Synthesis writing

Teaching high school students how to read, plan, draft, and revise

van Ockenburg, L.

Publication date
2022

Citation for published version (APA):

CHAPTER 4

A BRIEF INTERVENTION

Can a three-minute animation video promote the integration of source information in synthesis texts?[*]

Abstract. In secondary education, instruction in synthesis tasks is relatively scarce. Therefore, it is likely that students lack a well-developed task definition the first time they perform a synthesis task. The extent to which the task definition has been developed influences the quality of the resulting synthesis text. Therefore, we investigated whether it is possible to further develop students' task definition using a brief intervention, an animated three-minute video, that promotes integration of source information in synthesis texts. The study was conducted in four ninth-grade classes (n = 98), in an experimental design with random assignment to one of two conditions: viewing the video prior to the synthesis task or not. The results indicate that students in the experimental condition, who viewed the video, wrote better integrated synthesis texts, measured both analytically and holistically. It therefore seems that this brief intervention can lead to more integrated synthesis texts, based on a better developed task definition.

Keywords: writing education, synthesis texts, instructional design, design study, design principles

1   INTRODUCTION

Our current era is experiencing an explosive increase in the amount of available information. New technologies do not only make it possible to find more and more information, but also make it accessible to almost everyone. It is to be expected that both the amount of information and its accessibility will continue to expand and thus have an increasingly strong influence on all kinds of aspects of our lives (Sociaal Cultureel Planbureau [The Netherlands Institute for Social Research], 2004). Thus, people have to learn to find their way in this forest of information. Education plays an important role in this. At school, students should acquire information processing skills that are necessary, for example, to find relevant and reliable sources and then process them critically in their own texts. But there are serious concerns about precisely these skills. The 2018 PISA survey found that less than one in ten Dutch 15-year-olds could distinguish sufficiently

between fact and opinion, based on implicit clues regarding the content or source of the information (Gubbels et al., 2019). The reading skills of a large group of Dutch students are therefore currently insufficient. In addition, Dutch university teachers also experience the writing skills of students who enter higher education after pre-university education as insufficient (Bonset, 2010). Therefore, it is necessary to pay specific attention to the development of students' information processing skills. This can be done, among other things, by introducing hybrid tasks in secondary education, such as information synthesizing tasks, which require the integration of information from different sources. Producing written syntheses requires continuous alternation between reading and writing: exploratory reading and understanding the sources, selecting relevant information, arranging the selected information in a logical way, and integrating it into a new text (Klein & Boscolo, 2016). Thus, synthesis writing induces complex processes that call for recursivity and mediation (Vandermeulen et al., 2020b).

Students' information processing skills will remain insufficient if nothing changes in education (see for example Gubbels et al., 2019). It is therefore necessary to pay specific attention to the development of these skills (Koninklijke Nederlandse Akademie van Wetenschappen [Royal Netherlands Academy of Arts and Sciences], KNAW, 2012; Ontwikkelteam Digitale Geletterdheid [Development Team Digital Literacy], 2019). Learning how to perform hybrid tasks in education, in which reading and writing are practiced in conjunction, can certainly contribute to this (Graham et al., 2018).

An example of a hybrid task is composing synthesis texts. These are texts in which information from various sources is presented in an integrated manner. The synthesis task requires a combination of reading and writing skills: exploratory reading and understanding of the information, selecting relevant information, connecting, organizing, and integrating the selected information, and presenting this integration in a text that can be read independently. Synthesizing information from different sources as part of source-based writing is present in secondary education but teaching methods for learning how to synthesize are underexposed. Although relatively much attention is paid in education to the importance of a logical text structure, for example by learning to make a plan and to use connectives, there is currently too little attention for, for example, how to logically arrange and link source information. Students therefore do not have a well-developed task definition, and this influences the quality of their synthesis texts. For example, instead of synthesizing source information, they present the information as separate summaries per source, or in a series of superficially connected sentences taken quite literally from the sources (Boscolo, 1996).

Good synthesis texts, on the other hand, form a representative reflection of source information and simultaneously present this information in an integrated manner (Klein & Boscolo, 2016). Integration is seen as the core of the synthesis task (Barzilai et al., 2018). Writers must create an intertextual model based on the relationships between concepts in the source texts. An intertextual model is based on an overarching main idea. The structure of the model differs from the structure of the individual sources and is more than the sum of those individual sources (Klein & Boscolo, 2016). Intertextual integration, however, is a complex process that requires great cognitive effort (Mateos & Solé, 2009). Therefore, much research has been done on ways to teach students how to integrate (Barzilai et al., 2018).

Previous research has shown that students aged 14 to 15 have a much greater sense of intertextual integration than younger students, and that the greatest development in this area occurs between seventh and ninth grades (Spivey & King, 1989). That is why we have chosen to teach ninth-grade students how to write synthesis texts. To this end, we designed a learning unit (see Chapter 3). The focus of this unit is mainly on the three basic processes of synthesizing: a) selecting, b) integrating, and c) structuring source information (Spivey & King, 1989), because we want to give students the opportunity to first become proficient in these skills. After the basics have been mastered, important additional skills can be practiced, such as assessing the reliability of the sources, weighing conflicting source information and correctly citing the sources in one's own text.

Prior to carrying out a first synthesis task and participating in the unit, the students received a pre-instruction in the form of a short, animated video. This article focuses solely on this very first step prior to participating in the unit: the pre-instruction. We substantiate why we chose to provide a pre-instruction, describe its design process and measure to what extent a pre-instruction influences the integration of source information in synthesis texts.

2   THEORETICAL BACKGROUND

2.1 *The influence of task definition on task execution*

A Task definition is the mental image that someone has of the desired outcome and execution of a particular task. Task definitions can differ from the ideal task representation, especially if a task is new or complex (Flower et al., 1990). A task definition is not static, but is continuously revised based on experience and instruction, to reduce discrepancies between the current and ideal task definition. However, when we look at effective writing instruction, for example in the meta-

analysis of Koster et al. (2015), improving the task definition cannot be found as an individual instruction that has proven to be effective. Nevertheless, the task definition is an important part of the task process: a comprehensive task definition provides support for determining product and process goals.

In the MDTRACE model (Multiple Document Task-based Relevance Assessment and Content Extraction) developed by Rouet and Britt (2011), there is an important place for defining the task as well. This model describes how readers deliberately process and use various source texts. The first step is to construct a task definition and establish the objectives of the task. The way students interpret the task and set goals can influence all subsequent reading decisions and actions (Rouet et al., 2017). For example, the task definition, which varies from writer to writer, may include criteria that an acceptable text must meet.

In 1991, Wallace and Hayes showed that an instruction about the task definition can influence the quality of a text. In three experiments, Wallace and Hayes (1991) and Wallace et al. (1996) examined the relationship between revision and task definition. The reason for these experiments was that the researchers found that freshmen were mainly concerned with correcting local text problems, such as spelling and grammar, when asked to revise their text. They ignored global text problems, such as organization. Experienced writers, on the other hand, generally revised their texts both globally and locally. Bereiter and Scardamalia (1983) hypothesized that young writers ignore global text problems because they are not yet skilled enough to find and fix those problems. But Wallace and Hayes suspected this should not be the case with their students. They hypothesized that the difference between freshmen and seasoned writers had to do with how they defined the revision task. To test this hypothesis, they designed an eight-minute tutorial to explain to freshmen what experienced writers do when they revise. This explanation resulted in a significant improvement in the revision quality. This research suggests that the task definition can be developed with a relatively short instruction, in other words: with a mini-intervention.

Pre-university students in secondary education probably do not have a good task definition of synthesizing. Indeed, in a survey by Solé et al. (2013) among 15- to 16-year-olds, more than two-thirds indicated that they thought that a synthesis task involves summarizing each source text or the main ideas from the sources. When performing the task, these students therefore did not assume that they had to integrate the source information, despite the fact that they had been explicitly instructed to do so. Other studies such as Mateos & Solé, 2009, also showed that students confuse synthesizing with summarizing. That is why we gave the students a pre-instruction before the start of the synthesis writing intervention (see Chapter 3). By doing so, we wanted to increase the chance that

all students started from a valid task definition and generated 'on task' outcomes in the pre-test measurement.

2.2 *Animated videos as an instructional method*

Wallace and Hayes (1991) opted for a mini-intervention in which the differences between the revisions of experienced and less experienced writers were shown in eight minutes on an overhead projector. Today, students are no longer familiar with this method, but the intervention of Wallace and Hayes nevertheless formed the basis for the current study. Indeed, both interventions, that of Wallace and Hayes and the intervention investigated in the current study, are comparable in terms of a) the short duration of the intervention, b) the emphasis on observing a 'good example', c) the focus on task definition, and d) measurement of the effect. We are not aware of any other writing skills study in which all these things were combined.

Nowadays, students are accustomed to watching instructional videos instead of or in addition to written instructional texts. Since 'flipping the classroom' gained prominence (Mazur, 2009), more and more teachers are asking their students to explore new subject matter outside class so that class time can be used for higher-order assignments (Anderson & Bloom, 2001). Such teachers often make instructional videos in which they explain the new subject matter to their students. Some teachers do this so well that their videos are used by other teachers, or by students who search online for extra information in addition to their regular lessons. Moreover, many textbook publishers also offer instructional videos as a supplement to their regular teaching materials.

Instructional videos are a welcome addition to the usual written instruction, because we know from previous research (e.g., Solé et al., 2013) and from practical experience that students often do not read written instructions well. As a result, they often apply their own, incorrect, or incomplete task definition. We therefore expect a short instructional video to be a good addition to the written assignment.

There are many types of instructional videos. For this mini-intervention we chose to use an animated video. Moving objects play a role in animations and this makes them extremely suitable for efficiently representing dynamic systems, such as biological or technical processes. The way in which the verbal information is conveyed in the visualization can influence the learning outcome. Several studies confirmed a modality effect (e.g., Mayer et al., 2005), meaning that learning gains were greater when information was provided through two modes. For an optimal effect, the visualization should be linked to spoken text (auditory)

instead of written text (visual). The processing capacity of the working memory apparently increases when the cognitive load is distributed over two different channels, here an auditory and a visual channel (Mayer & Moreno, 2003).

Much research has been done on the effect of this type of conceptual understanding-oriented animation, but there is little empirical evidence regarding the expected learning outcomes (Berney & Bétrancourt, 2016). Nevertheless, we expect that instructional animations can contribute to the development of task definition in writing-reading education, provided they meet certain conditions. In every learning process, including learning by observing a video, four processes play a role that lead from observation to learning (Bandura, 1977). Students should first focus on what they have to learn and be able to follow and understand the offered subject matter well. Second, they must remember the most important information, third, they must be able to carry out what they have learned independently and finally, it is important that they should want to carry it out. De Grez et al. (2014) describe how these four processes, in combination with motivation, also play a major role in how instructional videos work. According to social cognitive theory, motivation arises from four variables: the self-efficacy that students experience, the value that students assign to the task, the extent to which the goals are clear to the students, and attribution. For example, students will be more likely to (continue to) focus their attention on the video and remember the information if they think the task is important and they believe they can (learn to) perform the task. Finally, based on earlier research, we can assume that the effect of animations can be enhanced by the addition of an evaluation component. Braaksma et al. (2001) analysed the activities of students who learned through observation while writing argumentative texts. They found that comparing, contrasting, and evaluating the observed is essential for the effectiveness of observational learning. Observational learning should therefore always contain an evaluation component (Rijlaarsdam et al., 2008).

2.3 *Research question*

Research seems to suggest that a well-developed task definition is often missing when students perform a synthesis task for the first time and that a short instruction could be sufficient to further develop such a task definition. Our research question is therefore:

> Can we promote the integration of source information in synthesis texts of students performing a synthesis task for the first time with a mini-intervention aimed at developing a task definition, namely a short, animated video?

3   METHOD

3.1 *Design and pilot of the video*

3.1.1   Design

We based the design of the instructional video on the learning-by-observation theory. Appendix E shows the complete scenario of the instructional video per scene. In scene 1, we first turned students' attention to what they had to learn, which was to perform synthesis tasks. Because this was a task they were unfamiliar with, we made it clear right from the start that synthesis tasks are source-based tasks. To increase motivation, we showed in scene 2 that writing synthesizing is a common practice in secondary and higher education and in many professions. In the following scenes we set two concrete goals as part of the task definition: all sources are represented in the synthesis text (scene 3) and the relevant source information is ordered per sub-topic (scene 4a). The video contains an evaluation activity at various moments. For example, in scene 4a we show what a good synthesis text might look like (in this case: well-integrated source information) and in scene 4b we show what is not the intention, but what often happens with more inexperienced writers (in this case: no integration, so three juxtaposed summaries). This part of the instruction has a structure similar to the intervention by Wallace and Hayes (1996): they showed the revision approach of an expert and a peer and clearly indicated the differences between the two approaches. By comparing both approaches, the students were able to picture the desired approach.

We made the first version of this video in an application for creating digital animations. We then showed the video to an expert group of writing researchers and teachers. They advised adding an extra element: "Use your brain: arrange the paragraphs in a logical way, use your own words and make sure that a reader who does not know the sources can follow the text well." (Scene 5). Finally, to help students even further, we added scene 6: an overview image which summarized the most important information.

3.1.2   First design and pilot

We first tested the effect of modality by creating different versions of the video (music only, voice only, voice + music). For practical reasons, we chose students from 10th-grade as the test group. A teacher from a school not involved in the main experiment taught four 10th-grade classes (N = 114) and had her students

assess the first design. Within those classes, one version of the video was ran-
domly assigned to each student. The animation was the same for each student,
but the auditory components were different. Group 1 (n = 37) watched a voice
only version, group 2 (n = 37) watched a music only version, and group 3 (n =
40) a version with both voice and background music. At the end of the video,
the students wrote down the most important information. Figure 4.1 shows that
97.3% of the students indicated that they had learned to combine several
sources into subtopics. 77.2% indicated that they had learned that a synthesis
text is a text based on sources and 43% that you must use your own words in a
synthesis text. Figure 4.1 shows the distribution of the comments across the dif-
ferent groups as well. The groups seem to focus on different information.
Whereas in group 1 (voice only) the focus was mainly on 'a synthesis text is a text
based on sources' and in group 2 (music only) on 'combining several sources
per paragraph', group 3 (voice and music) showed a more diffused picture.
Group 3 therefore seems to have focused less on one specific item than the
other groups.

*Figure 4.1 Overview of the distribution of the learning points mentioned (N =
342) in percentages over the different groups*
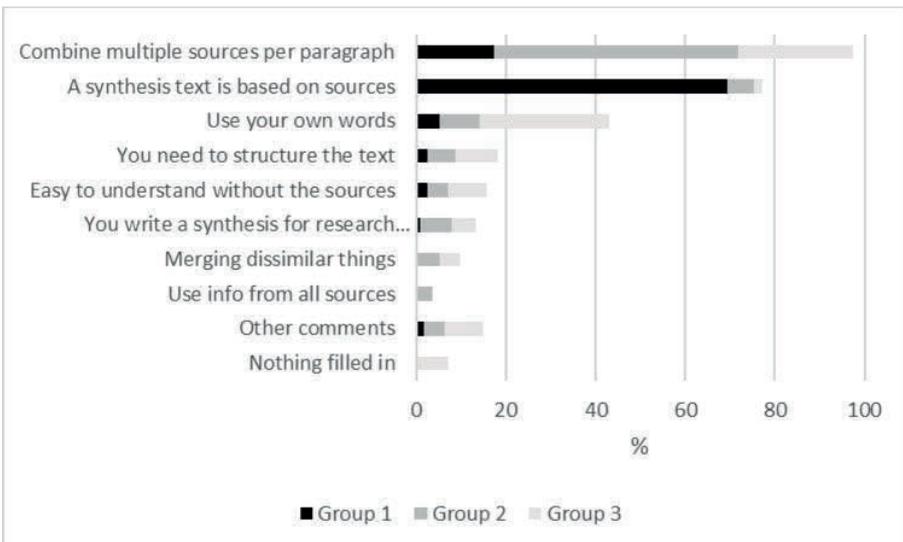
In addition, the students gave the developers of the video a suggestion for im-
provement (tip) and a compliment (top). Figure 4.2 shows the distribution of
these tips over the different groups. 25.3% of all tips concerned voice use. Of

course, none of the group 2 (music only) students made any comments about vocalization. The students from group 1 (voice only) found the voice 'dull' and 'monotonous' and the students from group 3 (music + voice) found the voice less intelligible due to the background music. At the same time, 12.5% of comments, all from Group 2, related to missing a voice. Furthermore, 22.5% of the tips were about making the video less elaborate or repetitive. At the same time, a small proportion of students (7.2%) thought the video went too fast. That seems contradictory, but the comments about the pace of the video were almost all from group 2 (music only). These students had to read along with the text that was visible in the animation. They indicated that it took them a lot of effort to read and at the same time process the content. Additionally, they had insufficient time to view the animations properly, which made them feel rushed. Overall, 14.8% of the comments were about distraction caused by the background music or superfluous animations in the video and 4.2% of the comments were related to the sound quality: the microphone we used made the voice sound hollow.

Figure 4.2 Distribution of the tips among the different groups.
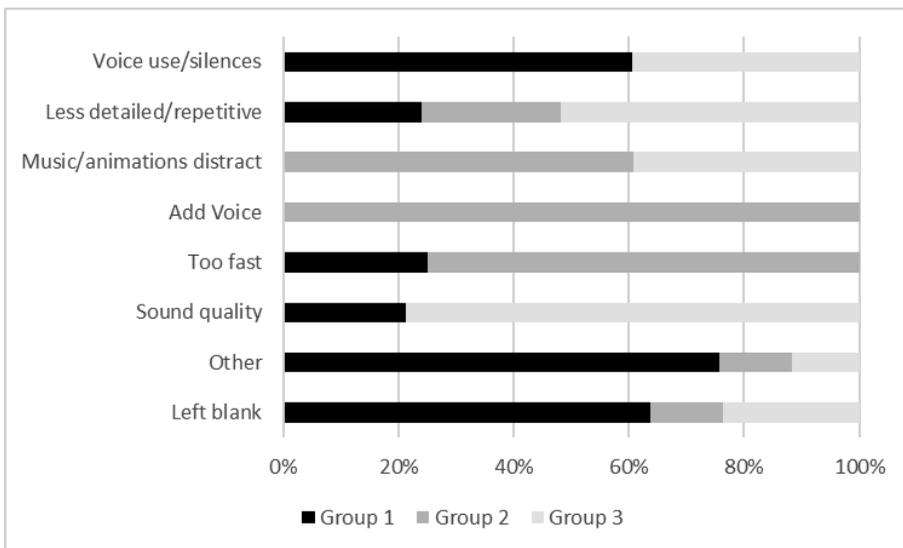
The compliments (tops) fell into three categories. 58.2% of the students found the explanation in the video clear. 32.5% liked the form: they liked the animations and/or thought the video was funny. The third category consisted of positive comments from some students (5.2%) about the background music.

3.1.3   Redesign and second pilot

At her own school, the teacher-researcher discussed these results with 10th-grade students during two extracurricular sessions. The students viewed the version of the video with voice and background music, confirmed that there was too much redundancy in the animation, and pointed out the parts which could be deleted. Subsequently, the video was shortened by 12 seconds (6%), to 3:02 minutes. We also removed some animated figures from the video that were not necessary for a proper understanding of the content. We chose quieter background music, played at a lower volume and re-recorded the voice with a better microphone.

We then pretested the redesign of the animation video, including voice and background music, in two ninth-grade classes from two schools that had not yet been involved in this study. We used this pretest to develop an instrument to measure the dependent variable, integration of source information. The students were randomly assigned to one of two conditions: performing a synthesis task after watching the video beforehand (test group) or without watching the video beforehand (control group). Table 4.1 shows the distribution of the students across the conditions.

Table 4.1. Distribution of students over conditions and average number of links

	School A	School B	Total	Links
				M (*SD*)
Test	5	12	17	4.71 (1.7)
Control	11	12	23	3.09 (1.4)

The test group received the instruction on paper, watched the video individually on a laptop and then started the synthesis task. The control group received the instruction on paper and immediately started the synthesis task. Both groups were given the same assignment and the same three sources about the use of mobile phones in traffic. At the end of the 50-minute lesson, the students submitted their text in digital form to their teacher.

This test yielded 40 texts. We counted the number of links in each text to get an idea of the possible effect of the video on integration. Links are places in the text where the content changes from information from one source to information from another source, or from the author's own text to strongly source-

based text (and vice versa). We describe this analysis method, which formed the basis for the analysis of the texts in the main study, in more detail in section 3.2.3.

Table 4.1 shows that the synthesis texts of students in the test group contained significantly more links on average than those of students in the control group $t(38) = 3.22$, $p = .003$, $d = 1.05$. Thus, we deduced that the video could have a positive effect on the level of integration in synthesis texts and thus possibly also on the quality of these texts. This was sufficient reason for us to want to investigate the effect of the animated pre-instruction on a slightly larger scale, in the target group, Grade 9. Therefore, after this second pilot, the video was used in the main study without any changes.

3.2 Main study

3.2.1   Participants

To test the effect of the video more extensively, we conducted a survey among 98 ninth-grade students (59.2% girls) from four different pre-university classes from one school. Students were between 12 and 15 years old. The research took place in a lesson in the first week of the school year. The research design was experimental with random assignment of participants to one of two conditions: with or without viewing the instructional video prior to starting with the synthesis task. The experimental group consisted of 48 students (41.7% girls), the control group of 50 students (76% girls). The distribution between boys and girls appeared to differ between the conditions ($\chi^2(1) = 11.95$, $p = .001$). We will come back to that when we discuss the results.

Prior to the lesson in which the experiment took place, the students were informed that they would participate in a study. They were told no more about the study than that it was about writing skills and that all texts were equally important for the study. The parents of one student did not give permission for participation. This student did write a text during the research lesson, but all their data were deleted afterwards.

3.2.2   Materials and procedure

Materials
All students were given the same assignment: they had to compose a synthesis text of 180-220 words based on three sources (number of words: $M = 193$; $SD = 34$) about the human-wildlife conflict in Africa. This task was a simplified version of a synthesis task that was developed for a national baseline study among

students in Grades 10, 11 and 12 (Vandermeulen et al., 2020a). The students received both the assignment and the sources on paper but wrote the synthesis text on the computer. All information the students needed to successfully complete the assignment was provided (See Appendix F).

Procedure

At the start of the lesson, the teacher asked half of the students waiting in the hallway to enter the classroom. The other half were asked to wait outside the classroom for a few minutes until the teacher signalled to enter the classroom. This group of students formed the control group. When the experimental group was seated, the teacher started the instruction video without further explanation and after watching it, there was no class discussion or discussion of the content. The chance of possible teacher effects was thus reduced to a minimum, while we were still able to investigate the effect of the instruction video in a natural school environment. After the video had ended, the control group was also invited to take a seat in the room. All students were given the assignment and sources on paper. They had about 45 minutes left to complete the assignment.

3.2.3   Text coding and assessing

The students' digital texts were anonymized and randomized. They were then assessed in two ways: analytically and holistically. In the analytical assessments, we assumed that integration of source information in a synthesis text can be measured by looking at the representation of the different sources. If a writer switches very little between information from different source texts, the synthesis text will probably consist of 'juxtaposed' summaries that are not or hardly integrated. However, when a writer switches a lot between information from different source texts, this can be an indication of a high degree of integration. That is why the teacher-researcher and a second, independently working coder indicated all links in the synthesis texts. Links are places in the text where the author switches content-wise from one source to another, i.e., between different sources, or from own text (with the author himself as source) to source information (and vice versa). If contiguous sentences in the text were clearly based on one particular source text and thus formed a thematic unit, the coders treated it as a unit. If a unit was followed by text based on a different source, this was indicated as a new unit. Units can be formed not only by information from one source text, but also by the author's own text. By own text we mean, in addition to overarching introductions and concluding final paragraphs, also sentences with meta-communicative, structure-indicating words, such as "There are various reasons for this, (such as)" and transition sentences, such as: "There is yet

another/second/third reason to kill/poach/kill (wild/endangered) animals)." We call the transition between two units a link.

Figure 4.3 shows the difference between a text with few links and a text with many links. We used different shades of grey to visualize the units. The shades refer to the different sources (see Appendix G).

Figure 4.3 Two sample texts;
relatively few links on the left, relatively many on the right

L*	Text with two links	L*	Text with eight links
1 2	Poachers kill wild animals in Africa, for example, to sell certain body parts. The poachers are not alone, but are in crime networks and thus have weapons and other necessities. A lot is sold, especially in Asia, for a lot of money, ivory is very popular. There is Trophy Hunting as well, a legal hunt for certain animals, remains are then taken back, such trips are offered in African countries, the government sells hunting permits, but nothing happens with that money. Nothing changed after the death of the lion Cecil, but many famous people took action. They want the US, where you can take the remains on return, to label the animals as officially endangered. On the edge of Masai Mara is the territory of the lions of the March Pride group, where cows of the Masai family graze at night. The lions killed those bows and of course the farmers were angry. They poisoned the lions and 3 died as a result. There are a few easy solutions for the dead cows, they could let the cows graze during the day so that the temptation to attack for the lions is less.	1 2 3 4 5 6 7 8	More and more animal species are threatened with extinction, and this has been known for a long time. But if it has been such a big problem for a long time, why is nothing being done about it? Well, that's complicated. There are several reasons why animals face extinction, but one of the biggest is poaching. These poachers, heavily armed and working together in large, well-organized gangs, slaughter endangered species. This may be one of the biggest factors, but not the only one. Another factor why these animal species are so endangered is the local population. Lions, for example, eat the livestock of the population, of course the population does not want that, so they want to get the lions out of the way. Another factor is trophy hunting. These are hunting trips offered by travel companies. But why do these people do that? Well, there's a big business behind the poaching. There are several body parts that are a lot of money. Of course there are a lot of rangers who try to prevent it, but the poachers are well prepared for them, which makes it difficult. It is also difficult for the local population to prevent it, mainly because the population does not want to listen. With trophy hunting it is perhaps strange that nothing is being done about it yet. People go trophy hunting purely for fun. This is often allowed by corrupt officials. There are still campaigns against trophy hunting
	* L = link; source a; source b; source c; writer's own text		

The two coders first analysed ten texts based on a protocol in which the proposed coding scheme was elaborated with examples (see Appendix H). After each independent coding of a text, the segmentation in units was compared and the protocol was supplemented with additional instructions if there were still

ambiguities. After coding ten texts together, the coders independently coded the other 88 texts, which were available in two sets, each in a random order.

For the assessment of holistic text quality, we asked fifteen experienced raters, ten teachers of Dutch and five writing researchers, to compare the texts in pairs in Comproved, a digital platform (https://comproved.com/) in which the rater always sees two randomly chosen texts side by side and indicates which of the two texts is the best. Research has shown that this judgement method is as reliable as an analytical assessment using criteria lists, with a comparable time investment by raters (Coertjens et al., 2017).

When judging comparatively, raters judge the text as a whole. Of course, they may be provided with a specific focus, which can vary from judging their overall holistic impression of the text to a specific facet, or 'primary trait': a characteristic property given the function of the text (Lloyd-Jones, 1977). For this study, we chose integration as a primary trait aspect. We tried to draw raters' attention to this aspect in three ways. First, in the written instruction for raters, we identified integration as an important characteristic of good synthesis texts, we explained what we meant by integration, and we asked the evaluators to watch the instructional video that the students had also watched before the writing assignment. Second, we included two sample synthesis texts on a different topic in the instruction, with varying degrees of integration. These texts were accompanied by a short explanation of their integration quality. In the sample texts, integration was made visible by means of colour, similar to the units in texts from the analytical assessment (see Figure 4.3). Finally, we made use of a facility in the *Comproved* application that required each rater to answer a question after the comparison about the degree of integration in the rated texts. The answer to this question could be short because the question mainly served to ensure that the raters would always include this aspect in their judgment. In *Comproved*, the raters found information on other aspects that determine the quality of synthesis texts as well, namely: a) To what extent is all relevant information included in the synthesis text? b) To what extent is the text well-structured? and c) To what extent is the style of the text attractive and are spelling and grammar correct?

3.2.4   Data analysis

For the analytical assessment, we counted the number of links in each text. The correlation between the ratings of the two coders was high (ICC .82; p < .001). Where the number of links counted differed, the coders discussed the difference and scores were adjusted in consultation. This led to a definitive determination of the number of links for each text.

For the holistic assessment, it was pre-calculated that each rater had to make 50 comparisons to create a ranking with a minimum Scale Separation Reliability between .70 and .80. This Scale Separation Reliability (SSR) is used because pairwise ratings are analysed using Rasch models. High SSR values indicate a smaller measurement error and a higher reliability of the final scale. Bouwer et al. (2017) showed that SSR can be used as a measure of inter-rater reliability as well. The reliability of the ratings in *Comproved* was .73 after completing all 800 comparisons. Based on the ranking, all texts were assigned a logit score and a confidence interval.

The number of links in synthesis texts appeared to correlate with the logit score of the text (r .54, p < .001). This indicates a common construct of integration, but not in such a way that the two scores are not separate operationalizations. Therefore, we kept both scores in the analyses as operationalizations of source information integration. The effects of the intervention were then determined via multivariate analyses with condition as factor and both scores as dependent variables. We checked which factors could possibly influence the effects, by looking first at the distribution of participants over the conditions. After all, it had already been shown that the two groups differed significantly from each other in the ratio between boys and girls. Therefore, we controlled for the effect of gender, but also for class and final grade of Dutch from the previous year in the same multivariate analyses. If the multivariate analysis was found to be statistically significant, we reported the univariate analyses for each of the two integration scores. We used SPSS version 24 to analyse the data.

4   RESULTS

4.1 *Multivariate analyses*

A multivariate analysis of variance showed that condition contributed significantly to the integration scores (Pillai's Trace = 0.083; $F(2, 93)$ = 4.193, p = .018, η^2 = .083). Multivariate analyses including class or gender showed no main effects (Pillai's Trace = 0.044; $F(6, 176)$ = .662, p = .681, η^2 = .022; Pillai's Trace = 0.010; $F(2, 91)$ = .662, p = .619, η^2 = .010) or interactions with condition (Pillai's Trace = 0.036, respectively; $F(6, 176)$ = .531, p = .784, η^2 = .018; Pillai's Trace = 0.000; $F(2, 91)$ = .003, p = .997, η^2 = .000). Furthermore, students' final grade for Dutch in eighth grade (covariate) does not contribute to the explanation of the scores (Pillai's Trace = 0.028; $F(2, 92)$ = 1.309, p = .275, η^2 = .028).

4.2 *Univariate analyses*

Analytical. The experimental group's synthesis texts contained more links (M = 5.22, SD = 1.7) than those from the control group (M = 4.20, SD = 1.7). This difference was significant (F(1, 94) = 8.04, p = .006, η^2 = .079).

Holistic. Scores for the experimental group's synthesis texts were higher (M = 24.17, SD = 131.92) than those of students in the control group (M = -32.44, SD = 138.90). This difference was significant (F(1, 94) = 4.18, p = .044, η^2 .043).

5   DISCUSSION, RECOMMENDATIONS AND CONCLUSION

5.1 *Discussion*

To what extent does a mini-intervention aimed at task definition promotes the integration of source information in synthesis texts? To answer this question, we conducted an experiment in which we had 98 ninth-grade pre-university students write a synthesis text. Half were shown a three-minute animated video beforehand, the other half were not shown that video. We then presented all students with the same writing assignment.
Every study has its limitations. In this study, some limitations stem from the way we analysed the data and others were related to the study's design. Each of these will be discussed below.

5.1.1   Data analysis

Effectiveness
The results indicate that the experimental group subsequently wrote better integrated texts than the control group, measured both analytically and holistically. How can a three-minute instructional video result in this effect? Part of the effect may be explained by previous research (e.g., Solé et al., 2013), which shows that students do not always read written instruction well and therefore start from an incomplete task definition. When students are shown a video in addition to their usual written instruction, this can make the task definition completer and more comparable for all students. The effectiveness of the video may also be due to the way in which this intervention fits in with proven effective writing didactics. Constructing a task definition comes close to what Koster et al. (2015) refer to as goal setting, the intervention with the largest proven effect size in their study.

What distinguishes the mini-intervention from goal setting, is that the instructional video does not contain any explanation of how students can achieve preset goals. Strategy instruction (Koster et al., 2015) is another proven effective intervention. Observational learning is often a first stage in strategy instruction. Comparing and contrasting is an essential cognitive learning activity within observational learning (Rijlaarsdam et al, 2008). In strategy education, the observational learning phase is often followed by a phase of explicit instruction. In the instructional video in this study, no strategies are offered, but the elements of observational learning and comparing/contrasting may have contributed to the effectiveness.

Assessment
Although both assessment methods showed significant effects of condition, we found clear differences between the analytical and the holistic assessment of the texts. During the analytical assessment of integration, the number of links was counted: the number of places in the text where the author changes their focus on content from one unit to another. Analyses showed significantly more integration of source information in synthesis texts of students in the experimental condition when assessed analytically. Texts in the experimental group contained an average of 5.22 links between information from different sources compared to 4.20 links on average for texts in the control group. The effect size was η^2 = .079, which means that the intervention had a medium to large effect on the degree of integration expressed in the number of links. Holistic text quality was also significantly higher in the experimental group, but this effect was smaller with a η^2 of .043, indicating a small to moderate effect. This difference in effect size may have arisen because assessing the texts holistically proved challenging, even for a group of experienced raters. We had calculated beforehand that 800 comparisons of text pairs would yield a reliability between .70 and .80. That reliability eventually came to .73. This is an acceptable, but not very high reliability. This may indicate, for example, that the raters did not base their assessments on the same quality aspects and thus that their assessments were less comparable than we expected. Alternatively, this could indicate that the raters found the quality of the integration of source information difficult to judge, even though they were focused on integration by the recurring question in Comproved.

We also found that the analytical assessments of the texts (the number of links in the synthesis text) and the holistic assessments were related (r .54, p < .001), but represented different qualities. On the one hand, there was consensus among raters about what integration in synthesis texts means, namely the merging of information from different sources, but on the other hand, integration in

the analytic and holistic approach was understood differently. We think that in the holistic assessment, the *quality* of the integration was probably assessed, while quality played no role in the analytical assessment. The coders only counted the number of links but were not asked to include the *quality* of the links in their assessment. In addition, aspects other than integration alone also play a role in a holistic assessment of text quality, namely the amount of relevant source information, the structure of the synthesis text and style/language use. The analytical and holistic assessment were thus complementary in this regard.

5.1.2   Research design

Prior to the intervention, the students knew that they would participate in a study, but they had received little background information; they only knew they had to write a text. Students from the four participating classes were immediately randomly assigned to a control and experimental group upon arrival at their classroom. The teacher let about half of the students into the room and asked them to sit down. The students in the control group had to wait outside the room without any further explanation. The teacher then closed the door of the room, which was situated in such a way that the control group could not hear or see anything that was going on in the room. In the classroom, the teacher showed the animation video to the experimental group, without further explanation.

We separated the groups in this way, as this was the most effective method to create an experimental setting in a normal school situation. By having the control groups wait in the hallway, the intervention took only five minutes (three minutes of animation video and about two minutes for the control group to come in and pick up a laptop), and all students could then spend the same time under the exact same conditions to work on the synthesis task. We found that these benefits outweighed potential drawbacks, such as potential interpretative difficulties associated with our decision to have the control group wait in the hallway. On the one hand, difficulties can arise from dividing a class into two groups. The results show, for example, that the distribution of boys and girls within the groups was not balanced. This may be due to the way the students were let into the room. The students waiting closest to the door formed the experimental group because they were admitted first and probably more boys were in the front than girls. The effect of the conditions could therefore be masked by the gender effect. However, gender did not appear to interfere with the condition effects.

Another potential threat to validity may be that the experimental group felt favoured. After all, these students received something extra compared to the control group that had to wait in the hallway. And even though the control group did not know what was going on in the classroom and were not told so later by the teacher, they may have felt slightly disadvantaged. These limitations should therefore be taken into account when interpreting the study's results.

5.2 *Recommendations for education and future research*

Despite the limitations we discussed, it is encouraging to see that pre-instruction on a new, unfamiliar task can be effective. In this study, a mini-intervention consisting of a video instruction of only a few minutes enabled students to write more integrated synthesis texts. It is plausible that these kinds of mini-interventions can also yield learning benefits for other writing skills; Wallace et al. (1996) already showed in a comparable study that a certain group of students revise better after a mini-intervention. Follow-up research could clarify the possible other forms that effective pre-instruction can take and its usefulness for other areas within (writing skills) education.

Although there is little empirical evidence that animated videos can be effective as a form of instruction, the results reported here seem to indicate that they can certainly contribute to effective (pre-)instruction. We believe that animated instructional videos can benefit a wide range of learners if their design is rooted in learning-by-observation theory, motivates learners, and includes components that require evaluation. In addition, to increase the learning outcome the modality effect must be taken into account when operationalizing the design, which ensures that the processing capacity of the working memory increases when the information is presented both auditory and visually. Comments from the pilots also indicated that many students find the combination of a visual and auditory approach more pleasant than a purely visual approach. Follow-up research could provide more clarity about the extent to which the various elements mentioned above contribute to the effectiveness of animations as a form of instruction.

However, such an intervention will only work under certain circumstances. On the one hand, students' task definition must still be under construction, otherwise the mini-intervention will probably have no effect. On the other hand, the students must already possess the skills that the intervention requires them to carry out. In other words, they must already be able to do it, to a certain extent. What we do not know is how long the effect of such a mini-intervention will last. If it brings about a change in the task definition, we can expect a lasting effect. In addition, the effect might be maintained for a longer period by involving

students in a class discussion after viewing it or by reminding them of crucial points from the video during the lessons in which they start carrying out the instruction.

5.3 *Conclusion*

Mini-interventions in writing education seem to be effective and easy to implement. Their use is in line with a trend to occasionally use (animated) videos as a replacement for or supplement to explicit instruction.