequently used in the regular lessons. These conceptual relations match the genre 'explanation' [28]. Therefore, the writing tasks required students to write explanatory texts, each focusing on a different conceptual relation.

GWPR instruction entails that each writing task is preceded by a teacher's explanation of a model text. Three model texts were derived from a textbook directed at the target audience of the writing tasks (grade 6) and rewritten to fit our needs. The topics of the model texts were related to that of the writing task, such that students could use the model as an example for how to realize the conceptual relations. However, the topics of the model texts were not identical with the writing tasks to avoid copying. For instance, the topic of a model text was 'using muscles for kicking a ball (soccer)' and the topic of the writing task was 'using muscles for throwing a ball (basketball)'. Each model text contained one conceptual relation (as in the writing task) that was expressed in various ways (e.g., the relation 'purpose and means' was expressed as: 'use ... to ... '; 'for that ... you use... '). The purpose was to familiarize students with various linguistic expressions of a conceptual relation.

In each writing task, the experimental participants were instructed to use the conceptual relation present in the model text when writing their explanation. Furthermore, students were instructed to direct their text at a younger audience of sixth grade students. By carrying out such tasks, students needed to reflect on how to transform their own thinking to adjust to their audience (the rhetorical plane of writing). We assume that this may lead to students' awareness of the need for a clear (and simpler) explanation of relations. GWPR assumes that this awareness results into more explicit topic knowledge and insight into subject matter [20].

The GWPR writing tasks consisted of the following parts. The first concerned pairwise planning activities by brainstorming on possible contents. Students had to take into account the assigned conceptual relation and the intended audience. The second part involved the writing of a draft individually. The third required students to give feedback on each other's draft referring to the conceptual relation and the intended audience. Finally, students had to revise their drafts using the received feedback.

The teachers received a manual for how to instruct students to do their homework and how to present the model texts and the writing tasks to the students. It contained suggestions for explaining the conceptual relation to the students, with the model text projected on a smart board, highlighting various linguistic expressions of the conceptual relation. In addition, students were asked to look for other examples of the conceptual relation in the model text and discuss these.

2.1.4. Instruments

2.1.4.1. Prior Knowledge Tests

In consultation with the two teacher students, the first author composed the tests for determining students' prior topic knowledge and insight. The tests were based upon subject matter taught in the first part of the seventh-grade academic year as well as in grade 6.

The grade 6 curriculum contains several different themes related to biology, varying from nature to sports. Eight items were derived from the biology textbook for grade 7 [29] and five were derived from a national test for investigating sixth grade students' knowledge level in biology [30]. Four items around the theme 'sports' were composed in collaboration by the first author and the teacher students.

The prior topic knowledge test consisted of 10 multiple choice items; the prior insight test consisted of seven open-ended questions. Two items were deleted from the prior topic knowledge test because one showed no variability in students' answers, and the second item was invalid, thus eight items were left. We did not compute interrater reliability for the open-ended questions because these were unambiguous. (For instance, 'What kind of clothes are best to practice sports in?').

The homogeneity of the tests measuring prior topic knowledge and insight is low (Cronbach's alpha is 0.36 and 0.34, respectively), which was expected because of the large variation in biological themes tested. Cronbach's alpha provides an underestimation of test reliability [31–33]. Therefore, the tests might still be sufficiently reliable to explain variance in our posttest measures. Consequently, we decided to include both tests as covariates in our analysis.

2.1.4.2. Vocabulary

A vocabulary test of 30 items derived from the Dutch version of the Peabody Picture Vocabulary Test [34] was composed. This Dutch version is based on frequencies per one million words. The words are ranked in 17 sets, each aimed at a specific age group. For the selection of words, we used three sets (nr. 6, 7, 8) aimed at ages below 15. We selected words of which the expected proficiency was $p = 0.50$ and higher. Nine items did not show variability and were therefore deleted. Cronbach's alpha over 21 items was 0.68, which is regarded as acceptable.

2.1.4.3. Topic Knowledge and Insight in the Post-Test

The post-test consisted of 14 multiple choice items (topic knowledge) and 10 open-ended questions (insight). The items and questions were derived from a test that was part of the biology textbook used in grade 7. Cronbach's alpha of the topic knowledge test was 0.20. Due to this low homogeneity, this test was not included in the analysis. In contrast with the tests for prior knowledge, the topic knowledge post-test consisted of items from the narrow domain of lesson contents about the musculoskeletal system. Therefore, a much higher homogeneity of the topic knowledge test was expected here.

Cronbach's alpha for the post-test for insight was 0.70, which is acceptable given the narrow domain from which the items were selected. Therefore, the test was included in our analysis. The open-ended questions were formulated in such way that only one answer was possible (for instance: which type of joint is in a knee?). Therefore, there was no need to calculate inter-rater reliability.

2.1.4.4. Post-Test Writing Task

Finally, a post-test writing task was administered. Students were asked to write an explanation about the correct body posture for lifting a heavy weight. The explanation had to be directed at grade 6 students entering high school the following academic year. We used this writing task for examining the process of writing-to-learn in the think-aloud study.

2.1.5. Procedure

Biology class was scheduled twice a week. The study involved six lessons of 45 min each. The first lesson was used for administration of the tests for prior insight, topic knowledge, and vocabulary; the fifth lesson was used for the post-test (topic knowledge and insight); and the sixth lesson for the post-test writing task. Lessons 2–4 were dedicated to the three writing tasks for the experimental condition and the business-as-usual lessons in the control condition. Table 3 presents the lesson structure for these three lessons in the control and experimental condition. The lesson structure in the control condition consisted of explanation of new theory and the completion of textbook assignments. For homework, students were required to finish assignments from the textbook that had not been completed. The three experimental lessons each comprised one writing-to-learn task consisting of an explanation of the model text, followed by planning and writing a first draft, and giving and receiving feedback for writing a final draft. In each lesson, experimental students received a writing task about theory they had studied as homework prior to the lesson.

Table 3. Lesson structure in control and experimental conditions.

Control Group	Experimental Group
Business as usual lessons 2, 3, 4	Writing-to-learn tasks lessons 2, 3, 4
• Explanation of new theory • Performing assignments from the textbook	• **Explanation of a model text** • **Planning and writing a first draft** • **Feedback and writing a final draft**
• Homework: completing unfinished assignments	• **Homework: studying new theory**

For securing treatment fidelity, the first author observed whether the teacher students in both schools carried out the instructions conforming the teacher manual. No deviations were encountered in that respect. However, in both schools, the observations also showed that students in the experimental condition at first worked seriously but became less concentrated when giving feedback to each other and revising their drafts. Therefore, the first author discussed this with the teacher students aiming to improve students' engagement. For instance, teacher students explained that students' feedback contained suggestions for replacing wordings that were incomprehensible for the intended audience. This served as a boost for the students' motivation to put effort in the review and revision of their texts.

For keeping time on task equal for both conditions, the elements 'explanation of new theory' and 'performing assignments from the textbook' were replaced in the experimental group by, respectively, explanation of the model text and planning, writing, feedback, and writing final draft (see Table 3). The homework 'completing unfinished assignments' was replaced by 'studying new theory', which was assigned in lesson 1. The teacher explained the experimental students how to study new theory individually by means of a roadmap on reading strategies because they were not used to performing such a task.

For investigating the process of writing-to-learn, a random sample of 10 students (five experimental and five control) performed the writing task thinking aloud. The sessions took place in lesson 6. These students performed their writing tasks individually in an empty classroom in the presence of the first author. She said that she was interested in how they addressed the writing task; for this reason, she asked them to think aloud while writing. She provided an instruction comprising a video clip of a student thinking aloud while writing to demonstrate what was expected from the students. When students kept silent for 10 s, the first author encouraged them to keep thinking aloud, and used prompts such as: please, say aloud what you are thinking. Students wrote their texts by hand. The duration of the sessions was 15–20 min. The sessions were video recorded.

2.1.6. Coding Students' Transcribed Utterances

Students' verbalizations were transcribed and represented in protocols as separate utterances in cases of verbal behavior and separate instances in cases of non-verbal behavior (for instance, frowning). An utterance was defined as a phrase containing one complete element of information [35].

The codes were based on Hayes and Flower's [36] writing model The following writing activities were distinguished: (1) planning, including four subcategories (orienting, generating, selecting and organizing); (2) formulating; (3) monitoring; and (4) evaluating. In total, 24 codes to describe students' writing and thinking processes were used. For instance, the utterance 'yes, this should be it' was coded as 'selecting content'. In addition, interruptions and utterances not focusing on the writing task were coded as 'other activities. Finally, the resulting coding scheme comprised 29 verbal and non-verbal activities divided over five main categories (see Appendix A).

By means of this coding scheme, a researcher and a research-assistant coded the utterances, one code per utterance or per instance. For determining inter-rater reliability, two protocols (one for each condition) were coded by two independent raters. There was agreement for 84% of all utterances/instances, which is acceptable for our purposes. Differences in coding were resolved after discussion.

Reflective activities such as reviewing, revising, goal setting, organizing, and generating contents can be regarded as indicators of writing-to-learn processes, according to previous studies [1,24,25]. The following specific codes from our list are regarded as indicators for these reflective activities: (1) using knowledge about audience; (2) thinking about content selection; (3) thinking about formulating; (4) revising while formulating; (5) revising after finishing an utterance; (6) rereading own text; and (7) rethinking task approach.

2.1.7. Data Analysis

Three analyses of variance (ANOVAs) were used for comparing prior topic knowledge, prior insight, and vocabulary in the two conditions. By means of analyses of covariance (ANCOVA), students' post-test scores on insight in the two conditions were compared. Prior insight, prior topic knowledge, and vocabulary were used as covariates. In all statistical tests, alpha level was 0.05.

Utterances indicative for the process of writing-to-learn and other utterances were systematically analyzed to investigate differences between the experimental and control students that were randomly selected from the sample. To determine whether indicators for writing-to-learn occurred more often in the experimental group than in the control group, the means of each code per condition were computed. Subsequently, the ratio of the mean frequency of each code to the total number of utterances in each condition was computed. Finally, effect-sizes (Cohen's d) were computed for estimating the magnitude of differences between conditions on these proportioned mean frequencies (0.20–0.50 small, 0.50–0.80 medium, $\geq$0.80 large).

2.2. Results

Table 4 presents the means and standard deviations for the four variables involved, i.e., prior insight, prior topic knowledge, vocabulary, and post-test insight.

Table 4. Means and standard deviations for prior insight, prior topic knowledge, vocabulary knowledge, and post-test insight.

N = 75 Variables	**Experimental Group (N = 43)** Mean (SD)	**Control Group (N = 32)** Mean (SD)
1) Prior insight	5.65 (1.71)	5.50 (1.92)
2) Prior topic knowledge	5.60 (1.31)	5.56 (1.64)
3) Vocabulary	19.65 (1.15)	19.41 (2.58)
4) Post-test insight	6.16 (2.46)	6.84 (2.43)

1) theoretical maximum score: 7; 2) theoretical maximum score: 8; 3) theoretical maximum score: 21; 4) theoretical maximum score: 10.

2.2.1. Prior Insight, Prior Topic Knowledge, and Vocabulary

ANOVA showed that experimental students' and control students' scores on prior insight did not differ significantly: F (1.73) = 0.13, p = 0.72, partial η^2 = 0.002. Scores of the two groups on prior topic knowledge did not differ significantly either: F (1.73) = 0.015, p = 0.90, partial η^2 = 0.00. Finally, the scores of the two groups on the vocabulary test did not differ significantly: F (1.73) = 0.31, p = 0.58, partial η^2 = 0.004.

2.2.2. Differences in Post-Test Insight

Two ANCOVAs were performed to compare experimental students' and control students' scores on post-test insight. In the first ANCOVA, prior topic knowledge, prior insight, and vocabulary served as covariates. Prior topic knowledge and prior insight predicted the scores on post-test insight significantly: F (1.73) = 5.09, p = 0.03, partial η^2 = 0.07, and F (1.73) = 5.85, p = 0.02, partial η^2 = 0.08. The remaining covariate (vocabulary) did not predict the scores significantly: F (1.73) = 1.25, p = 0.27, partial η^2 = 0.02.

Therefore, in the second ANCOVA, the latter covariate was omitted. The analysis showed again that prior topic knowledge predicted post-test insight significantly: F (1.73) = 5.34, p = 0.02, partial η^2 = 0.07; the same was true for prior insight: F (1.73) = 7.63, p = 0.007, partial η^2 = 0.10. However, experimental students did not differ significantly from control students on the post-test insight: F (1.73) = 2.51, p = 0.12, partial η^2 = 0.04.

2.2.3. The Process of Writing-to-Learn

Table 5 shows the proportioned means and standard deviations of the seven codes representing reflective activities that are considered indicators of the writing-to-learn process (see Appendix A for an overview of all 29 codes). The second, third, fourth, and fifth columns show the proportioned mean frequencies and standard deviations of codes for the experimental and control condition. The final column shows the effect size (Cohen's d).

Table 5. Proportioned means, standard deviations, and effect sizes of indicators of writing-to-learn in experimental (N = 5) and control groups (N = 5).

	Experimental		**Control**		**Effect Size**
Codes	Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
Planning: generating					
Using knowledge about audience	0.00	0.00	0.00	0.01	0.00
Planning: selecting					
Thinking about content selection	0.09	0.07	0.02	0.03	1.30
Formulating					
Thinking about formulating	0.01	0.02	0.01	0.02	0.00
Revising while formulating	0.00	0.01	0.02	0.03	−0.89
Revising after finishing an utterance	0.01	0.01	0.01	0.01	0.00
Monitoring					
Rereading own text	0.02	0.01	0.02	0.03	0.00
Rethinking task approach	0.01	0.01	0.01	0.02	0.00

Table 5 shows that only one reflective activity, thinking about content selection, was used more frequently by experimental students than control students, d = 1.30 (large effect). The remaining activities did not show differences in favor of the experimental condition. However, control students performed one reflective activity, revising while formulating, more often than experimental students (large effect). Students in both conditions appeared to be rarely inclined to reflect on their writing on most of the indicators in Table 5.

3. Study 2: Writing-to-Learn for Low Achieving Students in Mathematics Class

3.1. Materials and Methods

3.1.1. Participants

The study took place in four tenth grade mathematics classes in two schools for pre-vocational education in two cities in November and December 2014. Classes within a school were randomly assigned to an experimental and a control condition. From the start, 74 students (33 experimental and 41 control) participated in the study. Seven students were absent when the post-test was administered. Therefore, the final number of students participating in the study was 67 (30 experimental and 37 control students). Table 6 presents students' ages, genders, and mother tongues.

Table 6. Characteristics of the participants.

	Experimental Group (N = 30)	Control Group (N = 37)
Age	M: 15.8 (SD: 0.57)	M: 15.8 (SD: 0.64)
Gender	Female: 19	Female: 16
Mother tongue	Dutch: 27	Dutch: 35

For answering the second research question about the process of writing-to-learn, eight students were randomly selected from the sample of 67 students, four of whom belonged to the control group and four to the experimental condition. Table 7 shows the characteristics of these students.

Table 7. Characteristics of the participants of the think-aloud study.

	Experimental Group (N = 4)	Control Group (N = 4)
Age	M: 15.3 (SD: 0.50)	M: 15.5 (SD: 0.58)
Gender	Female: 2	Female: 1
Mother tongue	Dutch: 4	Dutch: 4

Two mathematics teacher students were involved in the study, one in each school. In both schools, the teacher students instructed students in the experimental condition, while an experienced colleague instructed students in the control group. The teacher students were fourth year students at a teacher education institute of a university of applied sciences. In the previous year, they had participated in an experiment examining the effects of GWPR instruction in mathematics teacher education [26] and had carried out writing-to-learn tasks themselves. Therefore, they were familiar with our approach of instruction in writing-to-learn and motivated to adapt this approach to the context of pre-vocational education.

3.1.2. Design

In a quasi-experimental, post-test only design, a control group and an experimental group were compared. The control group received business-as-usual lessons (without writing tasks) and the experimental group received GWPR lessons comprising writing tasks aimed at writing-to-learn. The dependent variables were insight and topic knowledge into the taught mathematics subject matter. Prior insight, prior topic knowledge in mathematics, and vocabulary knowledge were used as covariates.

For analyzing the process of writing-to-learn, a think-aloud multiple case study was carried out with four experimental and four control students who were randomly selected from the total sample. A post-test writing task was used for the think-aloud procedure (see Instruments). Utterances were coded and systematically analyzed to investigate differences between the writing processes of experimental and control students.

3.1.3. Treatment

The experiment took place in a lesson series aimed at insight and topic knowledge into mathematical relations. Preceding the experiments, the first author observed regular mathematics lessons in both schools to determine the usual proceedings and prepare the replacement of parts of the lessons with writing-to-learn tasks in the experimental condition. In both schools, teachers first explained new theory from the textbook about the topic at stake, after which students carried out assignments related to the explained theory. For homework, students completed assignments if they failed to complete them during the lessons.

In preparation for the intervention, the first author cooperated with the teacher students to design three writing tasks, accompanying model texts, and a teacher's manual. First, the first author and the teacher students determined collaboratively which lesson series from the textbook was most suitable for embedding writing tasks (both schools used the same textbook). They discussed which parts of the business-as-usual lessons to replace with writing tasks. Second, the first author proposed topics for the writing tasks and model texts. The teacher students decided whether these topics involved important aims of the lesson series. Finally, the teacher's manual composed by the first author was checked by the teacher students to ensure that suggestions for explaining conceptual relations in the model text and for lesson proceedings were clear.

For composing the writing tasks, it was decided to use the conceptual relations 'condition', 'sequence', and 'comparing' because these were most important in the business-as-usual lessons. These conceptual relations match the genre explanation [28]. Therefore, this genre was used in all three writing tasks. The writing tasks required students to write an explanation of how mathematical tasks can be carried out, each focusing on a different conceptual relation. Additionally, each writing task required students to write their explanation for an audience consisting of sixth grade students who have no prior knowledge about the mathematical tasks described. To enable their audience to understand the text, students needed to simplify their formulations thoroughly.

GWPR instruction entails that each writing task is preceded by a teacher's explanation of a model text. The model texts were based on texts from the students' mathematics textbook and were rewritten in such a way that they were comprehensible to an audience of sixth-grade students. The topics of the model texts were related to the topics of the writing tasks in such a way that students could use the text as an example for writing. To avoid copying, the topics of the model texts and aligned writing tasks were not identical. For instance, the topic of a model text was 'which shop is the cheapest when taking account of shipping costs and the price?' while the topic of the aligned writing task was 'where can you buy the most seeds for a specific amount of money?' Each model text contained one conceptual relation that was expressed in various ways. For instance, the relation 'condition' was expressed as: 'if you know ... you can ... '; 'if you make ... you not only give... but you ... '. The purpose was to familiarize students with various realizations of a conceptual relation. As is standard in texts about mathematics [27], the model texts contained graphical representations, such as a table, in addition to the text. The different stages of GWPR instruction on writing-to-learn (planning, individual formulation, review, and final draft) were the same as in Study 1.

3.1.4. Instruments

Prior Knowledge Tests

In consultation with the two teacher students, the first author composed the tests for determining students' prior topic knowledge and insight. The test was based upon test items from the textbooks students had used in grade 9 [37,38] and in grade 10 in the period preceding the experiment [39]. The items concerned mathematical skills, such as understanding and drawing graphs, and elements of the comprehensive theme 'mathematical relations', for instance 'equations'. Prior topic knowledge was measured using 10 multiple choice items; prior insight was measured using five open-ended questions (with fixed answers).

The homogeneity of the tests measuring prior topic knowledge and insight is low (Cronbach's alpha is −0.14 and 0.15, respectively), which was expected because of the large variation in mathematical subjects tested. Cronbach's alpha provides an underestimation of test reliability [31–33]. Therefore, the tests might still be sufficiently reliable to explain variance in our posttest measures. Therefore, we decided to include both tests as covariates in our analysis.

Vocabulary

A vocabulary test of 30 items derived from the Dutch version of the Peabody Picture Vocabulary Test [34] was composed. For the selection of words, we used four sets (nr. 9, 10, 11, 12) aimed at all ages above 15. Since the students belonged to the youngest in this category, we selected words with *p*-values on or above 0.80. Cronbach's alpha was 0.89.

Topic Knowledge and Insight in the Post-Test

Unfortunately, topic knowledge could not be measured because of test regulations in one school that excluded multiple choice questions from mathematics tests. Insight was measured by means of a post-test consisting of 17 open-ended questions, which were mathematic sums. These questions were selected in consultation with the teachers from two final tests completing the theme 'mathematical relations' in the used textbooks.

Scores on the post-test were based on one or more elements students' answer should comprise. For each element, one point was assigned. The number of elements varied per item. The maximum score for the 17 items was 33 points. Cronbach's alpha of the open-ended questions was 0.78. Given that the items of the post-test were derived from the narrow domain of mathematical relations that was taught in the lessons (in contrast with the prior knowledge tests), a rather high homogeneity was expected. Inter-rater reliability was not computed because the answers on the open-ended questions (sums) were fixed.

Post-Test Writing Task

Finally, a post-test writing task was administered. The task consisted of two parts: a computation and a writing task. First, students had to make the computation and next the writing task, which was writing an explanation of how they had carried out the computation in the first part. Students were asked to write this explanation for an audience of sixth-grade students. This writing task was used for examining the process of writing-to-learn in the think-aloud study.

3.1.5. Procedure

Mathematics lessons were scheduled twice a week, with a lesson duration of 45 min. In total, nine lessons were involved in this study. The first was used for administration of the tests for prior insight, topic knowledge, and vocabulary. The second, fourth, and sixth lesson were used for the first part of each writing task (writing the first draft). The third, fifth, and seventh lesson were used for the second part of the writing task (writing the final text). In the eighth lesson, the post-test for insight was administered, and in the ninth lesson, students performed the post-test writing task.

Table 8 presents the lesson structures for the control and experimental conditions. The lessons in the control condition consisted of the teacher's explanation of new theory, students' completing assignments from the mathematics textbook, and completing unfinished assignments for homework. To keep time on task equal in both conditions, the main elements of the business-as-usual lessons were replaced by a writing task in the experimental condition. Each of the three experimental writing tasks were performed in two lessons. Explanation of the model text, planning, and writing the first draft (the first part) took place in one lesson, and reviewing and revising the draft (the second part) in the subsequent lesson. Teachers' explanations of new theory were replaced by homework requiring students to study new theory from their textbooks. Since students were not used to studying theory individually, they received instruction on how to perform their homework in lesson 1.

For securing treatment fidelity, the first author observed whether the teacher students carried out the lessons as intended and as described in the teacher's manual in all experimental lessons. No deviations in delivery of the experimental condition were encountered. However, in one school, a change of scheduling because of unexpected staff meetings became necessary. Therefore, the two parts of the second and third writing tasks had to be merged into one lesson each. This resulted in the loss of two (of six) lessons from the series for the two classrooms involved.

Table 8. The lesson structure in control and experimental conditions in mathematics class.

Control Group	Experimental Group	
Business as usual (lessons 2–7)	First draft (lessons 2, 4, 6)	Revision (lessons 3, 5, 7)
• Explanation of new theory • Performing assignments from the textbook	• **Writing to learn task part 1: explanation of model text, planning and writing a first draft**	• **Writing to learn task part 2: feedback and writing final text**
• Homework: completing unfinished assignments	• **Homework: studying new theory**	

Eight students (four experimental and four control) were randomly selected and performed the writing task thinking aloud. The think-aloud sessions took place the day after the post-test had been administered. Students performed the writing task individually in an empty classroom, supervised by the first author. She provided them with the same type of instruction as in Study 1 (see Procedure). Students wrote their texts by hand. Their performance was video recorded. The duration of the sessions was 15–20 min.

3.1.6. Coding Students' Transcribed Utterances

The coding scheme consisting of 29 codes was the same as the one used in study 1 (see Appendix A). By means of this scheme, a researcher and a research assistant coded the utterances, one code per utterance or per (non-verbal) instance. For determining inter-rater reliability, two protocols (one for each condition) were coded independently by two raters. There was agreement for 78% of all utterances and instances. We consider this a sufficient reliability of coding. Differences in coding were resolved after discussion.

3.1.7. Data Analysis

ANOVA was used for comparing students' prior insight, prior topic knowledge, and vocabulary in the two conditions. By means of ANCOVA, differences between students' post-test scores on insight in the two conditions were compared. Prior insight, prior topic knowledge, and pre-test vocabulary were used as covariates. In all statistical tests, alpha level was 0.05. The analysis of the process of writing-to-learn was performed as described in Study 1.

3.2. Results

Table 9 presents the means and standard deviations for four variables involved: prior insight, prior topic knowledge, vocabulary, and post-test insight. Scores on the latter test were based on one or more elements students' answer should comprise. For each element, one point was assigned. The number of elements varied per item. The maximum score for the 17 items was 33 points.

Table 9. Means and standard deviations for prior knowledge (insight), prior knowledge, Vocabulary knowledge, and post-test scores on insight and topic knowledge.

N = 67 Variables	**Experimental Group (N = 30)** Mean (SD)	**Control Group (N = 37)** Mean (SD)
1) Prior insight	0.60 (0.62)	0.65 (0.72)
2) Prior topic knowledge	5.00 (1.46)	4.46 (1.33)
3) Vocabulary	21.67 (5.03)	22.05 (5.05)
4) Post-test insight	19.67 (6.23)	12.95 (7.51)

1) theoretical maximum score: 5; 2) theoretical maximum score: 10; 3) theoretical maximum score: 30; 4) theoretical maximum score: 33.

3.2.1. Differences in Prior Insight, Prior Knowledge, and Vocabulary

ANOVA showed that experimental students' and control students' scores on prior insight and prior knowledge did not differ significantly: respectively, F (1.65) = 0.086, p = 0.77, partial η^2 = 0.001, and F (1.65) = 2.51, p = 0.118, partial η^2 = 0.037. No significant differences in vocabulary knowledge between the two conditions were found either: F (1.65) = 0.098, p = 0.76, partial η^2 = 0.002.

3.2.2. Differences in Post-Test Insight

An ANCOVA showed that the covariates vocabulary and prior topic knowledge were not significantly related to post-test insight: respectively, F (1.65) = 1.77, p = 0.19, partial η^2 = 0.03, and F (1.65) = 1.13, p = 0.29, partial η^2 = 0.02. The covariate prior insight predicted the scores on post-test's open-ended questions significantly: F (1.65) = 9.48, p = 0.003, partial η^2 = 0.13. Therefore, in the final analysis, only the covariate prior insight was included. This showed that the control and experimental group differed significantly in post-test insight. The experimental group scored higher than the control group: F (1.65) = 18.84, p = 0.00, partial η^2 = 0.23. Additionally, the covariate prior insight was related significantly with the dependent variable: F (1.65) = 8.98, p = 0.004, partial η^2 = 0.12.

3.2.3. The Process of Writing-to-Learn

Table 10 shows the proportioned means and standard deviations for each of the seven selected indicators of writing-to-learn for experimental and control students and the effect size of the difference. The results of all codes in the think aloud study are given in Appendix B.

Table 10. Proportioned means and standard deviations and effect sizes of indicators of writing-to-learn in experimental group (N = 4) and control group (N = 4).

	Experimental		**Control**		**Effect Size**
Codes	Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
Planning: generating					
Using knowledge about audience	0.03	0.05	0.01	0.02	0.53
Planning: selecting					
Thinking about content selection	0.00	0.00	0.01	0.05	−0.29
Formulating					
Thinking about formulating	0.00	0.00	0.00	0.00	0.00
Revising while formulating	0.00	0.00	0.00	0.00	0.00
Revising after finishing an utterance	0.02	0.02	0.00	0.00	1.43
Monitoring					
Rereading own text	0.07	0.13	0.03	0.03	0.42
Rethinking task approach	0.16	0.19	0.10	0.09	0.40

Table 10 shows differences in favor of the experimental students on the following reflective activities: using knowledge about audience, d = 0.53 (medium effect); revising after finishing an utterance, d = 1.41 (large effect); rereading own text, d = 0.42 (small effect); and rethinking task approach, d = 0.40 (small effect). Two reflective activities, thinking about formulating and revising while formulating, were not used in both groups. Thinking about content selection took place in the control group only (small effect).

4. Discussion

We assume that writing can be a tool for learning that is suitable not only for high-achieving students but also for low achievers. When students write down their thoughts, they can read and reread these externalized thoughts and reflect on their intended meaning. In doing so, they may start questioning rhetorical issues, such as the comprehensibility of their ideas to the intended readers. In turn, this may lead to new insights and the acquisition of new topic knowledge.

GWPR instruction stimulates students' reflection repeatedly. For low achievers, a strength of GWPR is that reflection on planning and reviewing activities takes place in peer discussions that support them in learning [14]. Additionally, it provides students with the opportunity to discuss their thoughts using everyday language instead of academic language, which may stimulate low achievers' learning as well [14,21]. We expected that GWPR instruction would lead to new insights for low achievers, which was the case for mathematics, but not for biology. This outcome may be explained by the age difference between the biology (12–13) and mathematics students (15–16) in our studies. The younger students may have had more difficulty with reflecting on their texts. Although studies conducted with young students (grade 6) led to positive effects of writing-to-learn [40,41], these studies were not specifically directed at low achievers.

However, a more plausible explanation for the disappointing result in the biology classes is that time on task plays an important role in learning by writing. In mathematics class, the two parts of the writing tasks, 'planning and writing a draft' and 'reviewing and revising the draft', were spread over two 45 min lessons. In biology class, however, both parts of the writing tasks had to be carried out in one lesson of 45 min. This may have been too short for the students to really experience the importance of the different steps to be taken in GWPR instruction. Probably, the explanation of the conceptual relation at stake in each writing task and the peer discussions on planning and reviewing (feedback) require more time to be grasped than was available in one single lesson. The review phase in GWPR (including feedback from peers and text revision based on that feedback) is something that students are not especially used to do in regular education [42,43]. Therefore, they may have completed their writing tasks rushing without understanding the purpose of revising. In addition, an advantage of splitting each writing task in two lessons is that students have the opportunity to view their drafts with 'new eyes' because their memory of writing the first draft has somewhat faded. This provides a better motivation for them to take the revision of their texts more seriously.

Hand et al. [13] reviewed studies on the Science Writing Heuristic (SWH), a tool for learning by writing a lab report. They concluded that the time on task is an important factor for student learning. Their explanation is that students need sufficient time for becoming familiar with new ways of learning. This probably applies to GWPR instruction as well. It may be especially true for low achievers because they need more time for learning in general than average or high achievers do.

4.1. The Writing-to-Learn Process

The assumption behind the think-aloud studies was that GWPR instruction stimulates a recursive process between content and rhetorical goals [1,2] taking place during planning and revising activities. Seven reflective activities were labeled as indicators of this recursive process (see Table 11).

Table 11. Effect sizes of indicators for learning by writing in study 1 (biology) and study 2 (mathematics).

Indicators for Learning by Writing	Biology Cohen's d	Mathematics Cohen's d
Planning		
Using knowledge about audience	Not used *	Medium
Thinking about content selection	Large	Small negative effect
Formulating		
Thinking about formulating	No effect	Not used
Revising while formulating	Large negative effect	Not used
Revising after finishing an utterance	No effect	Large
Monitoring		
Rereading own text	No effect	Small
Rethinking task approach	No effect	Small

* Not used by both conditions.

The outcome of the think-aloud studies seems to echo the results of the quasi experiments. Experimental students in biology differed very little from control students in indicators for writing-to-learn (one positive and one negative effect), whereas experimental students in math differed substantially from control students on four indicators.

A remarkable difference found between the experimental biology and mathematics students is that the first frequently reflected on content selection (a large effect), whereas mathematics students did not (a small negative effect). This difference can be explained by the nature of the writing tasks. Biology students' writing task required them to use their knowledge of the musculoskeletal system for explaining how to act in a given situation. Therefore, students needed to select which parts of the system to use in their writing. Mathematics students, however, were instructed to describe how they solved a sum. They did not need to reflect on content selection because the content (the computation of the sum) was already there. This demonstrates that reflective activities analyzed in think-aloud studies are dependent on the specific demands that each writing task poses. One should not expect the same types of reflective activities to emerge for different writing tasks unless the cognitive demands of these tasks are similar.

4.2. Students' and Teachers' Evaluation

To explore students' and teachers' view on learning by writing, we asked them how they valued the writing-to-learn lessons. How did they experience this new way of learning? Did it support students? These questions were posed in interviews with three students in biology class, four students in mathematics class, and the teacher students that had instructed the experimental classes. Furthermore, 31 and 28 students, respectively, from the experimental biology and mathematics classes filled in a questionnaire containing four multiple choice questions.

Reactions by the students in biology class were mixed. Two (of three) interviewed experimental students found the writing-to-learn tasks quite difficult. One of them kept struggling during the experimental lessons because he did not understand what to do. The third student stated that the writing had improved her understanding of the biology subject matter involved. In their answers on the questionnaire, one (of 31) student found the writing to learn tasks difficult, 14 (of 31) students found the writing tasks a little complicated, whereas 16 students considered the tasks easy. A number of 10 students answered they had learned from executing the writing tasks, 14 students had learned a little, and seven indicated they had not learned anything.

Both biology teacher students stated that they viewed writing a good way to stimulate students' learning. In class, one teacher noticed that some students acquired new insights. However, he also had observed that students had to hurry to finish the writing tasks. He suggested planning a longer lesson duration.

The mathematics students who were interviewed stated that it does not matter which type of learning is used. However, writing took more effort. Writing a draft required deep thinking about organizing, as two students reported. One student who was interviewed remarked that giving feedback to her peer provided her with ideas just as her own writing. In this case, 10 (of 28) students answering the questionnaire found the writing-to-learn tasks difficult, 11 students found it a little complicated, and only seven students had no problem with them. However, 15 students believed they had learned from the execution of the writing tasks, while 11 students believed they had learned a little. Two students thought they did not learn anything.

The two mathematics teacher students considered writing-to-learn an effective way of learning. One teacher observed that students worked intensively when carrying out the writing tasks. She stated that more time for using writing-to-learn tasks might have made students feel more comfortable. The second mathematics teacher considered writing-to-learn a valuable addition for teaching mathematics because it is a different way of thinking than usual in mathematics.

In summary, students' reactions reflect the outcomes of the two studies. Their teachers took part in the study because they were enthusiastic about writing-to-learn and remained so. They attributed students' discomfort to a lack of time for executing the writing tasks.

5. Suggestions for Future Research

As previously discussed, research investigating effects of instruction in writing-to-learn specifically directed at low-achieving students is scarce. Nevertheless, there are good reasons to expect that low-achieving students can benefit as much from instruction in writing-to-learn as higher-achieving students [14,21,22]. Our studies give further support to this view. Although the biology study did not lead to the expected results on low achievers' learning, it certainly provided valuable information about the conditions that must be fulfilled for an elaborated approach such as GWPR to become successful. Therefore, future research allowing low achievers sufficient time on task in familiarizing themselves with the different steps of GWPR (such as recognizing genre specific conceptual relations, using them in writing tasks, and peer discussions directed at planning and revising a draft) is highly recommended. This research is essential to test the viability of assumptions underlying instruction in writing-to-learn and more specifically GWPR instruction for low-achieving students in diverse educational contexts, disciplines, and ages. The present research base needs more systematic trials to support the positive expectations of the effects of writing-to-learn for low-achieving students.

This type of future research can be strengthened by paying particular attention to the sample size, especially when the focus is on the processes of writing-to-learn. In our studies, there were only small samples selected for these analyses because of the workload involved in the analysis of think aloud protocols. In future studies, it is important that the generalizability of our findings regarding the indicators of writing-to-learn is tested with larger samples. In addition, other methods for tracking low-achieving students' writing processes such as keystroke logging [24] and eye movement tracking could be used to find additional marks of students' reflective processes while writing.

Another important issue for future research is randomization on the student level. While our studies were quasi-experimental in using intact classrooms for assignment to an experimental or control condition, a more profound design for arriving at causal interpretations of effects found is desirable. Although it is difficult to assign students in the same classroom to different conditions because of 'leakage' of the instructional components, there is a way of dealing with that problem. Care should be taken to make the materials self-instructing (such as the analysis of conceptual relations in model texts) and the lesson structure identical for both conditions. Additionally, the topics for the writing tasks should systematically differ between the conditions by using different overarching themes from different lesson series in the discipline (e.g., 'the human body' vs. 'bacteria' in biology). Effects can be analyzed by comparing topic knowledge and insight on the two different themes that students in the two conditions in the previous lessons had been writing about.

6. Pedagogical Implications

Regular mathematics lessons often consist of the teacher's explanation of new theory and thereafter students' completion of mathematics tasks. Students are supposed to learn by listening to the teacher and applying the explained theory by carrying out tasks from their textbooks. Low-achieving students may have trouble asking questions in response to the teacher's explanation because it can be difficult for them to formulate what they do not understand. This can result in their inability to carry out mathematics tasks.

Results of our studies show that low achievers benefit from writing-to-learn in mathematics as measured by their insight in how to execute mathematical tasks. Additionally, it seems that students receiving GWPR instruction show more indicators of reflective activities in a post-test writing task. Therefore, we assume that GWPR instruction for executing writing tasks allows low-achieving students to become more actively engaged in learning new theory. This engagement consists of discussions with their peers for planning, feedback, and revising a first draft directed at comprehensibility for an audience with no prior knowledge of the mathematical tasks involved. However, teachers may be hesitant to use writing in their math class because they believe that evaluating students' texts increases their workload [44]. In contrast, it must be emphasized that explicit evaluation by the teacher of students' written texts is not necessary because writing is aimed at reflecting on and learning of mathematics. In addition, the advantage of writing-to-learn tasks is that reading students' texts can inform about students' thinking, which may help teachers to focus on students' needs [44].

Teachers in the content areas (such as biology or math) do not seem at ease with developing writing-to-learn tasks for their classrooms because they are not used to writing as a tool for learning. As our studies show, writing tasks can quite easily be embedded in regular education as alternative for the more traditional approach of teaching. A more difficult aspect for content area teachers in GWPR instruction is the construction and use of model texts to demonstrate different ways of formulating genre specific conceptual relations. Implementation of that part of GWPR in classroom practice therefore requires the support of language teachers and/or the provision of a rich set of examples demonstrating the use and explanation of conceptual relations in context. Preferably, this type of support is offered on the school level as part of teachers' professional development program.

7. Conclusions

The present study was directed at writing as a tool for learning for low-achieving students in secondary education. We investigated whether GWPR instruction (Genre Writing with Planning and Revision) leads to positive effects on these students' learning.

Firstly, we investigated whether GWPR instruction focusing on explicit genre knowledge, cognitive and metacognitive strategies, and writing for a lay audience leads to more insight and topic knowledge in biology and mathematics classes compared to business-as-usual lessons (not including writing activities). Secondly, we investigated whether students that had received GWPR instruction showed more signs of the process of writing-to-learn, such as reflection directed at generating, selecting and organizing text contents, and reviewing and revising formulations.

Results showed that GWPR instruction for low-achieving biology students did not lead to effects on topic knowledge and insight, while GWPR instruction for low-achieving mathematics students did lead to a large effect on insight. Unfortunately, in mathematics classes, topic knowledge could not be measured because of school regulations prohibiting multiple choice questions in math examinations.

Think-aloud studies were carried out with a random selection of students to identify the process of writing-to-learn in both disciplines. Based on previous studies, reflective activities for planning, formulating, and monitoring are regarded as indicators of the process of writing-to-learn. Seven such indicators were identified. We investigated whether GWPR instruction had incited the process of writing-to-learn in a post-test writing task for experimental students in comparison to control students.

Table 11 summarizes the results for both studies. For biology, analysis of the think-aloud protocols showed that the activity *thinking about content selection* was used more frequently by experimental than by control students (large effect), while the activity *revising while formulating* was used more often by the control students (large effect). For mathematics, the following indicators were more frequently used by the experimental than the control students: *using knowledge about audience* (medium effect), *revising after finishing an utterance* (large effect), *rereading own text* (small effect), and *rethinking task approach* (small effect). The indicator *thinking about content selection* was used more often by the control students (small effect). Thus, in the mathematics study, experimental students showed several indications for writing-to-learn, while in the biology study they showed only one effect in favor of the experimental students and one opposite large effect in favor of the control students. These results are in line with the outcomes of the quasi experiments in the studies 1 and 2, showing no effect of GWPR instruction on learning for biology and positive effects on insight for mathematics.

Author Contributions: Conceptualization, A.V.D., A.V.G. and F.K.; methodology, A.V.D., A.V.G. and F.K.; software, inapplicable; validation, A.V.D., A.V.G. and F.K.; formal analysis, A.V.D., and A.V.G.; investigation, A.V.D.; resources, A.V.D.; data curation, A.V.D.; writing-original draft preparation, A.V.D.; writing-review and editing, A.V.G. and F.K.; visualization, A.V.D.; supervision, A.V.G. and F.K.; project administration, A.V.D.; funding acquisition, A.V.D. and A.V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from Rotterdam University of Applied Sciences.

Institutional Review Board Statement: Ethical review and approval were waived for this study because the studies involved started in 2011 under the responsibility of University of Amsterdam. At that time, there was no IRB at the University of Amsterdam and no regulations entailing approval and informed consent.

Informed Consent Statement: Informed consent (oral) was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Biology, Proportioned Means and Standard Deviations of Frequencies of Codes in Experimental (N = 5) and Control Group (N = 6), and Effect Sizes

		Experimental		**Control**		**Effect size**
Codes		Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
Planning: orienting						
1 Read. Assignment		0.03	0.04	0.05	0.06	−0.39
2 Thinking about task approach		0.03	0.04	0.00	0.01	10.03
Planning: generating						
3 Generating content		0.03	0.06	0.04	0.07	−0.15
4 Using knowledge about audience	**H ***	0.00	0.00	0.00	0.01	0.00
5 Using source text for new content		0.00	0.00	0.00	0.00	0.00
Planning: selecting						
6 Thinking about content selection	**H**	0.09	0.07	0.02	0.03	1.30
7 Selecting content		0.01	0.02	0.00	0.01	0.63
Planning: organizing						
8 Ordering ideas		0.00	0.00	0.00	0.00	0.00
9 Thinking about text structure		0.01	0.01	0.00	0.00	1.41
Formulating						
10 Thinking about formulating	**H**	0.01	0.02	0.01	0.02	0.00
11 formulating		0.21	0.20	0.28	0.13	−0.42
12 Revising while formulating	**H**	0.00	0.01	0.02	0.03	−0.89
13 Repeating formulation		0.00	0.00	0.00	0.00	0.00
14 Revising after finishing an utterance	**H**	0.01	0.01	0.01	0.01	0.00
Monitoring						
15 Rereading own text	**H**	0.02	0.01	0.02	0.03	0.00
16 Rereading assignment		0.02	0.03	0.03	0.03	−0.33
17 Using source text for correct understanding		0.00	0.00	0.00	0.00	0.00
18 Rethinking task approach	**H**	0.01	0.01	0.01	0.02	0.00
19 Checking task completion		0.01	0.02	0.03	0.01	−10.26
Evaluation						
20 Commenting own phrases		0.00	0.00	0.00	0.00	0.00
21 Commenting text structure		0.00	0.00	0.00	0.00	0.00
22 Commenting task performance incl. Writing process		0.03	0.03	0.03	0.03	0.00

Codes		Experimental		Control		Effect size
		Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
23 Commenting on assignment		0.00	0.00	0.00	0.00	0.00
24 Commenting on source text		0.00	0.00	0.00	0.00	0.00
Other activities						
25 Pause for thinking		0.12	0.10	0.16	0.14	0.24
26 Long pause (more than 50 seconds)		0.01	0.01	0.00	0.00	0.77
27 transcribing		0.12	0.03	0.07	0.02	0.00
28 Not task related remarks		0.23	0.15	0.22	0.09	−0.11
29 Expression of uncertainty		0.00	0.00	0.00	0.00	−1.55

* Hypothesized code.

Appendix B. Mathematics, Proportioned Means and Standard Deviations of Frequencies of Codes in Experimental (N = 4)) and Control Group (N = 4), and Effect Sizes

Codes		Experimental		Control		Effect size
		Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
Planning: orienting						
1 Read. Assignment		0.03	0.03	0.04	0.03	−0.33
2 Thinking about task approach		0.02	0.02	0.02	0.02	0.00
Planning: generating						
3 Generating content		0.00	0.00	0.00	0.00	0.00
4 Using knowledge about audience	**H** *	0.03	0.05	0.01	0.02	0.53
5 Using source text for new content		0.01	0.02	0.00	0.00	0.71
Planning: selecting						
6 Thinking about content selection	**H**	0.00	0.00	0.01	0.05	−0.29
7 Selecting content		0.00	0.00	0.03	0.05	−0.85
Planning: organizing						
8 Ordering ideas		0.00	0.00	0.00	0.01	0.00
9 Thinking about text structure		0.00	0.00	0.00	0.00	0.00
Formulating						
10 Thinking about formulating	**H**	0.00	0.00	0.00	0.00	0.00
11 formulating		.35	0.21	0.28	0.05	0.46
12 Revising while formulating	**H**	0.00	0.00	0.00	0.00	0.00
13 Repeating formulation		0.00	0.01	0.00	0.01	0.00
14 Revising after finishing an utterance	**H**	0.02	0.02	0.00	0.00	1.41
Monitoring						
15 Rereading own text	**H**	0.07	0.13	0.03	0.03	0.42
16 Rereading assignment		0.01	0.02	0.00	0.00	0.71
17 Using source text for correct understanding	**H**	0.00	0.00	0.00	0.00	0.00
18 Rethinking task approach		0.16	0.19	0.10	0.09	0.40
19 Checking task completion		0.01	0.02	0.02	0.02	−0.50
Evaluation						
20 Commenting own phrases		0.00	0.00	0.00	0.00	0.00
21 Commenting text structure		0.00	0.00	0.00	0.00	0.00
22 Commenting task performance incl. Writing process		0.03	0.03	0.01	0.02	0.78
23 Commenting on assignment		0.01	0.01	0.00	0.00	1.41
24 Commenting on source text		0.00	0.00	0.00	0.00	0.00

Codes	Experimental		Control		Effect size
	Mean	Std. Deviation	Mean	Std. Deviation	Cohen's d
Other activities					
25 Pause for thinking	0.08	0.03	0.13	0.09	−0.76
26 Long pause (more than 50 seconds)	0.00	0.00	0.00	0.00	0.00
27 transcribing	0.03	0.03	0.06	0.04	−0.85
28 Not task related remarks	0.13	0.10	0.25	0.08	−1.33
29 Expression of uncertainty	0.01	0.02	0.00	0.00	0.71

* Hypothesized code.

References

1. Bereiter, C.; Scardamalia, M. *The Psychology of Written Composition*; Erlbaum: Mahwah, NJ, USA, 1987.
2. Galbraith, D. Writing as discovery. *Br. J. Educ. Psychol.* **2009**, *6*, 5–26, Monograph Series II. [CrossRef]
3. Van Dijk, A.; Van Gelderen, A.; Kuiken, F. Which types of instruction on writing-to-learn lead to insight and topic knowledge in different disciplines? A review of empirical studies. *Rev. Educ.* **2022**, *10*, e3359. [CrossRef]
4. Britton, J. Shaping at the point of utterance. In *Prospect and Retrospect: Selected Essays of James Britton*; Pradl, G.M., Ed.; Boynton/Cook: Montclair, NJ, USA, 1982; pp. 139–145.
5. Langer, J.A.; Applebee, A.N. *How Writing Shapes Thinking: A Study of Teaching and Learning*; National Council of Teachers of English (NCTE): Urbana, IL, USA, 1987.
6. Mason, L. Sharing cognition to construct scientific knowledge in school context: The role of oral and written discourse. *Instr. Sci.* **1998**, *26*, 359–389. [CrossRef]
7. Klein, P.D. Reopening inquiry into cognitive processes in writing-to-learn. *Educ. Psychol.Rev.* **1999**, *11*, 203–270. [CrossRef]
8. Klein, P.D.; Boscolo, P.; Gelati, C.; Kirkpatrick, L.C. New directions in writing as a learning activity. In *Studies in Writing: Volume 28, Writing as a Learning Activity*; Rijlaarsdam, G., Klein, P.D., Boscolo, P., Kirkpatrick, L.C., Gelati, C., Eds.; Brill: Buckinghamshire, UK, 2014; pp. 1–14.
9. Bazerman, C. A response to Anthony Fleury's "Liberal education and communication against the disciplines": A view from the world of writing. *Commun. Educ.* **2005**, *54*, 86–91. [CrossRef]
10. Klein, P.D.; Boscolo, P. Trends in research on writing as a learning activity. *J. Writ. Res.* **2016**, *7*, 311–350. [CrossRef]
11. Graham, S.; Kiuhara, S.A.; McKay, M. The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Rev. Educ.Res.* **2020**, *90*, 179–226. [CrossRef]
12. Miller, D.M.; Scott, C.E.; McTigue, M. Writing in the secondary level disciplines: A systematic review of context, cognition and content. *Educ. Pschychol. Rev.* **2018**, *30*, 83–120. [CrossRef]
13. Hand, B.; Chen, Y.; Suh, J.K. Does a knowledge generation approach to learning benefit students? A systematic review of research on the Science Writing Heuristic approach. *Educ. Psychol. Rev.* **2021**, *33*, 535–577. [CrossRef]
14. Rivard, L.P. Are language-based activities in science effective for all students, including low achievers? *Sci. Educ.* **2004**, *88*, 420–442. [CrossRef]
15. Nückles, M.; Hübner, S.; Renkl, A. Enhancing self-regulated learning by writing learning protocols. *Learn. Instr.* **2009**, *19*, 259–271. [CrossRef]
16. Bangert-Drowns, R.L.; Hurley, M.M.; Wilkinson, B. The effects of school-based writing to learn interventions on academic achievement: A meta-analysis. *Rev. Educ. Res.* **2004**, *74*, 29–58. [CrossRef]
17. Newell, G.E. Learning from writing in two content areas: A case study/protocol-analysis. *Res. Teach. Engl.* **1984**, *18*, 265–287.
18. Klein, P.D.; Kirkpatrick, L.C. A framework for content area writing: Mediators and moderators. *J. Writ. Res.* **2010**, *2*, 1–46. [CrossRef]
19. Klein, P.D.; Samuels, B. Learning about plate tectonics through argument writing. *Alta. J. Educ. Res.* **2010**, *56*, 196–217.
20. Prain, V. Learning from writing in secondary science: Some theoretical and practical implications. *Int. J. Sci. Educ.* **2006**, *28*, 179–201. [CrossRef]
21. Akkus, R.; Gunel, M.; Hand, B. Comparing an inquiry-based approach known as the Science Writing Heuristic to traditional science teaching practices: Are there differences? *Int. J. Sci. Educ.* **2007**, *29*, 1745–1765. [CrossRef]
22. Faber, J.E.; Morris, J.D.; Liebermann, M.G. The effect of note taking on ninth grade students' comprehension. *Read. Psychol.* **2000**, *21*, 257–270.
23. Halliday, M.A.K.; Martin, J.R. *Writing Science: Literacy and Discursive Power*; The Falmer Press: London, UK, 1993.
24. Baaijen, V.M.; Galbraith, D. Discovery through writing: Relationships with writing processes and text quality. *Cogn. Instr.* **2018**, *36*, 199–223. [CrossRef]
25. Klein, P.D. Constructing scientific explanations through writing. *Instr. Sci.* **2004**, *32*, 191–231. [CrossRef]
26. Van Dijk, A.; Van Gelderen, A.; Kuiken, F. Writing to learn in biology and mathematics teacher education: Promoting students' topic knowledge and insight. 2022; *submitted*.

27. Veel, R. Language, knowledge and authority in school mathematics. In *Pedagogy and the Shaping of Consciousness: Linguistic and Social Processes*; Christie, F., Ed.; Cassell: London, UK, 1999.
28. Rose, D. Writing as a linguistic mastery: The development of genre based literacy pedagogy. In *Handbook of Writing Development*; Myhill, D., Beard, D., Nystrand, M., Riley, J., Eds.; Sage: Thousand Oaks, CA, USA, 2008.
29. Smits, G. *Biologie Voor Jou Vmbo-Kgt*; Malmberg: 's-Hertogenbosch, The Netherlands, 2008.
30. Thijssen, J.; Van der Schoot, F.; Hemker, B. *Balans Van Het Biologie-Onderwijs aan Het Einde van de Basisschool 4*; Cito: Arnhem, The Netherlands, 2011.
31. Boyle, G.J. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personal. Individ. Differ.* **1991**, *12*, 211–328. [CrossRef]
32. Sijtsma, K. On the use, the misuse and the very limited usefullness of Cronbach's alpha. *Psychometrica* **2009**, *74*, 107–120. [CrossRef] [PubMed]
33. Taber, K.S. The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Res. Sci. Educ.* **2018**, *48*, 1273–1296. [CrossRef]
34. Dunn, L.M. *Peabody Picture Vocabulary Test -III-NL*; Schlichting, L., Ed.; Dutch version; Harcourt Test Publishers: San Diego, CA, USA, 2005.
35. Pander Maat, H. *Tekstanalyse: Een Pragmatische Benadering*; Martinus Nijhoff: London, UK, 1994.
36. Hayes, J.R.; Flower, L.S. Identifying the organization of writing processes. In *Cognitive Processes in Writing*; Gregg, L.W., Steinberg, E.R., Eds.; Erlbaum: Mahwah, NJ, USA, 1980; pp. 3–27.
37. Reichard, L.A. *Getal en Ruimte wi/3 vmbo-K deel 2*; Noordhoff: Groningen, The Netherlands, 2011.
38. Van Berchum, C. *Netwerk: 3B Vmbo-Kader*; Noordhoff: Groningen, The Netherlands, 2010.
39. Reichard, L.A. *Getal & Ruimte: Wi/4 vmbo-KGT deel 1*; Noordhoff: Groningen, The Netherlands, 2011.
40. Ritchie, S.M.; Tomas, L.; Tones, M. Writing stories to enhance scientific literacy. *Int. J. Sci. Educ.* **2011**, *33*, 685–707. [CrossRef]
41. Martinez, I.; Mateos, M.; Martin, E.; Rijlaarsdam, G. Learning history by composing synthesis texts: Effects of an instructional program on learning, reading and writing processes, and text quality. *J. Writ. Res.* **2015**, *7*, 275–302. [CrossRef]
42. Kiuhara, S.A.; Graham, S.; Hawken, L.S. Teaching Writing to High School Students: A National Survey. *J. Educ. Psychol.* **2009**, *101*, 136–160. [CrossRef]
43. Van Gelderen, A.; Blok, H. De praktijk van het stelonderwijs in de groepen 7 en 8 van de basisschool; observaties en interviews. *Pedagog.Stud.* **1991**, *68*, 159–175.
44. Countryman, J. *Writing to Learn Mathematics: Strategies That Work*; Heinemann: London, UK, 1992.