education sciences

MDPI

Article

The Potential of Narrative for Understanding Protein Biosynthesis in the Context of Viral Infections

Jörg Zabel * and Cornelia Averdunk

Institute of Biology, Biology Education, Leipzig University, 04103 Leipzig, Germany; cornelia.averdunk@uni-leipzig.de
* Correspondence: joerg.zabel@uni-leipzig.de

Abstract: Based on the assumption that the process of understanding is partly narrative, this study explores the potential benefits and limitations of using narrative writing in biology education. We investigate what contribution a student-centered narrative intervention can make to the conceptual understanding of protein biosynthesis in the context of viral infections and virus replication. After a teaching sequence on this topic, 68 secondary school students (M = 15.7 years, SD = 0.57 years) explained virus replication in a written text. One subsample (n = 46) was instructed to write a narrative text, while the other one (n = 22) was asked to write an expository (non-fictional) text. Our data analysis encompassed an analysis of the structural narrativity in the student texts, as well as a concept-related rating of the level of scientific correctness in three categories. A post-test questionnaire (35 items) was used to depict the learners' viewpoints on their respective text production and the learning process that they experienced. Our findings indicate that most learners actually produced the text type they were supposed to, with exceptions in both sub-samples. As to the level of concept-related scientific correctness, we found no major differences between the two interventions. However, for two concepts, compartmentalization and levels of organization, the data indicate the significant advantage of the narrative intervention. We conclude from our results that to some extent, the effective learning properties of narrative texts, derived from the theoretical foundations, could indeed successfully be demonstrated in the field of virus replication. However, narrative text production is not equally beneficial for all aspects of the biological topic, and it also poses specific problems for some learners.

Keywords: narrative texts; text production; biology education; protein biosynthesis; secondary school

check for updates

Citation: Zabel, J.; Averdunk, C. The Potential of Narrative for Understanding Protein Biosynthesis in the Context of Viral Infections. *Educ. Sci.* **2024**, *14*, 521. https://doi.org/10.3390/educsci14050521

Academic Editor: Dorothe Kienhues

Received: 31 January 2024
Revised: 25 April 2024
Accepted: 6 May 2024
Published: 13 May 2024

1. Introduction

The idea of using narrative in science learning has already been discussed for many decades in science education research. The positive vision that narrative methods can improve the understanding and memorization of science content has been labeled the *narrative effect* [1]. There are theoretical foundations for this assumption, e.g., the constructivist approach to narrative by Jerome Bruner [2,3]. Consequently, the 'power of story' has become interesting for science teaching as well, and explanatory stories have been part of curricular recommendations [4]. However, there is still no comprehensive theory and too little data concerning the use of narrative in science classrooms. Meanwhile, it appears that the assumption of a narrative effect, as tempting as it may be, cannot adequately characterize the potential and the causal mechanisms related to the different forms of using narrative in learning science. Particularly, the great majority of existing studies have focused on interventions centered around the consumption of narrative material by the students, e.g., reading explanatory stories or listening to historical science stories. There is far less empirical evidence when it comes to the production of narratives by the students themselves. Therefore, the study presented here aims to contribute to this research gap by reporting results from a naturalistic quasi-experimental intervention study where students compiled narrative texts on protein biosynthesis in the context of virus reproduction. Our research design is essentially based on Jerome Bruner's theoretical approach and his notion of a *narrative mode of thought* [2]. Bruner's view on narrative can be very briefly subsumed, if not trivialized, by the phrase: 'it takes a story to make sense'. As we are adopting a constructivist perspective on science learning, we start from the assumption that, in adolescent minds, the process of making sense of viral reproduction and protein biosynthesis works even better if it is the learners themselves who compose the story.

1.1. Theoretical Framework

1.1.1. What Is a Story? More Than Just a Text Style

It is not at all trivial to define what a *story* is. Obviously, there is much more to a story than just a sequence of events. Linguists have identified some common traits and important elements of narratives [5,6]. The author composes a narration (*discours*) around a sequence of events that are temporally and causally linked (*histoire* or *fabula*). Their composition has a beginning and an end and is characterized by a certain viewpoint: even if the narrator does not add explicit comments or a 'moral' to their account, they give an implicit interpretation of the events by slowing down or accelerating the pace, using flashbacks and flashforwards, and so on. The process of turning a sequence of events into a story is called *emplotment* and is considered crucial for the overall meaning of the story. The plot represents a top-down structure that subordinates every part of the story to the storyteller's overall interpretation of the described events. Therefore, it is not the mere facts that create the meaning of the story but the way they are presented—or left out. In court for example, the same facts can be presented in very different ways for obvious reasons. Since our early childhood, we have all known what belongs to a story in general, as we dispose of a *story schema* in our minds. This *gestalt* structure of stories has been proved to facilitate memorization and recall of contents [6–8]. Narrative, through the perspective of cognitive psychology, is not just a text style, but a generalized cognitive structure of expectancy. The story schema is a mental reality that influences our perception and interpretation of the world around us. In this sense, stories can be compared to metaphors, which are no longer seen as ornaments of language, but fundamental tools of understanding [9–12]. In the past three decades, narrative psychology and related disciplines have extended the investigation of the role of narrative from the early, relatively narrow cognitive approaches to a variety of functions of the human mind, such as identity [13]. The research was influenced by the constructivist paradigm. Narrative is no longer reduced to a mode of writing or communicating, but is viewed as a mode of thought that creates an entire reality of its own. Bruner considers narrative particulars and narrative genres as a predominant way of interpreting everyday events [2]. Human beings "make sense of the world by telling stories about it" ([2], p. 130). He stresses the importance of story for meaningful understanding and the special characteristics of paradigmatic and narrative thinking [14]. The two *modes of thought*, the *narrative mode* and the *paradigmatic (scientific)* one, both have universal characteristics [2,15]. Hence, there is a gap that opens between the everyday logic applied by the learners and the scientific conceptions in the classroom.

1.1.2. Narrative Realities, and their Consequences for Learning Science

What are the consequences for science teaching? Meaningful narrative explanations cannot simply be replaced by scientific conceptions that are contradictory to the learner's experimental realism [9–11]. Instead, the key to conceptual change may lie in the use of both the narrative and the scientific mode as well as "metacognitive sensitivity" ([2], p. 149) in order to allow students not only to learn the appropriate scientific conceptions but also to give them an individual significance.

In "The Culture of Education", Bruner argues that for the very reason that we swim in a 'sea of stories', we are rarely conscious of how narratives influence our reasoning and acting [2]. Only through contrast, confrontation, and metacognition, e.g., by comparing different accounts of the same event, can we become aware that stories do not produce the truth but "provide a reasoned base for the interpersonal negotiation of meaning" ([16], p. 148).

There are still too few studies available to be able to specifically compare the effect of narration in different biological or scientific curriculum topics. To what extent does the thematic context determine the success of a narrative teaching strategy, i.e., how transferable are study results? One could assume that a fit between the structural characteristics of the topic and the story schema [6–8] can increase the learning effectiveness of the narrative intervention. If, for example, the topic already inherently offers a process-like sequence with a beginning and an end, then this linear sequence of events can in principle also be well represented narratively, which probably promotes the retention and networking of the factual structure of the target topic. Another structural aspect is the "centrality of trouble" for the narrative construal of reality ([2], p. 142): if the factual structure of the topic already suggests that a problem should take center stage, then this could make the use of stories more promising. In other words, while the narrative mode of thinking [2] is generally considered in opposition to the scientific mode, it can enable learners to engage in scientific problematization [17,18]. An empirically investigated example of such a narrative construction of *problem* and *solution* are the evolutionary adaptation processes in which populations react to striking and temporally definable changes in the environment in the course of their natural history [19–21]. It is obvious from the factual structure of the topic that a group of living beings has 'solved' an evolutionary 'problem' in a certain phase of its history, for example by colonizing a new habitat ("the whale ancestors conquered the water") and was subsequently exposed to completely new selection pressures. However, this narrative conceptualization neglects (1) that selection is individual, (2) that evolutionary fitness is always relative, and furthermore, (3) that the evolutionary race never stops, even if we humans like to perceive it episodically. In a didactic sense, or in the logic of 'highlighting and hiding', similar to metaphors [11], there would therefore be a price to pay for the narrative conceptualization of selection processes, despite or precisely because of the structural fit.

In line with Bruner's theoretical approach, our study is based on a constructivist perspective on narrative. We consider narrative as a 'construal of reality' of its own and attempt to explore the benefit of this construal for understanding the topic of virus replication in upper secondary biology education. However, even though stories are so prevalent and, according to Bruner, so essential to the process of meaning making, we do not expect a general narrative effect to be caused by our intervention. Instead, in line with Soares et al. (2023) [22], we will examine the specific effects of an intervention, closely related to the different aspects of the biological topics and to the student authors' own reflections of their writing process.

1.2. State of Research

Two major literature reviews have examined the previous evidence on narrative in science teaching in the last decades [22,23]. The number of empirical studies, however, is still quite limited, and characterized by a diversity of different contexts, target variables, and sample groups. This wide range of findings makes it impossible to draw quick conclusions on the use of narrative in science teaching. Regarding the context of protein biosynthesis and virus replication, we can find various barriers to understanding, and students' preconceptions in different studies, which should be taken into account for the intervention and discussion of our data.

1.2.1. The Impact of (the Production of) Narrative Texts on the Understanding of Biological Phenomena

Several studies examined the impact of the use of narrative texts in biology or science class on different aspects like understanding or interest and for different target groups e.g., preschoolers or undergraduates [24–28]. However, the studies' results are not consistent and moreover, only a few studies focused on the impact of not only text reception but text production on the understanding of a biological phenomenon [25–27].

As for an overview of the research field, the first significant review was presented by Avraamidou and Osborne [23]. In a non-systematic approach, they compiled a variety of aims and items of research on narrative as well as functions and effects of narrative methods. But their results also revealed that the existing web of studies often lacks connections and a systematization of the research field. Almost 15 years later, a theoretical review by Soares et al. (2023) examined 36 studies, focusing on the effects of narrating science (whether writing or reading) on memory and learning at different educational levels. The authors conclude that, in general, no narrative effect can be derived across all studies. Whether a narrative text benefits memory and learning in science class depends on the texts themselves, further activities in class, the wider context, educational level, prior knowledge, and other factors [22]. Some of the examined studies state that narrative texts can provide meaningful organization structures for activation of prior knowledge and integration of information [29,30]. However, it is also noteworthy that narrative texts can foster anthropomorphic explanations and naïve conceptions [26,30]. Nevertheless, Soares et al. wrote that most of the examined studies showed an advantage of narrative texts in some of the immediate and most of the delayed assessments of students' memory and learning from several educational levels in comparison to expository texts.

A study by Ritchie et al. (2011) with 6th graders ($n = 55$) examined the impact of writing narrative texts on the understanding of biosecurity in the context of microorganisms [27]. In a quasi-experimental intervention design, both intervention and comparison class received a learning unit on the topic of microorganisms first. Additionally, the intervention class worked with a website offering information on the topic of microorganisms and biosecurity. On this online platform the intervention class wrote and peer reviewed so-called BioStories on the same topic. Both intervention and comparison class used the website to fill in a BioQuiz as a pre- and post-test containing items for 'science interest', 'science self-efficacy', and 'familiarity with biosecurity'. The study found statistically significant differences between the intervention and comparison class in their results of the BioQuiz: writing the BioStories helped students to become more familiar with biosecurity issues and improved their interest in science. The results of the qualitative analysis of the students' written stories suggest that the students developed a deeper understanding of related biological concepts through writing the narrative texts [27].

Tomas et al. (2011) conducted an adapted study with 9th graders ($n = 152$) but without a comparison group [26]. As in the intervention study by Ritchie et al. (2011), the students participated in an intervention, including lessons about microorganisms and used an interactive website which provided more information about biosecurity and tools to write and peer-review two BioStories. Before and after the intervention, the participants filled in the BioQuiz containing items for students' 'interest in science', perceived 'general and personal value of science', 'familiarity with biosecurity issues', and 'attitudes toward biosecurity'. The study shows comparable results to the study by Ritchie et al.: the writing of the BioStories improved the students' interest in science, their perception of the values of science, and their familiarity with biosecurity issues [26].

An intervention study by Faria et al. (2015) with 4th graders ($n = 25$) examined the impact of writing diary entries on the students' understanding of the deep-sea fish's morphological characteristics and environmental adaptations [25]. The diary writing was embedded in an intervention focusing on the pioneering work of King Carlos I in oceanography and also comprised an exploration of the exhibits of two different science museums. Data were collected using observations during the intervention, worksheet answers, questionnaires, and additional interviews with teachers and students. Overall, the study results show that the students received insights into the characteristics and adaptations of organisms inhabiting the abyssal zone. It can be suggested that the intervention fostered not only the understanding of biological concepts but also encouraged the development of awareness regarding the nature of science and the process of constructing scientific knowledge [25]. However, it is not clear which part of the intervention fostered the students' understanding primarily.

To conclude, it can be said that there is no narrative effect in general [22]. From preschool age to university, according to the existing evidence, the reception or production of narrative texts in science class does not guarantee a better outcome for the understanding of scientific concepts. The impact of narrative texts depends on further factors like the text design or the wider teaching context. In view of the effects of narrative text production as a learning method in biology class, there is a lack on studies examining the relation between writing narrative texts and the understanding of biological phenomena. The studies by Ritchie et al. (2011) and Tomas et al. (2011) showed that writing narrative texts in secondary school can foster students' interest in science and their familiarity with biological issues [26,27]. Additionally, the study by Faria et al. (2015) supports the assumption that narrative text production can help students to deepen their understanding of biological concepts and of the nature of science [25].

However, the interventions often combine the activity of text writing with other learning activities like interactive website work or visiting an exhibition. As a result, it is partly difficult to ascribe the success of the intervention to the writing activity. Moreover, only a few studies examined the impact of writing narrative texts in comparison with other learning methods. In relation to that, Soares et al. (2023) encourage further research on specific narrative effects, such as (writing) which text type has a positive impact, in which context, on which target group, and on what? More detailed exploration could provide valuable insights into the conditions and causes of why different writing interventions work.

1.2.2. Students' Ideas about Protein Biosynthesis and Virus Reproduction

Situated within the 10th grade curriculum, the topic of protein biosynthesis is a mandatory component of the learning area of "genetics". The primary objective for students in our study is to understand six concepts by the end of the learning unit: flow of genetic information, the process and significance of replication, transcription and translation, compartmentalization, as well as the interaction between different organizational levels. Additionally, the students in our study were supposed to apply their new knowledge to the field of viral infections. This learning goal is linked to two more target concepts, namely virus structure and virus reproduction. So, what is already known about student conceptions in protein biosynthesis and virus reproduction?

In general, the consideration of students' pre-instructional conceptions about the biological phenomenon addressed in class is necessary for designing a fruitful learning situation [31]. Due to the low relevance of protein biosynthesis in students' everyday lives, preconceptions are usually absent for this concept [32]. Nevertheless, it is noteworthy that there are some preconceptions that students may hold in the broader context of protein biosynthesis. These preconceptions often center around associated concepts, such as the flow of genetic information, as well as ideas about the function of viruses and their effects on living organisms.

A literature review by Stern and Kampourakis (2014) on the genetic literacy of secondary students unfolds various conceptions in the context of genetics and genetic technologies [32]. Regarding the topic of our study, the review found that students are unaware of DNA providing information for the production of proteins. They also have difficulties describing the relationship between genes and DNA. The students imagine DNA as being made of proteins and tend "[. . .] to map the information included in DNA at higher levels of organisation rather than that of proteins" ([28], p. 200).

Jördens et al. (2016) also highlighted that students (n = 197; 18 years old) sometimes exhibit confusion of levels when it comes to phenomena at different organization levels [33]. This may manifest as students mixing up or disconnecting phenomena occurring at various levels of biological organization. This can also be observed in context of genetic information flow from DNA to the final protein: students (10th grade; n = 64) often have an idea of genetic instructions as containing information about both the structure and function of biological entities across multiple organization levels [34]. Moreover, students often overlook the mediating effects of proteins and their diverse functions, which are integral to cellular processes like protein biosynthesis. Another study by Lewis and Kattmann (2004) with 10th grade learners (n = 10) found similar results [35]: students may equate genes with characters, meaning they view genotype and phenotype as equivalent. An intervention study by Haskel-Ittah and Yarden (2017) also shows that 10th graders (n = 38) tend to ignore the mediating function of proteins and rather link genes directly to their encoded traits [36]. The authors found that although the students were able to illustrate the protein biosynthesis in concept maps, they were unable to connect genes and traits via proteins in their oral explanations. Similar results occurred in a study conducted by Gericke and Wahlberg (2013) with 11th graders (n = 12) [37]. Additionally, Gericke and Wahlberg found that students tend to consider the concepts of enzyme, polypeptide, and protein as synonyms and that they use the concept of polypeptide rather than protein to describe the result of translation. These pre-instructional conceptions can limit the students' understanding of why it is crucial to express genes through processes like transcription and translation and can also be obstructive for the understanding of viral protein biosynthesis [35].

Research on students' conceptions about viruses conducted by Trauschke (2008) with 7th graders (n = 11) revealed several common pre-instructional ideas [38]. Firstly, students often struggle to differentiate between living bacteria, non-living viruses (which existence depends on host cells), and other pathogens. Additionally, students may imagine pathogens as dirty, chemical, or harmful substances. Furthermore, there is a tendency to anthropomorphize viruses, perceiving them as evil organisms intentionally causing harm to their hosts, such as humans. This concept often involves the idea of a battle between the good immune system and the evil virus [38]. A study by Simonneaux (2000) about 5th graders' (n = 10) conceptions regarding microbes revealed comparable results [39]: students tended to describe microbes, viruses, and bacteria as intrusive agents fighting against defense agents like body cells or antibodies. They often used the warrior metaphor. Drawings of viruses convey an aggressive and evil impression. Although viruses and bacteria were partly described as identical, viruses have been attributed a more strategic proceeding when it comes to harmful behavior towards human bodies [39].

A latest study by Rönner et al. (2023), conducted during and after the COVID-19 pandemic, found that young students (n = 15; 10 to 12 years old) described the consequences of viruses and bacteria infection in a more differentiated way than before the pandemic: they ascribe colds and plagues to bacteria and flues and COVID-19 to viruses. They also tended to draw viruses, as well as bacteria, in a corona-like fashion. Additionally, the study concludes that the participants imagine viruses as superior microorganisms and bacteria as virus-generating cells [40].

In conclusion, the studies show that even if the topic of protein biosynthesis is not directly relevant to the students' everyday lives, there are still a large number of pre-instructional ideas that are related to this context. The studies reveal preconceptions ranging from confusion about levels of organization to misunderstandings about the flow of genetic information. Students' anthropomorphized view of viruses as evil creatures also plays a major role in their understanding of viral protein biosynthesis. It can be assumed that students will use some of these preconceptions to describe the processes of virus reproduction in their written texts in order to process the complexity of the topic and bridge missing everyday ideas. Especially for the narrative text writing, it can be assumed that anthropomorphic characteristics are being increasingly used by the participants.

1.2.3. Research Questions

Our overall research question is the following: What is the impact of a *narrative text production* on the *understanding of protein biosynthesis* in the context of *virus reproduction*, following a conventional teaching sequence, compared to writing an expository text?

This question generates several sub-questions:

(1) To what extent does the task of writing a narrative or expository text really lead to the students producing the required text type?
(2) What is the relationship between the produced text type (narrative or expository text) and the degree of scientific correctness of the written texts regarding eight concepts of protein biosynthesis in the context of virus reproduction?
(3) How do the students evaluate their own writing process, i.e., to what extent did the respective intervention help them understand the biological topic and its different aspects?

2. Materials and Methods

To examine the research questions, an intervention study with a quasi-experimental design was conducted in 10th grade. The methodology of this study is naturalistic, meaning that we conducted the survey as part of regular lessons in real classes instead of using randomized samples or laboratory conditions. Narrative or expository texts written by students during the intervention to explain protein biosynthesis in the context of virus reproduction were used for data collection. Additional data were collected by means of a post-test questionnaire with closed questions for reflection on the writing experience. With the help of a deductive analysis, the produced texts were classified as narrative or non-narrative and categorized regarding their scientific correctness. Followed by a *T*-Test and a Chi Square Test, differences in the students' explanations on protein biosynthesis and virus reproduction could be examined. The answers given in the questionnaire were analyzed statistically using a *T*-Test and a Fisher's Exact Test.

2.1. Sample and Context of the Study

The study was conducted with a cohort of 10th graders (n = 68, M_{age} = 15.7 years, SD = 0.57 years). Within the narration group (NG), 46 students were assigned the task of composing narrative texts, while the expository text group (EG) comprised 22 students tasked with writing expository texts.

Subsequently, this study was conducted with a teaching unit on the topic of protein biosynthesis in the context of virus reproduction. The topic of protein biosynthesis is embedded in learning unit 1: genetics in the biology curriculum for gymnasium schools in Saxony for 10th grade [41]. According to this curriculum, students in grade 10 are required to "apply the fields of structure and function as well as information to the cellular and molecular basis of the storage, transmission and realization of genetic information" ([36], p. 30, translated by the authors). In our course, the students were supposed to apply their new knowledge to the field of viral infections. This topic is not a mandatory part of the curriculum for 10th grade, but is already known to the students from grade 7 and is highly relevant to students' everyday lives due to the COVID-19 pandemic. The course took place before the intervention (text production task) and spanned 8 lessons of 45 min each (see Figure 1). Both NG and EG received these eight lessons from the same teacher. While the purpose of the first seven lessons was to provide students with a comprehensive understanding of protein biosynthesis, the eighth lesson was supposed to reactivate the students' previous knowledge on virus infections and make them transfer their new knowledge in the field of genetics and protein biosynthesis to this specific context. Together, these eight lessons were meant to ensure a solid foundation upon which the students could base their subsequently written texts.

2.2. Sampling Procedure

The participants were drawn from three classes at a rural gymnasium in Saxony, Germany. The sample size was determined based on practical considerations and the availability of participants within the chosen school. The allocation of participants to these groups was based on their division in regular classes. There was no option to separate the classes or to recruit a fourth class at this school.

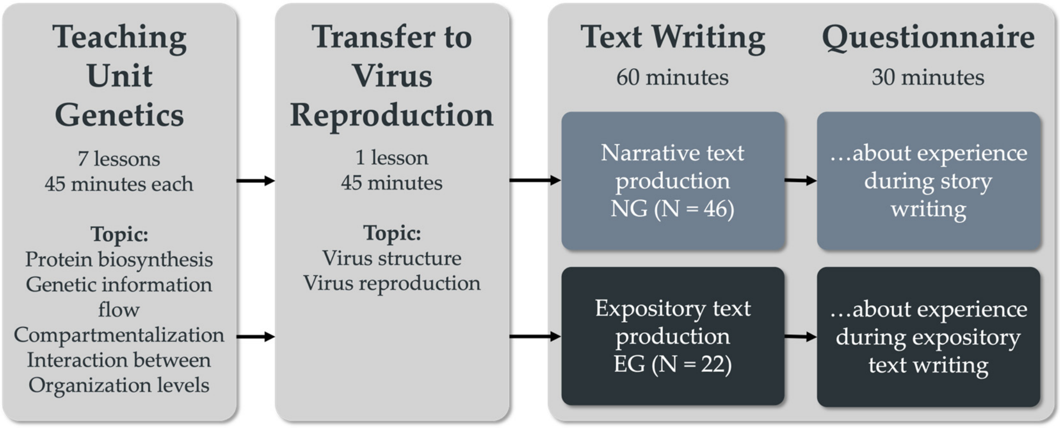


Figure 1. Study design featuring two different interventions (writing tasks).

Prior to the study, both students and their parents received all information regarding the study's objectives and processes. It was emphasized that participation is voluntary and students must actively agree to participate in the data collection. In order to adhere to data protection restrictions, analyses were conducted using anonymized texts and completed questionnaires. No supplementary information about the participants, except for their age, was gathered.

2.3. Intervention (Text Production)

The intervention itself was a text production task aiming to compare the effects of narrative and expository text writing on students' explanations of protein biosynthesis in the context of virus reproduction. The participants were divided into the two groups described above (Figure 1).

Both sub-samples were instructed to explain protein biosynthesis in the context of virus reproduction in a written text. While the initial prompt for the NG was "Explain the path and reproduction of a DNA virus in the human body by writing a story", the EG group was instructed in the following manner: "Your task is to produce a contribution for a biology textbook. Write an expository text about the path and reproduction of a DNA virus in the human body". In addition, both groups received additional information: "Use your knowledge of the following processes: Replication cycle of viruses including identical replication, transcription and translation".

The writing activity took place in a computer pool, providing students with a digital platform for text composition. In the beginning, the task was presented to the students, along with additional support cards for the NG offering inspiration for story settings and key story agents, such as ribosomes and viruses. This preparatory phase was followed by a dedicated 60 min period for text writing. Importantly, students were given the autonomy to choose whether to compose their texts on the computer or by hand, allowing for varied modes of expression. However, they were not allowed to choose whether they write a narrative or expository text. In the end, the students filled in a questionnaire to reflect on their writing experience.

Notably, no explicit introduction was provided regarding the specific properties of narrative and expository texts. It was assumed that students possessed a foundational understanding of these text types and their distinctive characteristics, acquired through previous lessons in both German and foreign language classes. It should also be mentioned that students could submit voluntarily their written texts for grading, which may have had an impact on their motivation for the writing process.

Additionally, it is worth mentioning that the use of narratives in biology classes requires a subsequent reflection on the texts to discern fictional content from scientific information, thereby fostering an awareness of their distinctions (the metacognition phase that Bruner [2] talks about, see also Zabel [19]). However, due to constraints in time resources, the reflection phase was omitted from this intervention. The reflection phase is primarily important for the development of a scientifically correct understanding and to target potential misconceptions fostered by narrative texts. For this study, an implementation and analysis of a reflection phase could give insights into the deeper meaning of the written texts. The decision to exclude the reflection phase from this study means that no conclusions can be drawn as to whether, for example, some of the texts seem scientifically inaccurate but are based on a high-level understanding may be hidden by the use of narrative language.

2.4. Data Collection and Data Analysis

The overall aim of this study is the examination of the relation between the degree of scientific correctness in the explanations of protein biosynthesis and virus reproduction and the written text type. For that purpose, the students' written narrative and expository texts and an additional questionnaire were used for data analysis. The main data analysis process was structured into three phases: firstly, the written texts were examined to determine whether the texts can be classified as (non-)narrative texts to ensure that the task of writing a narrative or expository text really led to the required text type. Secondly, the texts were examined for their degree of scientific correctness. The extent of scientistic correctness found in the students' texts was compared for the two groups, NG and EG. Thirdly, a post-test questionnaire was used to collect the students' reflective processes and writing experiences [42]. We designed the instrument to aim at the student authors' metacognition on their writing and learning processes, covering different aspects of the writing experience. In detail, the items covered three aspects: (A) the students' self-perceived increase in understanding aspects of the topic, (B) which specific concepts could be addressed in the written text, and (C) the overall utility of writing narrative or expository texts as a learning method, compared to the other text type. Thus, the questionnaire data are meant to contextualize the students' written texts and to give insights into cognitive and motivational aspects related to the text assignment.

2.4.1. Data Collection

Written Narrative and Expository Texts

This study aims to compare the impact of the text type on the scientific correctness of explanations for virus reproduction and protein biosynthesis within the written texts. The narrative or expository texts written by the students of both NG and EG provided the basis for this further analysis. The texts were produced during the intervention described in Section 2.3.

Post-test questionnaire

The questionnaire was handed to both NG and EG. In the questionnaire, the students were asked about their self-perceived increase in understanding the eight concepts (part A, 12 items), whether (partial aspects of) specific concepts could be included in the written text (part B, 16 items), and about the utility of writing narrative or expository texts as a learning method compared to the other text type (part C, 7 items). All items can be found in the results section, in detail.

2.4.2. Measurement

Narrativity of the texts

In order to examine the relationship between the written text type and the scientific correctness of the explanations within the texts, the extent to which the task of writing a narrative or expository text really led to the required text type should be explored. For this purpose, we used Zabel's "Category System for the Narrativity of Learner Texts" in our study [19]. This diagnostic system is based on aspects of narratology from literary studies [43] as well as on Bruner's constructivist approach to what he calls the two modes of thought [2]. The category system functions as a binomial determination key, comprising four distinct criteria that allow the classification of learner texts into *narrative* and *non-narrative*. The first criterion probes whether the text established a *connection between causal* *and temporal events*. The second criterion asks whether these causal connections are rooted in *underlying reasons or motives*. The third criterion centers on whether a *problem becomes the focus of attention* in the narrative's plot. The fourth criterion probes whether the text contains fictional elements and serves to differentiate between *fictional and non-fictional* narrative texts. The reliability of the diagnosis was checked communicatively. Two different coders initially worked independently of each other. In a second step, they compared their results, discussed deviations, and eventually reached consensus.

Scientific Correctness

Within the context of this study, understanding is operationally defined as the ability to explain biological phenomena [44]. This entails the articulation of scientifically accurate explanations for both virus reproduction and protein biosynthesis within written (narrative and non-narrative) texts. To examine the extent of scientific correctness in the texts, a descriptive category system was developed, based on eight selected concepts [36] (see Table 1). Six of these concepts refer to the mandatory grade 10 curriculum in protein biosynthesis and genetics, while the two other ones refer to virus structure and reproduction and thereby require the transfer of new genetic knowledge to a topic already known from grade 7 (see Section 2.1). Furthermore, two of the concepts from protein biosynthesis and genetics, namely *compartmentalization* and *interaction between different organization levels*, are not only relevant for protein biosynthesis but are also overarching key concepts of biology.

Table 1. Description of the eight concepts regarding protein biosynthesis and DNA viruses.

Concept	The Concept Contains the Description of…
Virus Structure	…viruses as organic structures with an outer protein coat with spike proteins containing DNA with genes in an inner protein coat (capsid).
Virus Reproduction	…virus reproduction as a complex process with multiple steps: virus' attachment on the host cells' surfaces (adsorption), penetration and the release of the DNA virus genome into the host cell, followed by the replication and protein biosynthesis of virus' genome into proteins, and assembly and release of new viruses.
Flow of Genetic Information	…genetic information flow as information transfer from virus DNA to mRNA to their realization into amino acids of virus proteins.
(Relevance of) DNA Replication	…DNA replication as continuously identical duplication of virus DNA by enzymes of the host cell.
(Relevance of) Transcription	…transcription as transfer of genetic information from virus DNA to functional mRNA by enzymes of the host cell.
(Relevance of) Translation	…translation as realization of mRNA into amino acids of virus proteins.
Compartmentalization	…compartmentalization as involvement of different cell components during the process of virus reproduction (cell membrane, cell nucleus, cytoplasm, ribosomes).
Interaction between Organization Levels	…interaction between organization levels as change of perspective to explain processes at subcellular, cellular, or higher levels of organization.

For the measurement of understanding, three levels of the scientific correctness of explanations can be differentiated: *scientifically correct*, *scientifically imprecise*, and *incorrect*.

- *scientifically correct* certifies that the student has a comprehensive understanding of the specific concept in accordance with the scientific ideas taught in the lesson. It means that they not only mention the concept but explain all the different aspects of the specific concept coherently.

- *scientifically imprecise* describes a basic understanding of the specific concept compared to the scientific ideas; however, with limitations in the detailed understanding or in the ability of presentation. It means that they not only mention the concept but at least describe some of the different aspects of the specific concept.
- *incorrect* certifies a lack of or a deficient understanding of the specific concept in comparison with the scientific ideas or indicates that the concept is not part of the learners' written text. It means that they only mention the concept without further detailed description of aspects of the specific concept or do not mention the concept at all.

The application of these three categories to the written texts provided detailed insights into the extent to which the students could explain virus reproduction and protein biosynthesis scientifically correctly. This analysis step was the basis for the further comparison of the impact of the written text type on the explanations' scientific correctness found in the students' texts. The reliability of the measurement was checked communicatively with the aim to reach consensus, as already described above for the narrativity of the texts.

Post-test questionnaire

The questionnaire contained three different item blocks named in Section 2.4.1. The items in part A and C use a Likert scale with the graduation *strongly disagree*, *disagree*, *neither*, *agree,* and *strongly agree*. Questions in part B could only be answered with *yes* or *no*. None of the items offered the option *no answer*.

2.4.3. Data Analysis

Narrativity of the texts

The written texts of NG and EG were analyzed by use of a deductive category system based on the "Category System for the Narrativity of Learner Texts" by Zabel [19]. The narrative and expository texts were coded according to the four categories mentioned in Section 2.4.2. When a text successfully meets the first three criteria, thereby indicating a seamless integration of causal, temporal, and motivational elements with a focal problem, it is classified as a narrative. Consequently, the third criterion emerges as a pivotal factor in classifying a text in terms of its narrativity.

This categorization process allows for a precise delineation of narrative and non-narrative texts. It should be mentioned though that this analysis step has no impact on the allocation of the students into the NG or EG. The crucial variable for this allocation in our study design is the respective writing task that was assigned to a given learner, not the actual type of text that he or she wrote as a result of this assignment. Nevertheless, the categorization of text narrativity serves two important purposes here: Firstly, it allows us to answer the question of whether the task of writing a narrative or expository text really led to the required text type, which is interesting information in itself. Secondly, knowing the type of all of the texts according to a theory-based category system will help us to better contextualize the results as to the scientific correctness of these texts, as well as their authors' writing and learning experience.

Scientific Correctness

Each written text was categorized across all eight concepts according to three distinct classifications: scientifically correct, scientifically imprecise, and incorrect. In order to statistically compare the degree of scientific correctness of the two groups and to transfer the categorization of the texts into metric values, every student received zero to two points for each of the eight concepts (scientifically correct = 2, scientifically imprecise = 1, incorrect = 0). In this way, an individual score between 0 and 16 was calculated for the overall performance during the text writing of each of the 68 study participants across all eight concepts.

To compare the results of the NG and EG, two different statistical tests were applied: Based on the individual total scores, the average score of the two sub-samples could be compared using a *T*-Test. The *T*-Test can be used because of the independence of the two intervention groups and the metric values based on the scoring. Additionally, the variances of the two groups for the mean score are homogenous.

In order to check for statistically significant differences within the frequency of each of the eight concepts between the groups, a 2 × 6 Chi-Square Test was used. For comparisons where the assumptions are not met (e.g., less than five observations per cell), we performed Fisher's Exact Test instead. In those cases where statistically significant differences are found on a global level, Bonferroni-corrected post-hoc comparisons were performed afterwards to identify for which level of scientific correctness the test yields statistically significant differences. All statistical analyses were performed with a significance level of alpha = 0.05.

Overall, the analysis allows to examine the relation between the produced text type (narrative/expository) and the extent of the scientific correctness of the explanations within the texts. Subsequently it provides an impression of the scientific correctness of the texts as a measure of the understanding of virus reproduction and protein biosynthesis.

Post-test questionnaire

The answers given in part A and C using a five step Likert scale were transferred into metric values: 1—strongly disagree, 2—disagree, 3—neither, 4—agree, and 5—strongly agree. The average score of the two sub-samples were compared using a *T*-Test. Depending on whether the variances for the specific item are homogeneous or heterogeneous, the *T*-Test for the comparison of NG and EG was performed accordingly.

To compare the given answers from part B, a 2 × 2 Fisher's Exact Test was used because there were often fewer than five observations per cell. All statistical analyses were performed with a significance level of alpha = 0.05.

3. Results

The aim of this study was to examine the impact of a narrative text production on the understanding of protein biosynthesis and virus reproduction, compared to writing an expository text. We will report our results step by step in this section and summarize them at the end (Section 3.4).

3.1. Text Analysis for Narrativity

In Both Interventions, Most Learners Actually Produced the Text Type They Were Supposed to, with Exceptions in Both Sub-Samples.

The learners' task in this study was to write either a narrative or an expository text on virus replication with reference to protein biosynthesis. Using the theory-based "category system for the narrativity of learner texts" developed by Zabel [19], all texts were checked for their text style and any deviations were identified (for anchor examples, see Table 2).

In the narrative text group, our analysis revealed that 37 of the 46 texts (80.4%) submitted could actually be categorized as narrative texts according to Zabel's category system. Of the nine texts that were identified as non-narrative, eight nevertheless exhibit superficial narrativity. This means that they exhibit features of narratives on the surface of the text, such as a narrator or opening formulae such as "Once upon a time...", but apart from such surface signals, they do not show a narrative text structure in the true sense, i.e., they do not place a problem at the center of the plot and do not indicate the motives and intentions of the acting characters [19].

In the EG, 20 of the 22 learners actually wrote expository texts. The two remaining student texts showed signs of structural narrativity according to Zabel's diagnostic system and were therefore classified as narrative. Interestingly, almost a third of the learners (7 out of 22, corresponding to 36% of the EG) used fiction signals in their texts and thus showed superficial narrativity.

Table 2. Four deductive categories for the analysis of narrativity of the students' written texts. Based on the category system by Zabel [19], with anchor examples from Hendrik's story. Underline: key text elements for the assignment of the respective category.

Category for Narrativity		Anchor Example
1	Connections between causal and temporal events	"He went to a mucosal cell and made himself comfortable on it with his protein spikes. When it was no longer cozy enough, it penetrated the mucosal cell [. . .]"
2	Causal connections that are rooted in underlying reasons or motives	"When he [virus] saw this, he became angry and wanted revenge. He devised a plan to do more harm to the human than the cell did to the virus. He began to reprogram it like a hacker so that the cell would do something very stupid for humans. Because suddenly the cell started to work and with the help of the enzymes the cell multiplied the virus DNA [. . .]"
3	Problems that become the focus of attention in the plot	"This shock unexpectedly dissolved the virus until only his DNA was left. When he saw this, he became angry and wanted revenge. He made himself a plan [. . .]"
4	Fictional elements	"Once upon a time there was a herpes virus and his name was Tim [. . .]"

3.2. Text Analysis for Scientific Correctness

3.2.1. Descriptive Statistics Suggest Differences between the Concepts Rather Than between the Interventions

Table 3 provides an overview of the scientific correctness in both intervention groups, measured with respect to the eight different concepts of the biological topic. The total number of texts sampled at the post-test was n = 68. The scientific correctness of the 8 different concepts in all these 68 texts was rated according to the classifications described in Section 2.4.2. Thus, our analysis encompassed a total of 68 × 8 = 544 individual ratings. Overall, 225 concepts in the texts (41.4%) were coded as scientifically correct and a further 169 (31.1%) were categorized as scientifically imprecise. In a total of 150 cases (27.6%), the text quality for the respective concept was classified as incorrect or missing (for anchor examples, see Table 4).

Table 3. Descriptive statistics of scientific correctness in narrative vs. expository texts. Comparison of NG and EG, individually rated for eight concepts. Categories of scientific correctness: corr = scientifically correct, impr = scientifically imprecise, inc/m = incorrect or missing. Absolute values (in brackets) are indicated below the percentages.

	Intervention "Narrative" (N = 46) Rel. Freq. (Abs. Freq.)			Intervention "Expository Text" (N = 22) Rel. Freq. (Abs. Freq.)			Total Sample (N = 68) Rel. Freq. (Abs. Freq.)		
Concepts	**Corr**	**Impr**	**inc/m**	**Corr**	**Impr**	**inc/m**	**Corr**	**Impr**	**inc/m**
Virus structure	65.2 (30)	10.9 (5)	23.9 (11)	77.3 (17)	4.5 (1)	18.2 (4)	69.1 (47)	8.8 (6)	22.1 (15)
Virus replication	52.2 (24)	19.6 (9)	28.3 (13)	27.3 (6)	22.7 (5)	50.0 (11)	44.1 (30)	20.6 (14)	35.3 (24)
Genetic information flow	19.6 (9)	30.4 (14)	50.0 (23)	22.7 (5)	22.7 (5)	54.5 (12)	20.6 (14)	27.9 (19)	51.5 (35)
DNA replication	32.6 (15)	47.8 (22)	19.6 (9)	36.4 (8)	45.5 (10)	18.2 (4)	33.8 (23)	47.1 (32)	19.1 (13)
Transcription	26.1 (12)	45.7 (21)	28.3 (13)	27.3 (6)	50.0 (11)	22.7 (5)	26.5 (18)	47.1 (32)	26.5 (18)
Translation	34.8 (16)	47.8 (22)	17.4 (8)	31.8 (7)	27.3 (6)	40.9 (9)	33.8 (23)	41.2 (28)	25.0 (17)

Table 3. *Cont.*

	Intervention "Narrative" (N = 46) Rel. Freq. (Abs. Freq.)			Intervention "Expository Text" (N = 22) Rel. Freq. (Abs. Freq.)			Total Sample (N = 68) Rel. Freq. (Abs. Freq.)		
Concepts	**Corr**	**Impr**	**inc/m**	**Corr**	**Impr**	**inc/m**	**Corr**	**Impr**	**inc/m**
Compartmentalization	78.3 (36)	8.7 (4)	13.0 (6)	59.1 (13)	36.4 (8)	4.5 (1)	72.1 (49)	17.6 (12)	10.3 (7)
Levels of Organization	41.3 (19)	30.4 (14)	28.3 (13)	9.1 (2)	54.5 (12)	36.4 (8)	30.9 (21)	38.2 (26)	30.9 (21)
Total	44 (161)	30 (111)	26 (96)	36 (64)	33 (58)	31 (54)	41 (225)	31 (169)	27 (150)

Table 4. Three categories used for analysis of scientific correctness in students' written texts, with anchor examples for the seventh concept, compartmentalization.

Category for Scientific Correctness		**Anchor Example. . .for Compartmentalization**
1	Scientifically correct	"A long, long time ago, a lonely little virus called *Herpes* lived in a small hut on the outskirts of town. [. . .] He was astonished when he stopped in front of a large building, the *Host Cell*. [. . .] The door was locked from the outside, but his hand, one of his spike proteins, fitted perfectly into the notch in the door. The moment Herpes touched the structure, he found himself inside the cell [. . .]. When one of these enzymes discovered him, he heard it call out loudly: "DNA, what are you doing here, you must be in the cell nucleus. I'll take you there." [. . .] The employees told Herpes to go to the employee named *RNA Polymerase* who was in the same section of the building [. . .] The employee sticks a sticker with the words *mRNA* on him and sends him out of the middle building. [. . .] Little Herpes arrives in the room where a doctor called *Ribosome* is apparently in charge. Unfortunately, Herpes couldn't see what he was doing there, but suddenly the employees brought her back to the main building." (story by Meike)
2	Scientifically imprecise	"Welcome to Operation DNA 3 [. . .] You are here to let us infect you with a DNA virus. Escape is not an option. [. . .] In 43 min, everyone in this room will have a DNA virus developed by us in their bodies. The DNA viruses then introduce their genetic material into the cell nucleus, where they multiply. This is called replication. In the cell nucleus, they produce new viruses, known as transcription. [. . .] Translation is the final step [. . .] [and] takes place on ribosomes [. . .]" (story by Hilde)
3	Scientifically incorrect	"They [DNA viruses] just go into my neighboring cells. [. . .] He said that they use the cells as so-called host cells [. . .] They enter the cell and release their genetic material. Many identical DNAs are formed there, but they still need a protective coating. They also get this and for this they produce mRNAs, which are then translated into virus proteins. The virus DNAs are then coated with their protective coating in the cell nucleus and then it's off to the next cell [. . .]" (story by Ben)

Evaluated across all concepts, we found only minor differences between the two intervention groups in terms of the scientific correctness of their respective texts. The most obvious imbalance between the sub-samples occurs in the highest category: While in the narrative texts 44% of the codes were coded as scientifically correct, the proportion was only 36% in the expository texts.

At this point, we noticed a trend for narrative texts to contain more scientifically correct concepts than the expository ones. The results of the statistical tests, which are presented in the following, show whether there is actually an advantage for the stories.

3.2.2. In a Score Analysis, the Two Interventions Do Not Differ as to Overall Scientific Correctness

In order to quantitatively compare the effect of the narrative vs. the expository text intervention on the two sub-samples, scores for the text quality in every concept were assigned as described in Section 2.4.3. Based on the individual scores for scientific correctness, the average score of the two sub-samples could then be compared using a *T*-Test (Figure 2).

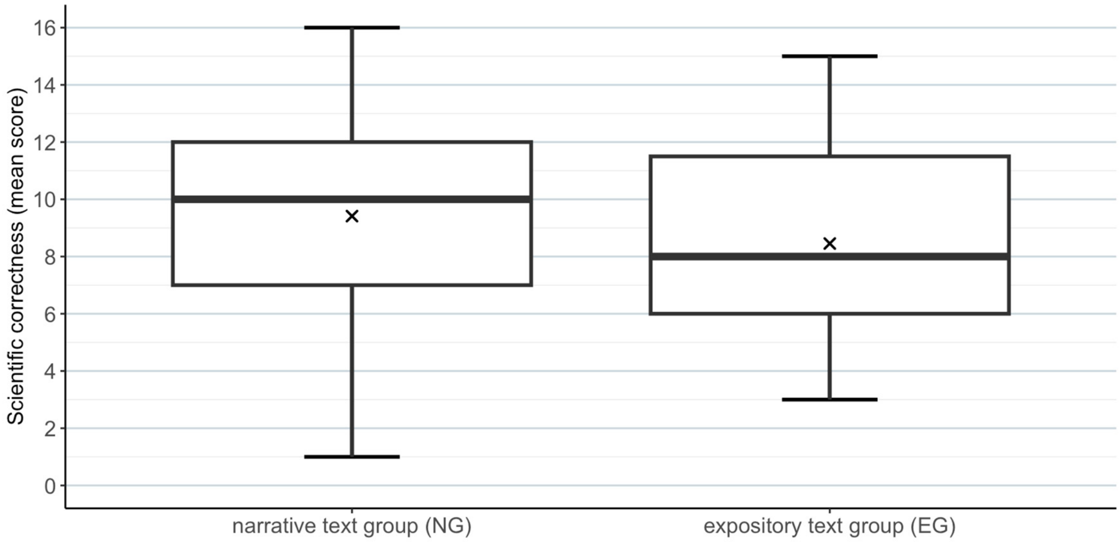


Figure 2. Boxplots for mean values of the two sub-samples for the scientific correctness of the texts. t (66) = 1.66, p = 0.158; narrative intervention group: median = 10, mean = 9.41; expository text group: median = 8, mean = 8.45; t (66) = 1.66, p = 0.158. Thick line: median value, cross: mean value.

There was no statistically significant effect for scientific correctness overall, t (66) = 1.66, p = 0.158, despite the narrative intervention group (M = 9.41, SD = 3.94) attaining higher mean scores than the students' from the expository text group (M = 8.45, SD = 3.39).

Of course, the small size of the sub-samples, especially the EG, must be taken into account with these results.

3.2.3. In a Descriptive View, There Are Considerable Differences between the Concepts Overall, but Much Less between the Interventions within These Concepts

Once the analysis is broken down into individual concepts, a more differentiated picture of scientific correctness in the texts emerges. Figure 3 shows the proportion of texts that reached the highest level of 'scientifically correct' in the respective concepts, compared between the two intervention groups. On the descriptive level, the percentage of scientifically correct descriptions differs considerably with respect to the concepts. For example, compartmentalization was characterized correctly by 72.1% of the total sample group (78.3% narrative vs. 59.1% expository). A similar picture emerges for the proportion of scientifically correct descriptions of the virus structure. In contrast, only 20.6% of the total sample managed to describe the concept of genetic information flow correctly (19.6% narrative vs. 22.7% expository). In contrast to these differences between the concepts, the differences between the interventions within the respective concept appear generally rather small. In fact, the only concept where the intervention caused statistically significant differences in the scientifically correct category is levels of organization. Here, 41.3% of NG participants reached this nominal category, compared to only 9.1% of the expository intervention.

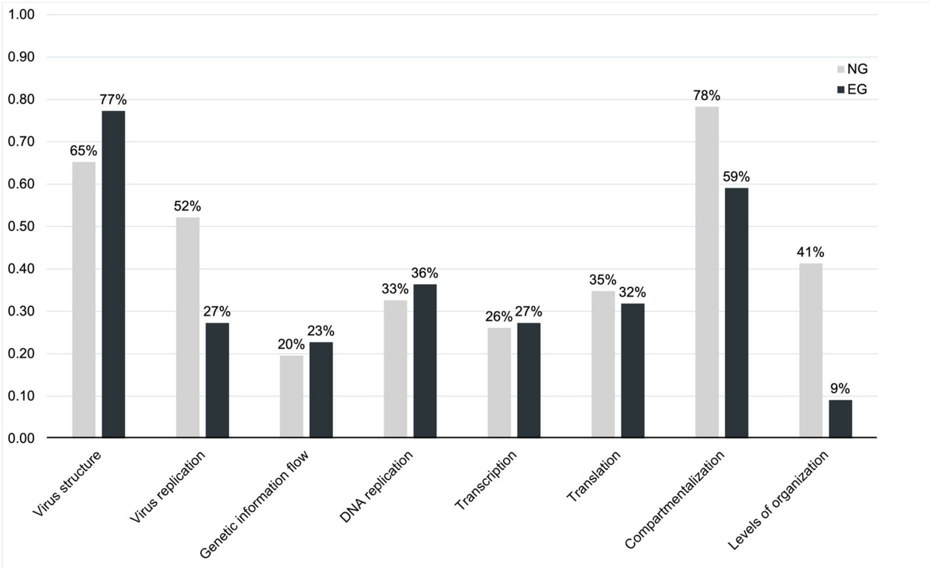
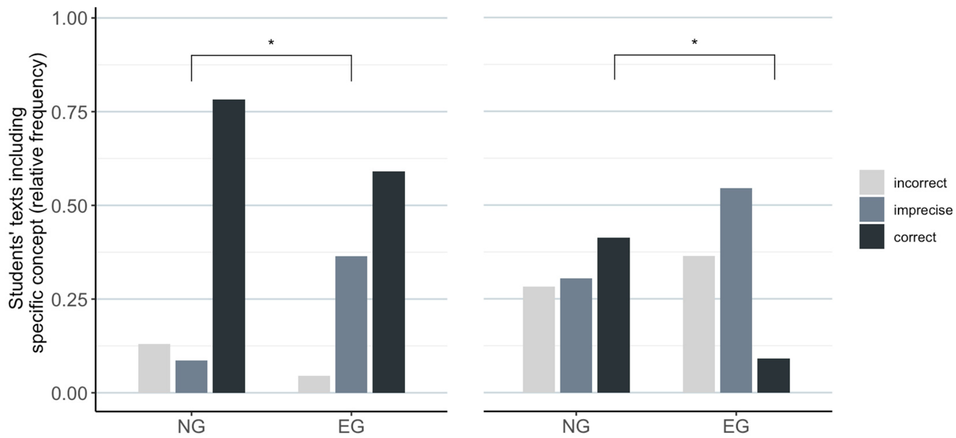


Figure 3. Descriptive statistics of the proportion of concepts rated as scientifically correct. Dark columns = NG, light columns = EG.

3.2.4. For Some Concepts, the Frequency of Scientifically Correct or Scientifically Imprecise Descriptions Differs Statistically between the Interventions

Table 3 and Figure 3 in the descriptive part already suggest that the three quality levels are not evenly distributed between the eight concepts and, in a few cases, between the interventions within one concept. In order to have a more accurate picture and to check the significance of the uneven distribution for every concept, the data obtained were tested using a Chi-Square and Fisher's Exact Test as described in Section 2.4.3.

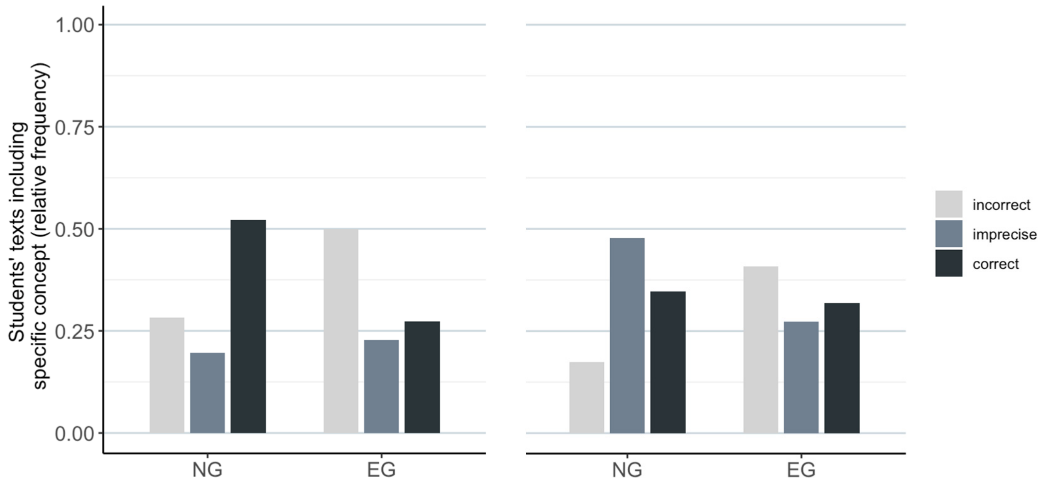
Only in two cases were statistically significant results obtained: (1) for the concept *levels of organization*, the number of narrative texts classified as scientifically correct is statistically significantly higher than the number of expository texts that achieved this classification χ^2 (2, N = 68) = 7.58, p = 0.0226 (Figure 4a). (2) The second statistically significant deviation concerns the concept *compartmentalization*. Here, there were statistically significantly more assignments to the medium quality level scientifically imprecise in the expository texts than in the narratives χ^2 (2, N = 68) = 8.26, p = 0.022 (Figure 4b).



(**a**) Compartmentalization (**b**) Levels of organization

Figure 4. Frequency of three levels of scientific correctness for the two concepts compartmentalization and levels of organization. (**a**) Compartmentalization, χ^2 (2, N = 68) = 8.26, p = 0.022. (**b**) Levels of organization, χ^2 (2, N = 68) = 7.58, p = 0.0226; *: $p < 0.05$.

Apart from compartmentalization and levels of organization, the frequencies of the quality levels found did not differ statistically significantly from the expected distribution in any other of the eight concepts. Nevertheless, tendencies in favor of the narrative intervention can be seen in two other concepts, namely *virus replication* χ^2 (2, N = 68) = 4.16, p = 0.133 (Figure 5a) and *translation* χ^2 (2, N = 68) = 4.86, p = 0.099 (Figure 5b). In the case of virus replication, the proportion of scientifically correct narratives is almost twice as high as that of expository texts (52.2% vs. 27.3%). In the case of the translation concept, the narratives show, above all, a much lower proportion of texts with incorrect or missing descriptions (17.4% vs. 40.9%).



(**a**) Virus replication (**b**) Translation

Figure 5. Frequency of three levels of scientific correctness for the two concepts virus replication and translation. (**a**) Virus replication, χ^2 (2, N = 68) = 4.16, p = 0.133. (**b**) Translation, χ^2 (2, N = 68) = 4.86, p = 0.099.

3.3. Analysis of Post-Test Questionnaires

The post-test questionnaire is aimed retrospectively at the student authors' experience of their own writing process. It is divided into three parts that cover different overall questions: how helpful was the respective text type in understanding certain aspects of the topic (part A, items 1–12), which sub-aspects were the learners able to address in their text (part B, items 13–28), and finally, a more global assessment by the learners of their respective text type and its usefulness for understanding biological content (part C, items 29 to 35).

3.3.1. The Text Authors Estimate That Stories and Expository Texts Promoted Their Comprehension to Roughly the Same Extent

Table 5 shows the students' metacognitive statements on the effect of the intervention. The learners rated their own writing process on a scale of 1–5 after writing the text, differentiated according to the eight concepts and four more aspects of virus replication (Table 5, No. 9–12) and protein biosynthesis (Table 5, No. 1–8). While most of the items focused on individual sub-steps or structures, items 3, 11 and 12 were aimed at larger contexts or networks of levels and sub-steps.

A statistically significant difference between the text types only occurred for item 4. When it comes to understanding the process of DNA transcription, the learners in the expository text intervention (M = 3.48, SD = 0.91) rated their writing process as statistically significantly more helpful than the students of the NG (M= 3.00, SD = 0.76), t (66) = 2.232, p = 0.014. In all other aspects, however, including the more cumulative or summarizing aspects, there was no advantage for either type of text from the authors' metacognition perspective.

Table 5. Students' metacognitive statements on the effect of the intervention on their understanding of specific biological topics. Post-test questionnaire, part A. Significant t-values are highlighted in gray and marked with asterisks. * $p < 0.05$.

Writing the Story/the Expository Text Has Improved My Understanding of the . . .	Mean NG	Mean EG	SD NG	SD EG	t-Value	*p*-Value
(1) . . .structure of the cell nucleus	3.37	3.77	1.07	0.79	1.549	0.063
(2) . . .DNA replication	3.48	3.59	0.85	1.03	0.468	0.320
(3) . . .genetic flow of information (from the gene to mRNA to protein).	3.59	3.73	0.92	0.91	0.580	0.282
(4) . . .transcription process	3.00	3.48	0.76	0.91	2.232 *	0.014
(5) . . .relationship between codon and anticodon.	2.43	2.91	0.80	1.28	1.566	0.061
(6) . . .function of the ribosomes	3.09	3.00	0.95	1.00	0.342	0.367
(7) . . .genetic code	2.74	3.09	1.03	1.20	1.227	0.112
(8) . . .translation process	3.37	3.55	0.96	0.89	0.711	0.234
(9) . . .structure of viruses	4.24	4.27	0.91	0.91	0.140	0.444
(10) . . .the reproduction cycle of viruses	4.30	4.36	0.83	0.83	0.272	0.393
(11) . . .protein biosynthesis as a whole	3.38	3.71	1.02	1.08	1.233	0.111
(12) . . .connection between processes at different levels of organization (DNA, cell nucleus, cell, organ, organism).	3.52	3.68	0.88	0.82	0.708	0.241

3.3.2. According to the Participants, 7 out of 16 Aspects of the Topic Were Addressed More Frequently in Expository Texts, Only 1 More Frequently Than in Narrative Texts

We compared the two sub-samples in terms of whether or not the students claimed to have addressed any of the 16 different aspects of protein biosynthesis and virus replication in their text (Table 6). In contrast to parts A and C (items 1–12 and 29–35) of the questionnaire, no graded response options were provided here in part B, but only yes or no. The differences in frequency were therefore tested for significance using a 2 × 2 Chi Square Test and Fisher's Exact Test.

Table 6. Abundance of 16 sub-aspects in the 2 text genres, based on the authors' metacognitive statements. Post-test questionnaire, part B: Item text NG: "Which topics or 'protagonists' were you able to address/process in your story (using suitable metaphors and symbols)?"; Item text (EG): "Which thematic areas and functional components were you able to address/process in your text?"; χ^2 (1, N = 68). Significant χ^2-values are highlighted in gray and marked with asterisks. * $p < 0.05$; ** $p < 0.01$.

Thematic Area	NG Abs. Freq.		EG Abs. Freq.		χ^2	*p*-Value
	Yes	No	Yes	No		
Cell nucleus (structure)	23	23	5	17	4.57 *	0.033
Enzymes of DNA replication	27	19	14	8	0.15	0.699
RNA replication (process)	31	15	20	2	4.39 *	0.036
Genetic information flow	21	25	17	5	6.04 *	0.014
Transcription (process)	22	24	18	4	7.10 **	0.008
mRNA	28	18	18	4	2.98	0.084

Table 6. *Cont.*

Thematic Area	NG Abs. Freq. Yes	NG Abs. Freq. No	EG Abs. Freq. Yes	EG Abs. Freq. No	χ^2	*p*-Value
Translation (process)	22	24	16	6	3.74	0.053
Ribosomes	22	24	13	9	0.76	0.383
tRNA	6	40	9	13	6.72 **	0.009
Amino acids and protein	16	30	11	11	1.44	0.230
Codon and anticodon	5	41	9	13	8.21 **	0.004
Start codon, stop codon	3	43	6	16	5.58 *	0.018
Genetic Code	6	40	7	15	3.39	0.066
Structure of viruses	35	11	19	3	0.96	0.327
Reproduction of viruses	43	2	21	1	0.10	0.752
Change of organization levels	16	30	9	13	0.24	0.624

According to the students' own assessment, 9 out of 16 sub-aspects were addressed equally frequently in both types of text. According to the students' answers, six other topics were integrated statistically significantly more frequently in the explanatory texts than in the stories. In contrast, the aspect *structure of the cell nucleus* was processed more frequently in stories than in expository texts, χ^2 (1, N = 68) = 4.57, $p = 0.033$.

3.3.3. The Students' Metacognition Does Not Indicate Any Advantage for One of the Two Text Types, unless They Are Compared Directly

Table 7 compares both sub-samples with regard to the subjective statements of the participants on the potential of the text type assigned to them for learning. Items 29 to 35 were aimed at the question of how helpful the students rated 'their' text type (narrative or expository text) for understanding biological relationships or learning in the natural sciences.

Table 7. Students' metacognitive statements on the effects of their respective intervention. Post-test questionnaire, part C, five-step-Likert scale: 1 = strongly disagree, 2 = disagree, 3 = neither, 4 = agree, 5 = agree strongly. The bracket [own text type] is a placeholder for "a story" (NG) or "an expository text" (EG), respectively; Item NG 34 only included in the NG questionnaire. Significant t-values are highlighted in gray and marked with asterisks. ** $p < 0.01$.

Item	Mean NG	Mean EG	SD NG	SD EG	t-Value	*p*-Value
(29) I find it useful to write [own text type] in biology lessons.	3.26	3.09	1.13	1.08	0.579	0.282
(30) Writing [own text type] helped me to understand and comprehend protein biosynthesis as a whole.	3.37	3.23	1.09	1.28	0.469	0.320
(31) I can understand and comprehend biological content more easily with the help of [own text type] than with the help of [other text type].	3.54	2.73	1.19	1.01	2.730 **	0.004
(32) I would like to write [own text type] about biological issues more often in class in the future.	2.85	2.64	1.41	1.30	0.583	0.281
(33) It was easy for me to write a [own text type] about the reproduction of viruses.	3.16	3.50	1.25	0.90	1.144	0.128
(34/35) By writing the [own text type], I have dealt with the topic in depth.	3.78	3.71	1.14	0.88	0.244	0.404

Table 7. *Cont.*

Item		Mean NG	Mean EG	SD NG	SD EG	t-Value	*p*-Value
(34 NG)	When learning scientific content, I have already occasionally used stories (story-like elements) (also mentally).	2.27	-	1.31	-	-	-

As a result, there was no advantage in favor of one of the two text types from the respondents' point of view. One exception is item 31, the only item in which one's own text type is directly compared with the other text type. Here, a statistically significantly larger proportion of the narrative intervention found the text type 'story' to be more conducive (M = 3.54, SD = 1.19) to understanding with regard to biological content than participants from the EG (M = 2.73, SD = 1.01), t (66) = 1.66, p = 0.004.

3.4. Summary of Results

3.4.1. Narrativity and Scientific Correctness of the Texts

In both interventions, the majority of the students actually produced the text styles that they were asked to write. However, there were exceptions in both sub-samples. For example, in the narrative intervention, 9 out of 46 texts were actually not true narratives as they did not place a problem at the center of the plot and did not indicate the motives and intentions of acting characters [19]. In the expository text group, on the other hand, 2 out of 22 student texts were classified as narratives according to the criteria of structural narrativity mentioned above. Furthermore, seven texts in this intervention are superficially narrative, as they contained fiction signals.

As to the scientific correctness of the texts, the overall impression is that the influence of the intervention was rather small. While the students performed quite differently across the eight concepts of protein biosynthesis and virus reproduction, within most of these concepts it made no difference whether the students were supposed to write a narrative or an expository text. This impression from the descriptive viewpoint was confirmed by the score analysis, as the mean values of scientific correctness do not differ between the two interventions.

However, in some cases there are interesting exceptions from this overall picture. The observed frequencies in the 3 × 2 matrix allowed us to reject the null hypothesis for two out of eight concepts, namely levels of organization and compartmentalization. In the case of levels of organization, the narrative intervention led to a higher frequency of scientifically correct descriptions in the texts, χ^2 (2, N = 68) = 7.58, p = 0.0226. For the concept compartmentalization, the only statistically significant difference in the matrix is that scientifically imprecise descriptions were much more frequent in expository texts than in narrative ones, χ^2 (2, N = 68) = 8.26, p = 0.022. However, the overall picture of this matrix suggests that the high proportion of scientifically correct descriptions in the narratives (78,3% vs. 59,1%, see Figure 4a) has contributed considerably to this picture.

Furthermore, tendencies in favor of the narrative intervention can be observed in two other concepts, namely virus replication and translation. In the case of virus replication, the proportion of scientifically correct narratives is almost twice as high as that of expository texts (52.2% vs. 27.3%). Still, this difference is not statistically significant, which is probably also due to the small sample sizes.

3.4.2. Post-Test Questionnaires

The post-test questionnaire was aimed retrospectively at the authors' experience of their own writing process, directly following the writing exercise. The students estimate that stories and expository texts promoted their comprehension to roughly the same extent. A statistically significant advantage for the expository text intervention (M = 3.48, SD = 0.91) only occurred for one aspect, the process of DNA transcription rather than the narrative intervention (M = 3.00, SD = 0.76), t (66) = 2.232, p = 0.014. In all other aspects, however, including the more cumulative or summarizing aspects such as *protein biosynthesis as a whole*, there was no advantage for either type of text from the students' metacognitive perspective.

According to the participants, 7 out of 16 aspects of the topic were addressed more frequently in expository texts, and only 1 (cell nucleus structure) was more frequently addressed in stories.

In their overall judgement, the students see no advantage for one of the two text types, unless these types are compared directly, as in item 31 "I can understand and comprehend biological content more easily with the help of X than of Y", t (66) = 1.66, p = 0.004.

4. Discussion

The overall research goal of our study was to examine the impact of a narrative text production on the understanding of protein biosynthesis and virus reproduction, compared to writing an expository text. We see the relevance of our results in being able to provide indications for (1) in which specific phases or partial aspects of learning and understanding the production of one's own stories could possibly reveal potential for teaching biology, (2) where exactly this potential lies, and (3) which mechanisms of action the patterns found in the data possibly speak for. However, before we discuss our findings in detail, we will revisit the current study situation in order to recall the gaps in previous studies and some of the challenges that characterize this research field.

4.1. Research Gaps and Challenges of the Field

As outlined in the introduction, our knowledge about the effect of narration on learning and understanding in science lessons is still very patchy. A reliable overall picture from empirical results of intervention studies is not yet available and the results are inconsistent [22,23]. The reasons for this are manifold. We will roughly explain some of these reasons here at the beginning of the discussion section in order to better assess the significance of the findings described further below. So, what are the obstacles and challenges that characterize the field of research in general? We see three main ones.

Firstly, there is the complexity of the phenomenon of narration itself, which is not easy to delimit empirically. Depending on the discipline in question, e.g., linguistics [45] or literary studies, there are different definitions of what a story is. Overall, the theoretical perspectives on narration are drawn from many academic disciplines; e.g., narrative psychology, psycholinguistics, and classical narrative theory. For this study, we chose a theoretical approach that focuses on the construction of knowledge [2,14]. Bruner's definition of narrative as a universal everyday mode of world construction places it at the center of teaching. However, this upgrading from a constructivist perspective also brings empirical problems with it. Because the narrative mode, as Bruner describes it [2,14], deeply permeates and determines our everyday thinking, it is practically impossible to create a control group that is narrative-free in the strict sense within an empirical study.

A second research gap is linked to the nature of the intervention, in particular the question of whether it is receptive (i.e., learners are given a prefabricated story to read or listen to) or whether they are asked to write a story themselves, as in this study. The vast majority of existing studies investigate the effect of receptive interventions. From the constructivist perspective on narrative outlined above, it would make sense to encourage students to construct new knowledge through productive tasks.

A third complication, which makes the overall research situation less meaningful, arises from the diversity of the target variables in intervention studies on narration. What exactly the study authors want to promote with their intervention differs from case to case, as does the role of the narrative phase for the learning arrangement as a whole. The spectrum ranges from prefabricated stories, which are intended to promote motivation as an introduction to a new teaching topic without affecting the actual learning phase, to narratives written by the students themselves, in which explanations for phenomena are designed and which are aimed directly at a conceptual change process in the sense of constructivist understanding tools [20,46].

Against the background of all these limitations and difficulties, we cannot speak of a well-defined research gap that could be closed with a few empirical studies. Rather, we are only at the beginning of what we might cautiously call an 'enlightened' phase of research into narration in science didactics. We will return to this idea at the end of the discussion.

4.2. The Biological Topic and Methodology Used in This Study

As we explained in Section 1.2.2, the topic of protein biosynthesis and virus reproduction bears some inherent challenges for the learners. It is a compulsory topic in the biology curriculum and certainly somewhat demanding, as it requires at least basic knowledge in cell biology, molecular biology, and biochemistry. Particularly the size and time scale of the structures and processes make it impossible to experience them directly. Furthermore, a basic understanding of the interaction between different organizational levels is a prerequisite and a difficulty [33] linked to the rather abstract concept of genetic information flow from DNA to the final protein [34]. Students tend to confuse genes with characters, mixing up genotype and phenotype [35]. As an additional challenge, many of the technical terms used to explain biosynthesis, such as ribosomes or mRNA, are new, complex, and hard to contextualize for the students. In contrast to the more organismic curricular topics such as evolution or ecology, cell biology uses dimensions and time scales that are difficult for the learners to imagine. While there do exist preconceptions about viruses [38], we still assume that, e.g., in contrast to evolutionary phenomena on a macroscopic scale that are linked to well-known animals or plants, the topic of protein biosynthesis and virus reproduction is not likely to facilitate a narrative conceptualization in a world close to human experience, as most of the structures and processes in this microscopic and molecular realm are not directly accessible and hard to make sense of or even imagine for the students.

As to the methodology of this naturalistic study, two sub-samples were compared with each other who had written different types of texts about the previously covered learning material. In contrast to previous studies that are roughly comparable [20], the students in our study were not allowed to choose the text type but were randomly assigned to one of the two interventions instead. From our own years of teaching experience not only in biology, but also in German or Ethics lessons, we assume that while there are in fact some pupils who really enjoy writing stories, especially in lower secondary school, many students tend to reject stories and prefer concise and factual forms of written expression instead. Some learners do not write long texts simply because they generally do not feel confident in written expression and like to keep writing at school to a minimum. Although there is no proof, as these preferences were not covered by our questionnaire, chances are we might have forced a considerable proportion of the sample group to write a narrative text against their own preferences. Furthermore, many students presumably did not fancy writing an extensive text in a science lesson at all.

Another distinctive feature of our intervention is the fact that it only encompassed the final phase of the teaching unit, not the main one. It was not a 'narrative teaching unit' overall. Instead, the narrative method was only used in the sub-sample in order to engage the students in subsuming the lesson content and make sense of it in a productive and constructive manner. We need to keep this in mind while discussing the findings in detail.

4.3. Discussion of the Results, Related to the Research Questions 1–3

The overall goal of this study was to explore the potential of a productive narrative method for understanding a classical biological topic. We will now discuss our findings, guided by the research questions formulated in Section 1.2.3.

4.3.1. To What Extent Does the Task of Writing a Narrative or Expository Text Really Lead to the Students Producing the Required Text Type?

Even if instructed to write a narrative, roughly 20% of the student authors did not produce a structural narrative according to the criteria by Zabel [19]. On the other hand, 10% of the student texts in the expository text intervention were actually narratives. From a methodological point of view, this aspect could also be called *intervention fidelity*, but of course there are some grey zones to this fidelity when it comes to narrative text writing. Stories are characterized by a temporal and intentional causal linking of events; i.e., also by a central problem, its complication, and the corresponding approaches to resolving the problem situation [2,19,43]. The criterion of structural narrativity is therefore suitable as a decisive factor for classifying a text in terms of its narrativity. As we reported in our results, while the great majority of the 10th graders was able to produce the text style they were supposed to, there were deviations in both directions. Some students from the NG produced texts that only contained superficial but no structural narrativity, whereas some students from the EG actually included narrative elements in their texts. As we found no evidence in the data that would support other assumptions, we tend to assume here that both deviations from the text style in the instruction happened more or less unwillingly. Given that our intervention did not encompass any further explanation of the text style but simply relied on the student authors' previous knowledge, these deviations probably reflect the sample group's writing competence, more precisely their ability to use the two genres narrative and expository text within the respective limits and rules. Consequently, we conclude from our results that (1) the intervention fidelity in our sample group (10th graders from a German gymnasium) appears sufficiently high in order to make a real differentiation between the two interventions, and (2) that it seems to be both reasonable and empirically feasible to identify 'real' narratives in a constructivist sense by using the criterion of structural narrativity. The distinction between structural narrativity and superficial story features appears significant for future studies, in particular if they adopt a constructivist approach similar to the one chosen here.

4.3.2. What Is the Relationship between the Produced Text Type (Narrative or Expository Text) and the Degree of Scientific Correctness of the Written Texts Regarding Eight Concepts of Protein Biosynthesis in the Context of Virus Reproduction?

As we reported in Section 3, the overall descriptive view shows differences between the eight concepts rather than between the two interventions. Actually, in a score analysis, the two interventions did not differ as to the overall scientific correctness of the texts produced by the students. We conclude from these results that the influence of the intervention was rather small. Within most of the eight concepts it made no difference whether the students were supposed to write a narrative or an expository text. Apparently, the concept itself and its representation in the teaching sequence were more influential than the text style when it comes to scientific correctness. The exceptions were the two concepts of levels of organization and compartmentalization, where the narrative intervention showed some advantages. Furthermore, there were noticeable tendencies in favor of the narrative intervention for virus replication and translation.

As far as the two concepts levels of organization and compartmentalization are concerned, the results reported here encourage a special consideration. It is conceivable that the narrative text format systematically foregrounds the organizational level of the organism in the text, because stories tend to deal more often with people and their experiences than with viruses and molecules. Primo Levi's famous novel "Carbon" [47] provides a counterexample, but perhaps this masterpiece is the exception to the rule. In a similar way, the concept of compartmentalization could also be given more weight in narrative texts, because the narrative mode turns the compartments into places of a narrated plot, and thus into important narrative elements. Whether certain basic concepts benefit from non-fiction or narrative texts could be an interesting question for future studies.

All in all, our impression as to this research question is that (1) the non-scientific text style of the narrative apparently was no hinderance for the students in expressing their biological knowledge on the topic in a scientifically correct way, even though most of them actually engaged in writing 'real' stories that featured structural narrativity. Furthermore, for some of the concepts at stake, the narrative intervention even improved the text quality in terms of scientific correctness. This is particularly interesting in the case of levels of organization, a crucial and demanding concept that is representative of the difficulty of this learning area (see Section 2). Due to their typical and self-contained beginning–middle–end structure (see Section 1), narratives offer a tangible structure for the integration of scientific content. This 'template' is particularly suitable for the presentation of processes with a definable start and end, such as virus replication. The structure facilitates the meaningful processing, organizing, and linking of information [48,49]. This possibly can be related to the positive tendency of the narrative intervention group for the concept of virus replication. The concepts of compartmentalization and levels of organization, however, cannot easily be framed in this explanation, as they do not contain an obvious beginning–middle–end structure.

In addition to these interesting details in the data set, our findings do not indicate a general advantage of the narrative intervention. We assume that the small sample size, coupled with the methodological conditions described in Section 4.2, i.e., the very limited extent of the narrative intervention in our design and the fact that the students did not have a choice of text styles, contributed to the overall effect on scientific correctness being rather weak. It is remarkable, against this background, that some statistically significant advantages (and no disadvantages) of the narrative intervention could be found.

4.3.3. How do the Students Evaluate Their Own Writing Process, i.e., to What Extent Did the Respective Intervention Help Them Understand the Biological Topic and Its Different Aspects?

According to the questionnaire data, the student authors in our study assessed that stories and expository texts promoted their comprehension to roughly the same extent. Based on the memory of these young text authors, 7 out of 16 sub-aspects of the topic protein biosynthesis and virus replication were addressed significantly more often in expository texts, and only 1 sub-aspect was addressed more frequently in narrative texts. The students' own assessment does not indicate any advantage for one of the two interventions, with one exception. If the text styles are compared directly, as in item 31, the students believe narratives to be more conducive to understanding biological content than expository texts. The majority of items in the questionnaire, however, do not support this advantage for stories. It is therefore possible that the students answered item 31 with a more general scope, as the wording of the item does not refer to the actual text that they just wrote but rather provokes a general statement. It could even be understood as a comparison of given (not self-produced) text types. So, we need to state here, as a conclusion, that in the eyes of the students, neither the narrative nor the expository text intervention in our study had any advantages over the other. Interestingly, although the authors of narratives claim to have mentioned significantly less concepts in their texts, this memory does not resonate with analysis of their texts as to scientific correctness. If in fact so many concepts had not been addressed in the NG groups' texts, this would have had a negative impact on the results as to the scientific correctness. Actually, the authors in the NG group seem to underestimate the scientific quality of their texts, at least the range of topics mentioned therein, which is an interesting observation.

4.4. Implications for Research on Narrative in Science Teaching, and for Teaching Practice (RQ 4)

At the beginning of the discussion, we characterized the research situation in order to be able to better classify our own findings within it. For the current phase, the latest review article by Soares et al. (2023) obviously marks the transition to a more unexcited, demystified phase of research into the phenomenon of narration in science didactics. The present study can make a contribution to this clarification and patient measurement, also and precisely because its results are rather unspectacular. Our study did not aim to clarify the question of whether or not narrative should have a place in science lessons. This would be completely out of place due to the small sample size alone. Other studies have chosen a more comprehensive design that corresponds much better to an overall narrative intervention [26,27]. It is all the more surprising that the minimal intervention in our study brought significant advantages in favor of the narrative texts in at least two of eight concepts, and at least tendencies in two other concepts. It could therefore be said that the design of the study was more of a scalpel than a mallet from the outset. Our aim was not to make a broad narrative intervention potentially reveal a strong contrast in the comparison between 'ordinary' teaching in the EG and a progressive narrative method, which could then be positioned as evidence for the effectiveness of narrative in general. Rather, our aim was to look with a magnifying glass and to obtain indications of possible causal relationships through the differentiated concept-related examination of the texts and the precise questioning of the text authors. We estimate that our results will encourage more small-scale naturalistic studies like this, given that our intervention could be integrated relatively easily into regular lessons and yet provided a differentiated picture of the effect of a narrative phase on certain aspects of a biological topic. For example, it would be interesting to know whether the ability of students to cope with the widespread confusion of the levels of organization [33] can also benefit from narrative interventions in other contexts or in bigger studies.

As to the question of which conclusions can be drawn as to the future use of narrative for the development of biology lessons (a) in the context of virus reproduction and (b) in general, we conclude from our findings that it is too early for general recommendations regarding the use of narrative in the science classroom—at least if they are to be empirically sound. Rather, our results once more underline the fact we need to clarify exactly what we expect from a narrative intervention and how we can measure precisely this effect. Still, we think it is remarkable that despite the gap between the narrative and the scientific mode that Bruner describes, and even despite the lack of a reflective phase in our lesson plan, the narrative intervention in our study did not affect the results of the lessons, but rather showed a few positive effects. This means an encouraging message for the teaching practitioners: they should feel free to experiment with individual narrative phases even in a conventional setting. It will not immediately jeopardize the success of their lessons just because narrative, at first glance, seems so out of place in a science class.

To what extent can our findings be transferred to other scientific topics? The empirical research situation is currently still too thin and the respective study designs are not comparable enough to make clear assignments as to which topics are particularly suitable for narration and which are less so. In addition, the declared goal of the educational processes is also important here: If one decides to use narratives in science lessons for reasons of developmental psychology, e.g., because learners are currently located in the romantic phase [48], then the objective of a narrative intervention goes beyond the classical learning objectives of the biological curriculum in a narrow interpretation. If, on the other hand, one concentrates more on cognitive understanding processes and the acquisition of specialist knowledge, then predictions about the effectiveness of narration in individual subject areas could refer to an assumed fit between the factual structure of the topic at stake and the story schema, as we pointed out in the introduction section. What does this mean for the interpretation of our results here? First of all, the technical structure of protein biosynthesis and virus replication basically offers a suitable form for a linear narrative. Secondly, potential main characters in the plot, such as the virus and its host cell in our case, also facilitate narrative realization and anthropomorphic representation. Since there is also an antagonistic relationship between them, this promotes the use of a good-and-evil scheme, as the students' texts show. An in-depth, case-based analysis of the texts was not part of our method. However, the overview of the texts that we obtained from the deductive analysis allows us to formulate the following observations: an anthropomorphic representation of the virus as a villain or as a positively cast hero of a story was particularly attractive for the learners. The temporal and spatial structure of protein biosynthesis and virus replication, on the other hand, was usually only described, even by the story authors, in terms of a technical process. The complicated processes in the cell therefore did not 'automatically' benefit from a narrative realization, possibly because the learners' prior knowledge was not sufficient in many cases. E.g., their less concrete idea of a ribosome did not inspire them to a narrative realization. Consequently, it would be too simplistic to assume that using narratives in the science classroom is a promising strategy to use narratives when they are structurally close to the topic. In addition to the structural fit of the topic to story grammar, the learners' prior knowledge of the specific topic is probably also very important for the success of a narrative strategy. Furthermore, Bruner ([2], p. 147) describes "contrast, confrontation and metacognition" as a didactic strategy when dealing with the narrative mode. Stories can therefore be particularly effective for learning where there are breaks and inconsistencies with the factual structure of the topic, or overly crude anthropomorphisms and moral judgments. However, these must then be reflected on together in class later on. It follows that the question of whether or not anthropomorphic and incorrect elements of learner stories can be specifically reflected on later in the lesson plays a role, may be as important as the structure of the topic itself. Narration is therefore not a didactic "all-purpose weapon" that would be equally conducive to learning for all scientific topics [22]. In addition to the mere writing or reading of a story, it is the tension between narrative and scientific thinking that should be used productively in the science classroom [2].

A narrative effect can only be examined with regard to the educational goal, the age group, the specific role of the stories in the teaching sequence and, of course, the respective scientific context. For example, our intervention was not designed to reflect the texts and thus emphasize the "contrast" described by Bruner ([2], p. 147). The stories written by our sample group primarily served to deepen and link prior knowledge. Our results are therefore only transferable to a limited extent to other scenarios in which narration is used. However, they at least provide indications that thematic fit is only one of many factors when it comes to the effectiveness of stories for learning. We need much more empirical evidence before we can draw general conclusions.

With this study design, we are responding to the findings of Soares et al. (2023), who were unable to find sufficient empirical evidence for a general narrative effect in their current review. Therefore, they suggest that future studies should be more nuanced; i.e., more specific about what the narrative intervention consists of and what its expected effect should be. We believe that our study and its results are in line with this suggestion. They can therefore contribute to a more differentiated research situation in which narration is no longer evaluated in a generalized way.

4.5. Limitations

This study is a naturalistic quasi-experimental intervention study in a school context, using a convenience sample group. Its most important limitation is a comparatively small (and unevenly distributed) sample. Furthermore, the test design is simple in that it relies only on the students' self-written texts and a post-test questionnaire. We did not perform a pre-test, a limitation which makes the comparison of the two groups less reliable. As they were natural classes and not randomized sample groups, we cannot assure that a pre-test would have shown homogeneous sub-samples with respect to previous knowledge in biology or other relevant variables. Also, our category system for scientific correctness does not directly reflect understanding or knowledge but is also sensitive to the language use of the students. For example, a text in a more colloquial writing style could have led to an underrating of their student's scientific understanding. Whether the extent of scientific understanding can be derived from a text analysis is debatable: if someone is able to reproduce specialist content in an explanation, this does not always mean that they have understood it well. Other factors such as linguistic fluency, perhaps creativity, etc., have a significant influence on the writing process.

Another relevant limitation is the reliability of the coding process. The two coders compared their results and eventually reached consensus, so that no intercoder reliability was calculated. In hindsight, this would have made the quality of the category systems and the coding process more transparent.

As to the explanation of our results, our study design does not provide enough data about the extent to which (narrative) text production enabled in-depth understanding. For example, we cannot distinguish here between the understanding that the course already achieved in advance and the understanding specifically fostered by our intervention. Process data from the teaching sequence, coupled with individual interviews, could answer the question of what the text production actually contributed to individual conceptual change processes. With the existing data set, we cannot discern whether the significance of the (narrative) writing process lies more in the linking of existing knowledge or whether new knowledge (in the sense of a superordinate context) was generated during writing.

Author Contributions: Conceptualization, J.Z. and C.A.; methodology, J.Z. and C.A; software, C.A.; validation, J.Z. and C.A; formal analysis, J.Z. and C.A.; investigation, J.Z.; resources, J.Z.; data curation, J.Z. and C.A.; writing—original draft preparation, J.Z. and C.A; writing—review and editing, J.Z. and C.A.; visualization, J.Z. and C.A.; supervision, J.Z. and C.A.; project administration, J.Z.; funding acquisition: J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Funded by the Open Access Publishing Fund of Leipzig University supported by the German Research Foundation within the program Open Access Publication Funding.

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

Data Availability Statement: The data presented in this study are available on request from the corresponding author (joerg.zabel@uni-leipzig.de).

Acknowledgments: The authors acknowledge Lili Genzel and Kenny Teucher for their extensive help with conducting the research and the investigation process, specifically performing the teaching unit and the data collection in the classroom. Furthermore, we would like to acknowledge Alexander Bergmann-Gering for his help and advice with our statistical analysis.

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

References

1. Norris, S.P.; Guilbert, S.M.; Smith, M.L.; Hakimelahi, S.; Phillips, L.M. A theoretical framework for narrative explanation in science. *Sci. Educ.* **2005**, *89*, 535–563. [CrossRef]
2. Bruner, J.S. *The Culture of Education*; Harvard University Press: Cambridge, MA, USA, 1996.
3. Bruner, J.S. *Actual Minds, Possible Worlds*; Harvard University Press: Cambridge, MA, USA, 1986.
4. Millar, R.; Osborne, J. *Beyond 2000, Science Education for the Future. The Report of a Seminar Series Funded by the Nuffield Foundation*; King's College London, School of Education: London, UK, 1998.
5. Rumelhart, D.E. Notes on a Schema for Stories. In *Representation and Understanding*; Elsevier: Amsterdam, The Netherlands, 1975; pp. 211–236. [CrossRef]
6. Mandler, J.M.; Johnson, N.S. Remembrance of Things Parsed: Story Structure and Recall. *Cogn. Psychol.* **1977**, *9*, 111–151. [CrossRef]
7. Stein, N.L.; Nezworski, T. The effects of organization and instructional set on story memory∗. *Discourse Process* **1978**, *1*, 177–193. [CrossRef]
8. Mandler, G. *Mind and Body: Psychology of Emotion and Stress*; W W Norton & Co Inc.: New York, NY, USA, 1984.
9. Gropengießer, H. Lernen und Lehren. In *Gehirn Und Lernen*; Siebert, H., Ed.; Bertelsmann: Bielefeld, Germany, 2003; pp. 29–39.
10. Lakoff, G. The Invariance Hypothesis: Is abstract reason based on image-schemas? *Cogn. Linguist.* **1990**, *1*, 39–74. [CrossRef]
11. Lakoff, G.; Johnson, M. *Metaphors We Live By*; University of Chicago Press: Chicago, IL, USA, 1989.
12. Zabel, J. Narrative and Metaphor as Conceptual Tools for Understanding Evolution Theory. In *Innovazione Nella Didattica Delle Scienze Nella Scuola Primaria e Dell'infanzia'*; Universitas Studiorum: Mantova, Italy, 2015; Volume 3.
13. Echterhoff, G.; Straub, J. Narrative Psychologie: Facetten eines Forschungsprogramms (Teil 1). *Handl. Kult. Interpret.* **2003**, *12*, 317–342.
14. Bruner, J.S. *Acts of Meaning: Four Lectures on Mind and Culture*; Harvard University Press: Cambridge, MA, USA, 1990; Volume 3.
15. Kurth, L.A.; Kidd, R.; Gardner, R.; Smith, E.L. Student use of narrative and paradigmatic forms of talk in elementary science conversations. *J. Res. Sci. Teach.* **2002**, *39*, 793–818. [CrossRef]
16. Bruner, J.S. *Search of Pedagogy Volume I*; Routledge: London, UK, 2006. [CrossRef]

17. Lhoste, Y.; Boiron, V.; Jaubert, M.; Orange, C.; Rebière, M. Le récit: Un outil pour prendre en compte le temps et l'espace et construire des savoirs en sciences ? *RDST* **2011**, *4*, 57–82. [CrossRef]
18. Orange, C.; Orange Ravachol, D. Problématisations scientifiques fonctionnalistes et historiques en éducation relative à l'environnement et au développement durable: Le cas de l'évolution climatique. *Rev. Des. Hautes Écoles Pédagogiques Inst. Assim. Suisse Rom. Du. Tessin* **2017**, *22*, 21–38.
19. Zabel, J. *Biologie Verstehen: Die Rolle der Narration beim Verstehen der Evolutionstheorie*. 2009. Available online: https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa2-796484 (accessed on 3 May 2024).
20. Zabel, J.; Gropengießer, H. What can Narrative contribute to Students' Understanding of Scientific Concepts, e.g., Evolution Theory? *J. Eur. Teach. Educ. Netw.* **2015**, *10*, 136–146. Available online: https://etenjournal.com/2020/02/07/what-can-narrative-contribute-to-students-understanding-of-scientific-concepts-e-g-evolution-theory/ (accessed on 3 May 2024).
21. Paulin, F.; Bruguière, C.; Zabel, J. La sélection naturelle racontée: étude de cas chez des lycéens. In *ACTES XIIe Rencontres Scientifiques ARDIST.*; Venturini, P., Pélissier, L., Eds.; ARDIST: Toulouse, France, 2022; pp. 97–102.
22. Soares, S.; Gonçalves, M.; Jerónimo, R.; Kolinsky, R. Narrating science: Can it benefit science learning, and how? A theoretical review. *J. Res. Sci. Teach.* **2023**, *60*, 2042–2075. [CrossRef]
23. Avraamidou, L.; Osborne, J. The Role of Narrative in Communicating Science. *Int. J. Sci. Educ.* **2009**, *31*, 1683–1707. [CrossRef]
24. Altun, D. From Story to Science: The Contribution of Reading Fiction and Hybrid Stories to Conceptual Change with Young Children. *Child. Soc.* **2019**, *33*, 453–470. [CrossRef]
25. Faria, C.; Guilherme, E.; Gaspar, R.; Boaventura, D. History of Science and Science Museums. *Sci. Educ.* **2015**, *24*, 983–1000. [CrossRef]
26. Tomas, L.; Ritchie, S.M.; Tones, M. Attitudinal impact of hybridized writing about a socioscientific issue. *J. Res. Sci. Teach.* **2011**, *48*, 878–900. [CrossRef]
27. Ritchie, S.M.; Tomas, L.; Tones, M. Writing Stories to Enhance Scientific Literacy. *Int. J. Sci. Educ.* **2011**, *33*, 685–707. [CrossRef]
28. Csikar, E.; Stefaniak, J.E. The Utility of Storytelling Strategies in the Biology Classroom. *Contemp. Educ. Technol.* **2018**, *9*, 42–60. [CrossRef] [PubMed]
29. Negrete, A.; Lartigue, C. The science of telling stories: Evaluating science communication via narratives (RIRC method). *J. Media Commun. Stud.* **2010**, *2*, 98–110.
30. Prins, R.; Avraamidou, L.; Goedhart, M. Tell me a Story: The use of narrative as a learning tool for natural selection. *EMI Educ. Media Int.* **2017**, *54*, 20–33. [CrossRef]
31. Duit, R.; Treagust, D.F.; Widodo, A. Teaching Science for Conceptual Change: Theory and Practise. In *International Handbook of Research on Conceptual Change*; Vosniadou, S., Ed.; Taylor & Francis: London, UK, 2013; pp. 487–503.
32. Stern, F.; Kampourakis, K. Teaching for genetics literacy in the post-genomic era. *Stud. Sci. Educ.* **2017**, *53*, 193–225. [CrossRef]
33. Jördens, J.; Asshoff, R.; Kullmann, H.; Hammann, M. Providing vertical coherence in explanations and promoting reasoning across levels of biological organization when teaching evolution. *Int. J. Sci. Educ.* **2016**, *38*, 960–992. [CrossRef]
34. Duncan, R.G.; Reiser, B.J. Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *J. Res. Sci. Teach.* **2007**, *44*, 938–959. [CrossRef]
35. Lewis, J.; Kattmann, U. Traits, genes, particles and information: Re-visiting students' understandings of genetics. *Int. J. Sci. Educ.* **2004**, *26*, 195–206. [CrossRef]
36. Haskel-Ittah, M.; Yarden, A. Toward Bridging the Mechanistic Gap Between Genes and Traits by Emphasizing the Role of Proteins in a Computational Environment. *Sci. Educ.* **2017**, *26*, 1143–1160. [CrossRef]
37. Gericke, N.; Wahlberg, S. Clusters of concepts in molecular genetics: A study of Swedish upper secondary science students understanding. *J. Biol. Educ.* **2013**, *47*, 73–83. [CrossRef]
38. Trauschke, M. Krieg im menschlichen Körper—Überfälle böser Viren. *MNU* **2008**, *61*, 493–499.
39. Simonneaux, L. A study of pupils' conceptions and reasoning in connection with "microbes", as a contribution to research in biotechnology education. *Int. J. Sci. Educ.* **2000**, *22*, 619–644. [CrossRef]
40. Rönner, A.C.; Jakobsson, A.; Gericke, N. "Bacteria are not viruses; viruses are more malicious"—Young pupils' understanding of bacteria and viruses in the aftermath of COVID-19. *J. Biol. Educ.* **2023**, 1–14. [CrossRef]
41. Staatsministerium für Kultus des Freistaates Sachsen, Ed.; Lehrplan Gymnasium Biologie; 2019. Available online: https://www.schulportal.sachsen.de/lplandb/index.php?lplanid=83&lplansc=aUFARA6HvFKFs2nxWtO3&token=f6612663b50335ac23f99c83f437d1c7 (accessed on 3 May 2024).
42. Zohar, A.; Barzilai, S. A review of research on metacognition in science education: Current and future directions. *Stud. Sci. Educ.* **2013**, *49*, 121–169. [CrossRef]
43. Martinez, M.; Scheffel, M. *Einführung in die Erzähltheorie*; Beck: München, Germany, 2003.
44. Schilcher, A.; Krauss, S.; Kirchhoff, P.; Katharina, A.-M.; Christina, E.; Michael, E.; Mario, F.; Lisa, G.; Maria, G.-P.; Eileen, G.; et al. FALKE: Experiences From Transdisciplinary Educational Research by Fourteen Disciplines. *Front. Educ.* **2021**, *5*, 1–14. [CrossRef]
45. Labov, W.; Waletzky, J. Erzählanalyse: Mündliche Versionen persönlicher Erfahrung. In *Literaturwissenschaft Und Linguistik. Eine Auswahl. Texte Zur Theorie Der Literaturwissenschaft*; Athenäum Fischer: Frankfurt/Main, Germany, 1973; Volume 2.
46. Zabel, J. Geschichten für das Lernen nutzen. In *Biologie-Methodik. Handbuch Für Die Sekundarstufe I Und II*; Spörhase-Eichmann, U., Ruppert, W., Eds.; Cornelsen Verlage: Berlin, Germany, 2016; Volume 3, pp. 63–67.
47. Levi, P.; Rosenthal, R. *The Periodic Table*; Penguin Books: London, UK, 2000.

48. Egan, K. *Primary Understanding: Education in Early Childhood*; Routledge: London, UK, 1988.
49. Hadzigeorgiou, Y. *Imaginative Science Education. The Central Role of Imagination in Science Education*; Springer: Berlin/Heidelberg, Germany, 2016.