Literacy in vocational education and training: scenario-based reading and writing education

Liana Konstantinidou[1] · Karin Madlener-Charpentier[1] · Aleksandra Opacic[1] · Curtis Gautschi[1] · Joachim Hoefele[1]

Abstract

Many students in vocational education and training (VET) have difficulties with reading and writing. To date, there is little research on whether and how the development of VET students' writing skills may benefit from teaching approaches that integrate reading and writing. This study reports results from a semester-long intervention study conducted in Switzerland in 2018/19 (N = 285) in which we investigated the impact of a scenario-based integration of reading-to-write-tasks on the development of VET students' text quality. In the approach, problem-solving processes are set in motion by scenarios representing real- or work-life situations. Reading-to-write tasks form part of the student-initiated problem-solving process, and result in situated argumentative writing. A small experimental intervention effect was found where text quality developed significantly better in the experimental group than in a matched control group ($F_{1,178} = 7.40$, $p < .01$, Cohen's $f = 0.20$), as measured in a writing test before and after exposure to the teaching method. Outcomes suggest tangible benefits may result when applying this approach to literacy education in VET, particularly for academically weaker students. We discuss the consequences and implications of these findings, as well as open questions to be addressed by further research.

Keywords Reading and writing scenarios · Vocational education and training · Literacy education · Intervention study · Reading-to-write-tasks · Text quality

Introduction

The vocational education and training (VET) system is the most widespread form of upper-secondary education in Switzerland, providing high quality professional training to prepare young people for integration into the labor market (SERI, 2022). The

✉ Liana Konstantinidou
  liana.konstantinidou@zhaw.ch

Extended author information available on the last page of the article

so-called *dual system* in VET combines practical vocational training in a company with in-school learning. However, in VET schools, many students have difficulties with reading and writing, which may have an impact on their professional career and overall participation in society (Efing, 2008; Baumann, 2014; Conrad et al., 2018; Milke et al., 2013).[1] VET students in Switzerland tend to have heterogeneous language backgrounds and diverse language abilities. Many grow up in multilingual environments, using several languages simultaneously, or have learned German as an additional language later in life (Konstantinidou et al., 2016). Furthermore, students in the German-speaking part of the country are faced with a diglossic situation with significant differences, in terms of lexis, grammar, and idioms, between written German and the more widely spoken local dialects (Ammon et al., 2004; Dürscheid et al., 2011).

Despite the linguistically diverse backgrounds of students, literacy teaching in VET schools still follows an approach for homogenous L1-student groups (for a more general criticism of the *monolingual habitus* in education, see Gogolin 2008) and VET-teachers have thus little practical experience coping with the plurilingualism that is common in their classrooms (Krekeler, 2002; Hoefele & Konstantinidou, 2018).

Given the importance of writing and reading skills to successful societal and workplace integration, the difficulties met by VET students, and the challenges faced by VET teachers, there is a need to develop teaching approaches for reading and writing education which reflect and are suited to the plurilingual profiles of VET students. In turn, it is necessary to examine the effectiveness of such approaches. To date, little is known about whether and how the development of VET students' writing skills may benefit from teaching approaches that integrate reading and writing, in what we here refer to as *Scenario-based reading and writing education*. To address this gap, the present study examines the impact of this approach on text quality.

Scenario-based reading and writing education

Scenario-based reading and writing education meets the specific needs of students in vocational education, as scenarios (a) allow reading and writing tasks to be situated in various social areas or professional domains (OECD, 2016), (b) link topics of general or professional interest with language education as required by the curriculum framework and school curricula (BBT, 2006), and (c) result in communicative tasks that can be solved by learners with different levels of proficiency. It has

---

[1] 2018 PISA data show that 47% of lower secondary students (level 2 according to UNESCO ISCED, 2011) have difficulty identifying, understanding, evaluating, and reflecting on key messages in texts (PISA 2018), and thus do not reach a basic proficiency level in reading comprehension. Although there is no statistical data specifically on Swiss VET students' literacy skills, it is likely that many have considerable difficulty in reading and writing given that two-thirds of Swiss lower secondary education students (SERI, 2022, p. 11) attend vocational school (level 3 according to UNESCO ISCED, 2011).

been successfully implemented in prior vocational education research (Hoefele & Konstantinidou, 2018; Hoefele et al., 2018).

A scenario is a description of a real- or work-life situation that is likely to occur (OECD, 2016; Piepho, 2003). Scenarios can be developed for a wide range of possible situations (OECD, 2016). Even though scenarios do not explicitly call for reading and writing activities, both reading and writing tasks are often an integrated part of the problem-solving process (BBT, 2006; Council of Europe 2001) set in motion by the scenario.

To successfully manage the problem-solving processes emerging from a scenario, students require literacy competence in a broad sense. Literacy involves knowledge, skills, and attitudes (European Commission, 2018). In terms of knowledge, it includes knowledge of reading and writing as processes and knowledge of the linguistic means of expression, from vocabulary and functional grammar to genre awareness. In terms of skills, literacy covers (a) effective oral and written communication in different situations and domains, (b) finding and evaluating resources and information, and (c) formulating and expressing "one's oral and written arguments in a convincing way appropriate to the context" (European Commission, 2018, p. 2). Attitudes refer to positive interactional dispositions and to the "awareness of the impact of language on others and a need to understand and use language in a positive and socially responsible manner" (European Commission, 2018, p. 2).

Through the problem-solving processes that emerge from scenarios, reading tasks are primarily used for goal-oriented action (OECD, 2016). Goal-oriented 'functional reading' is thereby focused within the scenario approach. Functional reading refers to reading in everyday and professional contexts with an intention to act (Ziegler et al., 2012). Therefore, students do not refer to entire texts as they do, for example, in an interpretation, summary, or commentary task (Feilke et al., 2016; Sturm, 2017). Rather, they use texts to gather selective information (OECD, 2016; Ziegler et al., 2012). To do this, they must scan texts to locate, identify, and understand relevant information (Ziegler et al., 2012). This means readers must evaluate whether the information gathered from the text supports the intention to act: "To be able to convert written information into action, its integration into a mental model is necessary, i.e., the reader must develop a concrete idea of what is to be done." (Ziegler et al., 2012, p. 5). In this process, information is reconceptualized, restructured, and linguistically transformed to a writing product which should inform and convince the reader, in order to achieve a personal goal (functional writing).

Reading-to-write

This reciprocal interaction between literacy skills, which involves the interplay between reading and writing processes, describes, according to Ascención Delaney (2008), the reading-to-write construct. In scenario-based reading and writing education, reading-to-write involves reading to integrate information for the purpose of producing a written text. Readers choose information from the source text, evaluate it and use it in their own writing (Ascéncion Delaney, 2008). At the same time, reading-to-write may involve reading to integrate linguistic resources

wherein readers analyze the texts of more skilled writers and use their linguistic resources (e.g., vocabulary, text structure, genre formal conventions) in their writing.

Reading-to-write tasks can also be considered as part of a scaffolding concept for the development of writing skills. For example, in a previous study in VET context, reading tasks were integrated as reading instructions, as well as the scaffolding of language skills was implemented and integrated into the writing process (Hoefele & Konstantinidou, 2016; Hoefele et al., 2017). This was based on the principle that more skilled language users can, for example, provide, through their texts, a scaffold that helps novice writers to handle more complex tasks (Applebee & Langer, 1983). Positive effects on writing were found in the experimental group that implemented the method.

The ability to perform reading-to-write not only requires basic reading skills, but also the ability to regulate reading and writing processes and to apply effective strategies. VET students, in general, tend to lack strategies and routines when dealing with reading texts (Roche, 2017), requiring teachers to provide reading instruction. Graham et al., (2018) examined in a meta-analysis whether reading instruction leads to improvements in students' writing performance. The analysis showed that in 19 of 20 studies, preschool to high school students improve their overall writing performance, when they were taught how to read. The reading instruction consisted of phonological awareness, phonics and reading comprehension instruction. 12 of the studies measured the impact on text quality. The average weighted estimated effect in these studies (0.63) was found to be statistically significant (Graham et al., 2018). On these grounds, it may be assumed that integrating reading instruction into writing instruction will also improve the text quality of the VET students.

Previous studies on reading-to-write show that performance also depends on individual variables. Literacy expertise, language proficiency and educational level affect the way students organize and connect information from source texts when using them in their own writing (Ascención Delaney, 2008; Kennedy, 1985) found that higher reading ability has an influence on how many notes college students take and how elaborate notetaking is. When undergraduate college students are better readers, they also produce better texts according to a study conducted by Risemberg (1996). This is also the case for lower education levels. Better readers at the sixth, eighth, and tenth grade levels have been found to use more ideas from source texts in their essays and show a better understanding of content (Spivey & King, 1989).

In the L2-context, Watanabe (2001) and Plakans (2009) found a similar effect of reading comprehension ability on integrated writing tasks. Plakans (2009) also examined the role of reading strategies in L2-writing tasks. The results reveal that higher scoring writers are higher scoring readers. They also use global strategies, such as goalsetting before dealing with texts, and mining strategies, such as looking purposely for useful information in texts. However, Ascéncion Delaney (2008) found only a weak relationship between reading ability and performance on reading-to-write tasks, based on low but significant correlations between a measure of reading ability and scores on two types of integrated reading-to-write tasks.

Reading and writing abilities have both been found to depend on overall language proficiency, as part of a multifaceted general language proficiency construct

(see Harsch 2014). Less advanced L2 readers/writers generally lack the linguistic resources (vocabulary, grammar, syntax) necessary for managing reading and writing processes. Corbeil (2000), exploring the effects of language proficiency on summary writing in French as a second language, found that both the ability to summarize in the first language and second language proficiency have an impact on the quality of the second language summaries. Similar effects of second language proficiency on summary writing were also found in previous studies of Johns & Mayes (1990) and Cumming et al., (1989). Ascéncion Delaney (2008) also found proficiency to have modest effects on essay writing. However, language proficiency and linguistic migration background generally appear related to overall educational level (Federal Statistical Office, 2021; OECD, 2017). In any case, with respect to reading-to-write, very little research has confirmed this relationship (Ascéncion Delaney, 2008), and only in the university context. Little is known about the impact of educational level on reading-to-write performance in lower education levels. Furthermore, even if stronger students perform better in reading-to-write tasks, a previous intervention study in the VET context demonstrated that students with lower language proficiency and in academically weaker training programs benefited more from a process-oriented writing instruction approach which included reading-to-write tasks (Hoefele et al., 2017). This suggests that writing improvement may occur independently of overall higher academic or proficiency levels.

Study goals and hypotheses

In summary, few studies have to date examined the development of VET students' literacy skills or the concrete challenges of VET students in reading and writing. (Becker-Mrotzek et al., 2006; Efing & Kiefer, 2018). Furthermore, little is known about the effects of different teaching approaches, let alone how factors such as linguistic background or the type of training may impact effects, as intervention studies are lacking. Specifically, there is no research evidence on whether and how the development of VET students' writing skills benefits from scenario-based reading and writing education where reading and reading comprehension instruction is used to leverage writing (Schneider et al., 2013).

In response to these gaps, the research project 'Integrated Reading and Writing Support in Vocational Education and Training' (2018–2020) examined the effects of a scenario-based reading and writing education, which integrated reading and writing tasks, aiming to improve students' performance on writing tasks. In this project, integrated reading and writing approaches, which are common in second language learning (Hirvela, 2016), were not only seen as a source of content knowledge, but also as a way to develop language competence (Hoefele & Konstantinidou, 2016, 2018; Hoefele et al., 2017). Reading instruction was used to enhance common sources of knowledge for planning, drafting and revising students' writing. Reading tasks were used to raise awareness not only with respect to content but also the linguistic resources available in the reading texts. This approach was particularly intended to meet the needs of linguistically diverse multilingual VET students in ABU general education classes. It is the results of this project that are reported here.

Thus, as the overarching purpose of this study was to measure the impact of a scenario-based reading and writing education. Stepwise, the three goals are, first, to determine whether scenario-based reading and writing education, which is grounded in reading-to-write, improves writing in the VET context. Second, since reading and writing are related, as identified in the literature review above, it is important to determine whether scenario-based reading and writing education has an impact on reading in itself. Third, do individual variables (i.e., linguistic background and the educational level which are relevant to the VET context) have any influence on the impact of scenario-based reading and writing education. Our interest here is not the direct effect of language proficiency and linguistic migration background on writing, but in whether these variables have an influence on the intervention effects. Three research questions guide the study:

RQ1 Does scenario-based reading and writing education yield a significant positive effect on VET students' text quality?

RQ2 Does scenario-based reading and writing education yield a significant positive effect on VET students' reading ability?

RQ3 Which factors moderate effects revealed from RQ1 and RQ2?

The answers to these questions were sought by means of an intervention study wherein an experimental group was exposed to the scenario-based reading and writing education and a control group was exposed to the regular curriculum teaching approach that did not use the scenario-based approach over the period of a semester. Text quality and reading skills were measured by means of a writing test and a reading test prior to and after the semester in question in both groups. Student profile data (training profile and linguistic background) were additionally considered, and gathered by means of a student background questionnaire. Linear mixed modelling, in which variables of time (pretest and posttest), group (experimental and control), training profile, linguistic background, and reading skills were considered as fixed effects, was used to answer the research questions.

Hypothesis 1 With respect to RQ1, while some improvement in text quality from pretest to posttest as a product of language instruction was anticipated in both groups, we expected students in the experimental group (EG) to display significantly greater improvement in text quality than students in the control group (CG), from pretest to posttest.

Hypothesis 2 With respect to RQ2, although reading, within the scenario-based reading and writing education, is integrated together with writing, the function of reading within the approach is to provide the means, through gathering information found in texts, to enhance students' argumentative writing for explicit purposes, rather than to improve reading skills *per se*. Reading is thus used to leverage writing. Given that reading practice is also integrated into the regular educational program,

we expected reading scores to behave similarly in both the control group and the experimental group.

Hypothesis 3 With respect to RQ3, as there is some evidence that (linguistically) weaker students in particular may benefit from process-oriented instruction that involves reading-to-write (Hoefele et al., 2017), differential effects were expected as a function of learner-related variables. These effects were explored to determine which students benefited most from the scenario-based reading and writing education teaching approach. The two main student variables (see section *Student profiles in VET programs* below for details) of training profile (VET program: EBA, EFZ) and linguistic background (L1, 2L1, L2) were considered, as well as reading ability as measured in the reading test. Our interest in the present study is not the direct effect of language proficiency and linguistic migration background on writing, but in whether these variables have an influence on the intervention effects.

Methods

Research design

Design

To answer the research questions, that is, to determine whether the teaching approach under study results in a measurable difference in writing output within the VET context, for different groups with differing profiles, the study employed a quasi-experimental pretest-posttest control group design. In this design, the experimental group was exposed to the scenario-based reading and writing education teaching approach, and the control group was exposed to the regular curriculum teaching approach. Measurements of reading and writing constituted outcome variables. Data collection took place in the first semester of the participating students' first year of apprenticeship (October 2018–January 2019) as part of their ABU (general education) classes. Four to five weeks after school began, and 1–2 weeks before the start of the intervention, the reading and writing tests were administered via a paper-and-pencil pretest (T0; see Table 1 below). Test sessions were conducted by trained test administrators. During the first session, all students and teachers in the control group (CG) additionally answered respective questionnaires, the former to gather background information and the latter to profile teaching behavior within the CG. The pretest was followed by the pedagogical intervention of 12 to 14 lessons in the experimental group (EG). During this time, the control group (CG) received no special treatment, that is, it followed the regular ABU school curriculum taught by teachers who were not introduced to the scenario-based reading and writing education. Three to four weeks after the intervention, the reading and writing posttests were administered to the EG and CG.

Table 1 Study design summary: timeframe for both experimental and control group activities, with pre- and posttests of reading and writing

	Sept.	Oct.–Jan.	Feb.–Mar.
EG	Pretest	INTERVENTION PROGRAM in the context of ABU Same learning goals as the CG 12–14 lessons for reading and writing Writing products: 3 Three-day teacher training (introduction to scenario-based reading and writing education; development of the intervention learning material)	Posttest
CG	Pretest	Language lessons based on the regular ABU curriculum Same learning goals as the EG 6.5–16 lessons dedicated to reading in 18 weeks Reading of shorter texts (shorter than one page): 10–22 texts Reading of longer texts (longer than one page): 2–10 texts 6.5–13 lessons dedicated to writing in 18 weeks Shorter writing products (shorter than one page): 4–25 Longer writing products (longer than one page): 0–10 No teacher reading-to-write training	Posttest

Setting: regular reading and writing practice in vocational schools

In Swiss VET schools, in addition to vocational education subjects, students take a general education course (*Allgemeinbildender Unterricht*, ABU). It consists of two to three lessons per week and combines the teaching of language and communication with various general subjects such as economics, civic education, and law. The corresponding curriculum framework (BBT, 2006), its implementing regulations (SERI, 2014), and the school curricula (BBW, 2017) include basic instructions on the relationship between the teaching of general education content and the teaching of language and communication.

Setting: student profiles in VET programs

Swiss VET classes are highly heterogeneous with regard to students' linguistic background. In addition to monolingual (L1) students whose first language is (Swiss) German (and who, actually, in the Swiss diglossic context, use at least one local Swiss German dialect and the Swiss-standard variety of German), we find substantial proportions of bi-/multilingual (2L1) students (who grow up with two or more languages, for instance, at least one more home language other than (Swiss) German), as well as students who learned German as an additional language (L2) later in life (e.g., first generation child or adolescent migrants, refugees).

Students in VET education are also characterized by the educational program they select. These are here referred to as *training profiles*. There are two main training profiles in Swiss VET education: Two-year VET programs leading to the *Federal VET Certificate* (*Eidgenössisches Berufsattest* EBA) focus on the development of practical skills and issue a recognized qualification for a specific

occupational profile. Three- or four-year VET programs leading to the *Federal VET Diploma* (*Eidgenössisches Fähigkeitszeugnis* EFZ) target more in-depth professional skills and open access to tertiary-level professional education. EFZ programs are generally assumed to be more academically challenging than EBA programs. Their learners have the option of preparing for the *Federal Vocational Baccalaureate* (FVB). This preparatory course covers general education subjects and qualifies learners for enrolment in a Swiss university of applied sciences (SERI, 2022).

For the intervention study, VET classes from five schools in the greater Zurich area were assigned to the experimental group (EG) and control group (CG). From these five schools, a total of ten first year classes, whose teachers volunteered to participate in the study, represent the EG. Within the same schools, an equal number of classes that matched the EG as closely as possible with respect to training profile (EBA, EFZ) and training profession were chosen to constitute the CG. For the EG and CG sample described below (N=285 participants), at a recommended power of 0.8 (Cohen, 1988), an effect size of 0.167 (Cohen's *f*) is detectable (alpha=0.05).

Experimental group participants

The EG comprised 134 VET students from 10 classes. Students' ages ranged from 15 to 35 (M: 18.3, SD: 4.13). 88 students were male (65.7%), 36 were female (26.9%), and 10 provided no gender information. 68 students (50.7%) were attending the 2-year EBA program (hereafter, "EBA students"), with training professions such as hairdressers, construction workers, carpenters, and mechanics. 66 students (49.3%) were attending the 3-year EFZ program (hereafter, "EFZ students"), with training professions including biology lab assistants, bakers, and bicycle/scooter mechanics.

Control group participants

The CG comprised 151 VET students from 10 classes. Students' ages ranged from 15 to 40 (M: 17.5; SD: 3.54). 93 students were male (61.5%), 32 were female (21.2%), and 10 provided no gender information. 88 students (58.3%) were EBA students, with training professions including logisticians, company technicians, hairdressers, and construction workers. 63 students (41.7%) were EFZ students, with training professions such as logisticians, chemical lab assistants, bakers, masons, and electricians. 47 students (31.1%) indicated German/Swiss German to be their first language (L1). 38 (25.2%) were 2L1 students. 35 (23.2%) were L2 students, with a range of different L1s (e.g., Arabic, Tigrinya, Tamil, Kurdish). 31 students provided no language information. Both groups are summarized in Table 2.

Table 2 Overview of participants by group

	N classes	N students	Age	Gender			Training profile		Linguistic background			
				Male	Female	n.a.	EBA	EFZ	L1	2L1	L2	n.a.
EG	10	134	15–35 (M: 18.3, SD: 4.13)	88	36	10	68	66	54	31	35	14
CG	10	151	15–40 (M: 17.5, SD: 3.54)	93	32	10	88	63	47	38	35	31

Independent variables

As mentioned in the *Study goals and hypotheses*, the independent variables considered in the study were: time (pretest and posttest), group (experimental and control), training profile (EFZ and EBA), linguistic background (L1-, 2L1-, L2-students), and reading skills. *Reading skills* was also considered as an outcome variable for RQ2.

Intervention design

Conditions in the experimental group

The intervention for the EG was designed based on the theoretical considerations described in the "Introduction" section, and was implemented in three teaching scenarios that were piloted in a class with a similar academic and linguistic profile to the experimental classes. The scenarios and the corresponding teaching materials were developed in collaboration with VET teachers and involved the following communicative situations:

Scenario 1 "Solving a conflict with the company".
Scenario 2 "Reacting to a modified school cafeteria menu".
Scenario 3 "Reporting ethnic discrimination in the company".

All scenarios followed roughly the same action-based problem-solving process:

Step 1 Presenting the scenario,
Step 2 Discussing the problems involved,
Step 3 Planning action,
Step 4 Reading text(s) as the source(s) of content and linguistic knowledge,
Step 5 Generating and structuring ideas,
Step 6 Drafting writing while consulting text resources,
Step 7 Giving (peer) feedback and revising,
Step 8 Final drafting and editing (Konstantinidou et al., 2016; Hoefele et al., 2017).

The first three steps introduce the scenario, stimulating the planning of actions needed for solving the problem. These reflect the theoretical underpinnings of scenario-based reading and writing education, described in the "Introduction" section. Reading (Step 4) then results as a natural action from the scenario and is functional in that it is used to selectively gather information that is subsequently organized and transformed into ideas and argumentation (Step 5), before being integrated into the students' texts (Steps 6 and 8). Students' writing intends to solve the problem which was described in the scenario. Peer feedback is implemented to support students to improve their writing and to enhance their capacity to reflect the impact of language on others (Step 7). During the whole process (Steps 1 to 8), the teachers provide the

reading and writing support they consider necessary for the respective group of students (macro-scaffolding).

The scaffolding of writing through reading tasks was explicitly based on reading instruction and was designed in collaboration with teachers according to Hammond & Gibbons' concept of macro-scaffolding (Hammond & Gibbons, 2005). Macro-scaffolding refers to scaffolding for groups or subgroups of learners planned by teachers in advance. Micro-scaffolding, on the other hand, refers to the concrete interaction between teachers and learners within the classroom (Hammond & Gibbons, 2005). Thus, teachers first assessed the expected linguistic, textual, and content-related challenges for learners and compared these collectively to their experience of learners' current proficiency levels. In this way, the tasks, their sequencing, and the materials were designed to support students' steps into the "zone of proximal development" (Gibbons, 2015, p. 89; Hammond & Gibbons 2005, p. 8). Given the short duration of the intervention (12–14 lessons), the supporting materials could not be phased out gradually, as foreseen in Gibbons' scaffolding approach (Gibbons, 2015). To build up reading-to-write routines, learners were thus given similar support materials in all three scenarios. Individual scaffolding on the micro-level was not prioritized in order to keep the conditions of the intervention comparable for all classes involved.

Example scenario

To illustrate the approach, we describe Scenario 1 in detail.

A VET trainee is regularly late for soccer practice or misses it completely because he has to work overtime at his vocational training company (bicycle shop). The soccer coach now gives him an ultimatum. If he misses any more practices, he will be kicked off the team. However, the vocational training instructor informs him by letter that he must work overtime and will be compensated later. The trainee's classmates are asked to help him write a letter to the vocational training instructor to find a solution (Step 1).

After a short discussion of the problem (Step 2), students designed an action plan, concluding that they needed information on overtime regulations (Step 3). With the help of two informative texts, one about employment law and the other about conflict management, the students prepared a presentation (Step 4). This presentation task was supported by a worksheet with instructions to take notes about the topic and to find relevant information, possible arguments, and conflict resolution strategies. Thus, they collected, structured, and presented the information needed for writing the reply to the vocational training instructor (Step 5), in preparation for their letter.

During writing, students also received support, in the sense of macro-scaffolding (Gibbons, 2015), for formulating and structuring their own text, while arguing their position. Before writing (Step 5), they were encouraged to check the training instructor's letter (from the scenario introduction) for useful content, genre markers, and linguistic means of expression. During drafting (Step 6), the students were given a letter template with example language for writing an introduction and conclusion, and for argumentation. Students could refer to their notes or to the reading

texts at any time. After writing the draft (Step 6), the students were instructed to use a checklist to reflect on whether their letter contained relevant formal, linguistic, and content aspects. Writers also received a classmate's peer feedback which was supported by a checklist with the same criteria (Step 7). The draft texts were then revised based on the peer-feedback, resulting in the final letter to the vocational training instructor (Step 8).

Scenarios 2 and 3 followed the same problem-solving steps and included similar types of reading and writing tasks. In Scenario 2, students received a letter from the school cafeteria operator, which was to be answered by email. Students were asked to give their opinion on the proposed change to regional and organic dishes, with more vegan or vegetarian options, and the resulting increase in prices (Step 1). Again, the students were required to discuss (Steps 2–3) and prepare a presentation based on three informative texts about the advantages and disadvantages of organic and/or vegetarian/vegan nutrition (Steps 4–5). To support the presentation task, students used worksheets similar to those described above. To support writing, students were required to fill in a form containing the three steps of convincing argumentation: stating the problem, substantiating one's demand/position (including an example), and concluding. They then followed Steps 6–8 (drafting, giving peer-feedback, revising, and writing the final version).

In the third scenario, one of the students' classmates described how he was being bullied in his vocational training company by several colleagues because of his ethnic origin. He asked his classmates for help (Step 1). The students discussed the problem and planned the next steps (Steps 2–3). This scenario contained more challenging reading and writing tasks writing tasks, as students worked on three different legal/advice texts on their rights relevant to workplace bullying (Step 4). In small groups, they worked on the different texts, using worksheets to prepare presentations. During classroom presentations, students took notes, supported by a worksheet again, to prepare and pre-structure their drafts (Step 5). Before writing the draft (Step 6), students filled in a form with the same three-step argumentation formula described above, then completed steps 7–8.

Overall, the three scenarios increased in source-based writing complexity in terms of the number of sources. At the same time, the same steps of problem-solving (1–8) and support (Steps 4–7) were repeated and subsequently reflected on by students across all scenarios. This was meant to help students (1) build reading and writing routines for mastering writing and (2) understand writing as a natural approach to solving problems in everyday and professional contexts.

Conditions in the control group

Language lessons in the CG, as in the EG, were for the purpose of achieving the learning goals outlined in the ABU curriculum. The manner and specific teaching methods employed to reach these goals are not explicitly described in the ABU curriculum, however. Teachers or groups of teachers are thus free regarding the design of their lessons. Because of this variation in the design of language lessons in VET schools, a questionnaire was administered to teachers in the control group to characterize their reading and writing teaching practice (e.g., time spent on reading/writing

in class, amount and types of reading/writing assignments given to students, amount and types and aspects of feedback provided).

Teachers in the control group reported that they always (29%) or often (71%) included reading instruction within ABU lessons. Most estimated that approximately 50% of the ABU lessons were dedicated to reading instruction. The number of short texts (less than one page) read in CG classes during the intervention period ranged from 10 to 22. The number of longer texts read was between 2 and 10. Teachers in the CG also reported that they often (83%) included writing instruction within ABU. Most estimated that 20 to 25% of their ABU lessons are devoted to writing instruction. During the intervention period, students in the control group wrote between 4 and 25 shorter texts (less than one page) and 0 to10 longer texts.

While these data verify that both reading and writing took place in the control group in a somewhat balanced fashion, there are key differences compared to the EG. In terms of text reception and output, the CG on average read more texts and produced more writing products than the EG (see Table 1). However, more importantly, neither reading nor writing instruction took place within scenarios, as described above, in the CG. Furthermore, reading and writing were not integrated in any way in terms of reading-to-write, as described earlier. Instead, the CG is characterized by reading and writing instruction as distinct elements, with no explicit link between reading tasks and writing tasks. This is in opposition to the EG where the full integration, by design, of reading and writing elements took place within a scenario-based teaching approach.

Measurements

Measurements of reading and writing were taken prior to and after exposure to the intervention programs in the CG and EG in the form of a test that incorporated a reading task and a writing task. The tasks were scenario-based, and included reading comprehension as a pre-writing task. Students needed to read through a short text (one page, five clearly separated paragraphs), providing relevant legal background for the writing task that followed (pretest: consumer rights in case of product damage; posttest: holiday regulations for VET students).

The graded reading task consisted of three parts designed to measure reading comprehension. In the first subtask, five subheadings had to be appropriately assigned to the paragraphs of the reading text. The second subtask required short text productions, as students either had to provide definitions, in their own words, for two core concepts from the reading text (in the pretest) or answer two comprehension questions (in the posttest). The third subtask was a multiple-choice task consisting of four items. For each item, students had to pick the correct statement from four options (one statement consistent with the reading text, three close distractors). The reading scale consisting of 11 items was found to have minimally acceptable reliability (Kuder-Richardson KR20=0.7) at both pretest and posttest (Kline, 2000).

The graded writing tasks arose from the scenario and the action plan that were introduced in the introductory sections of the pre- and posttest, and followed up on the reading comprehension tasks. In both cases, students were asked to write a

(formal) email, formulating and convincingly substantiating a complaint and/or request.

In the pretest, the email was addressed to an electronics retailer. Students were expected to explain the problem (i.e., an incorrectly charged express repair service for a new computer) precisely and in their own words, and to formulate and justify their complaint or request (i.e., not to pay for the express repair service). In the post-test, students addressed their email to their vocational training instructor. They were expected to explain the situation (i.e., request to participate in a 3-day school trip) precisely and in their own words, and to argue for their request (i.e., participation in the school trip without deduction of salary and vacation days) convincingly.

For both the pre- and posttest, students were given two blank pages to write a draft. For their final text, they were given an email form, including boxes for the recipient's email address, subject line, and a two-page text box with lines to write on. This format was meant to implicitly draw students' attention to the target genre. Writing was supported by a checklist stating the evaluation criteria for student texts: clarity, comprehensibility, persuasive power, structure and coherence, word choice and register/style, and formal correctness. The writing scale (was found to have strong reliability (Kuder-Richardson KR20 > 0.9) at both pretest and posttest (Kline, 2000).

Data coding and text quality rating procedures

For the evaluation of the students' texts, we used a combination of analytic and holistic scoring (Weigle, 2002). While some of the subscales demanded a holistic judgement (e.g., global score for communicative impact), others included more detailed criteria (e.g., analytic scale for genre convention knowledge) or even a scale based on numbers of errors (e.g., correctness). Text quality was thus evaluated using six subscales, clustered into three main dimensions or competences (Konstantinidou et al., 2016).

1. Linguistic competence with two subscales: correctness and style/word choice.

Correctness refers to students' ability to produce texts with correct grammar, spelling, and punctuation. *Style* mainly refers to students' lexical abilities, that is, their ability to use precise words from a broad vocabulary range in an appropriate register.

2. Genre competence with two subscales: genre convention knowledge and structure/ coherence.

The *genre conventions* subscale refers to students' ability to conform to the formal conventions of the genre in question. *Structure/coherence* refers to the degree to which students' texts are "meaningfully and logically structured and can be read fluently" (Konstantinidou et al., 2016).

3.   Pragmatic competence with two subscales: content and communicative impact.

Content refers to students' ability to clearly state the purpose of writing and to provide all relevant information. *Communicative impact* refers to students' ability to convince the reader and was judged for the entire text overall.

These dimensions and subscales are based on previous studies from German-speaking countries (e.g., Becker-Mrotzek & Böttcher 2012; Harsch et al., 2007), which have been adapted to and reliably used in prior projects in the VET context (e.g., Konstantinidou et al., 2016).

Each text was evaluated on the six subscales. For each subscale, a score ranging from 0 (very poor/unscorable to 4 (excellent) was attributed. Mastery levels (0 to 4) were defined beforehand for all subscales and made available to raters in a detailed rating scale (see Table 3 for the *content* and *communicative impact* subscales). This allowed for the evaluation of each text along the six subscales, independently of other test takers' performances (Konstantinidou et al., 2016). Very short texts (under 50 words) were scored 0 on the three subscales of *correctness*, *style*, and *structure/coherence*, since the quality of these parameters depends on text length. For example, very short texts tend to have inflated lexical diversity (Jarvis, 2013, 2017) and contain fewer orthographic errors.

All texts were rated by a team of five trained raters. Each text was independently rated by two raters. Classes were randomly assigned to a main rater, counterbalancing schools, training profiles (EBA, EFZ), groups (EG, CG), and test versions (pretest, posttest representing the times of test administration). Second raters were also assigned to classes counterbalancing schools, training profiles, groups, and test versions. To avoid rater bias, the assignment was blind, that is, raters received the texts without knowing which class, group or training profile texts belonged to. To ensure reliability, all texts were first evaluated independently by the two raters. After scoring independently, raters had the opportunity to discuss their scores in rater teams (main rater and second rater), applying consensus scoring (NAEP, 2008; Robinson 2000). In case of disagreement, the score of the main rater prevailed. Such disagreement, where a consensus could not be found, was rare, occurring 2% of the time.

Treatment fidelity

All teaching materials – scenarios, reading texts, supporting worksheets for reading and writing, contextualized language exercises, and checklists for self-evaluation and peer feedback – were developed in collaboration with the VET teachers who conducted the intervention in the EG. In two full-day workshops, teachers were introduced to the teaching concept and used their practical knowledge to integrate the concept into their teaching practice. This *teachers-as-mediators* approach resulted in intensive teacher involvement and strong commitment, which had previously proven to be an important prerequisite for treatment fidelity (Hoefele & Konstantinidou, 2018).

In addition, teachers were provided with teaching guidelines for each scenario, including the individual phases of teaching, timing, as well as notes on content and the use of worksheets, interaction, and additional comments. To ensure standardized

Table 3 'Pragmatic Competence' scoring criteria and descriptors (based on the codebooks of the German studies DESI (Deutsch Englisch Schülerleistungen International) and VERA (Vergleichsarbeit)

Scoring criteria	Score	Descriptor
Content	4	Texts at this level refer to all issues in sufficient detail (e.g., reason for writing, class description, description, description and justification of intentions, gratitude). The purpose of writing is clear. Argumentation is strong and weighty.
	3	Texts at this level refer to all issues (e.g., reason for writing, class description, description and justification of intentions, gratitude) and deal in sufficient detail with at least three of them. The purpose of writing is clear. Argumentation is good and weighty enough.
	2	In texts at this level, only one issue is missing. However, all other issues are dealt with in sufficient detail. OR: Texts at this level refer to all issues (e.g., reason for writing, class description, description and justification of intentions, gratitude), but two of them are not dealt with in sufficient detail.
	1	In texts at this level, more than two issues are missing. However, all other issues are dealt with in sufficient detail. The purpose of writing is clear. OR: Texts at this level refer to all issues (e.g., reason for writing, class description, description and justification of intentions, gratitude) but not in sufficient detail.
	0	Texts cannot be scored.
Communicative impact	4	Texts at this level would definitely persuade the reader. They are very convincing regarding content, form and rhetoric.
	3	Texts at this level are likely to persuade the reader. They are convincing regarding content, form and rhetoric.
	2	Texts at this level could fail to persuade the reader. They are only partly convincing regarding content, form and rhetoric.
	1	Texts at this level are likely to fail to persuade the reader. They are only slightly convincing regarding content, form and rhetoric.
	0	Texts cannot be scored.

implementation of the intervention, systematic classroom observations were conducted in seven of the EG classes. Observations were evaluated with reference to the teaching guidelines. Short implementation reports showed that treatment fidelity was high. The only teaching guideline deviation noted was concerning the timeline, mainly in the EBA classes, where teachers invested more time in some of the teaching phases.

With respect to treatment in the CG, as discussed in the "Conditions in the control group" *section above*, the CG invested time on both reading and writing, with some differences compared to the EG in terms of input and output. More significantly however, the treatment in the control group did not involve any form of scenario-based integration of reading-to-write, treating reading and writing as distinct elements.

Data analysis

As the core outcome variables of interest, data inclusion and exclusion for analysis were based on whether test scores were measured (both writing and reading). Research questions were investigated using linear mixed effects regression modeling (LMM) rather than repeated-measures ANOVA given its advantage in specifying random effects, and dealing with missing data. For each LMM fit for RQ1 and RQ2, a manual step-up (forward-selection) process was applied, wherein to the null model (intercept only) random variables were added, followed by main effect terms and their interactions. Leading up to the interaction effect representing the intervention effect, model selection was based on likelihood ratio tests to compare models' goodness of fit. The principle of parsimony was applied, in which additional terms were accepted only if the resulting model exhibited improved fit and a significant test result. If tests of pairs of nested models were not found to be significant, the simpler model was preferred. Significance of effects by means of F-tests and p-values were calculated using Satterthwaite's method, following Crawley (2007) and Mangiafico (2016). The normality of residuals and homoscedasticity were checked by means of plots and model diagnostics.

If the intervention effect investigated was confirmed (RQ1 and RQ2), models were subsequently tested to determine whether additional factors moderate the intervention effect (RQ3). Here, based on the principle of marginality (Fox, 2016) the moderation effect was tested directly by fitting models consisting of the three-way interaction between candidate factors and the intervention effect plus all lower-order terms marginal to them. Here, only the three-way interaction is tested rather than the constituent main effects or lower-order interactions. The significance of effects, calculation of p-values, and model diagnostics were performed as for RQ1 and RQ2. Following, the significance of effects is reported for all models either by means of t-tests and/or F-tests.

To answer RQ1, that is, whether scenario-based reading and writing education yields a positive effect on VET students' text quality, the following variables were considered:

(a) Text quality (measured at pretest and posttest – dependent variable).
(b) Time (pretest and posttest).
(c) Group (EG and CG).

To determine the effect on reading comprehension (RQ2), Time and Group were considered together with reading test scores as the dependent variable.

To answer RQ3, that is, whether other factors moderate the intervention effect if confirmed in RQ1 and RQ2, the following additional variables were considered:

(d) Training profile.
(e) Linguistic background.
(f) Reading comprehension (if the intervention effect for Text quality from RQ1 was confirmed).

Results

The achieved sample used to produce the outcomes of the linear mixed models was the result of a dropout rate of 37%. This reduction from the intended sample (N = 285) to the usable dataset (N = 180) was caused by two factors. The first was one entire class dropping out from the study midway through the intervention program. The other was overall individual absences either at pretest or posttest due to, for example illness or work program requirements on the day of the tests. In spite of this dropout, the balance of distributions of factors within both the EG and CG remained largely unchanged. A second power analysis for linear mixed model terms showed that a sample of 180, using Cohen's (1988) recommended power of 0.8 and an alpha of 0.05, is sufficient to detect effect sizes of 0.202 (Cohen's *f*, equivalent to a Cohen's *d* of 0.404).

All analyses were run in R (R Core Team 2021, version 4.04). The following packages were used (a) for regression modelling, plotting and diagnostics: *lme4* (Bates et al., 2021), *lmerTest* (Kuznetsova et al., 2020), *sjPlot* (Lüdecke, 2021); (b) for power analysis and effects testing: *effectsize* (Ben-Shachar et al., 2021), *pwr* (Champely, 2020).

RQ1: development of text quality intervention effect

For RQ1, Text quality was taken as the dependent variable, Class and Participant within Class were added as random effects, plus fixed effects of Time (pretest, posttest) and Group (EG, CG) and the interaction between them, each in turn and tested for goodness of fit. The final model specification for Text quality was as follows:

Model 1: *Text quality ~ 1 + (1|Class) + (1|Class:Participant) + Time + Group + Time:Group*.

Table 4 Contrasts in estimated marginal means for Text quality from Model 1 (RQ1), from pretest to posttest, by Group (EG = experimental group, CG = control group) with t-statistics and p-values. SE = standard error, df = degrees of freedom

Contrast	Group	Difference	SE	df	t	p
Pretest - posttest	EG	2.34	0.48	178	− 4.86	0.00
Pretest - posttest	CG	0.49	0.48	178	− 1.01	0.31

Table 5 Summary of Model 1 with beta estimates, standard errors, confidence intervals and *p* values

Effect	Estimate	SE	95% CI		p
			LL	*UL*	
Fixed effects					
Intercept	8.73	1.24	6.3	11.16	<0.001
Time	1.66	0.34	0.99	2.33	<0.001
Group	-0.59	1.72	-3.97	2.79	0.73
Time:Group	-1.31	0.48	-2.26	-0.37	0.007
	Variance				
Random effects					
Class:Participant	6.82	2.61			
Class	12.45	3.53			
Residual	10.47	3.24			

Number of observations = 360

CI Confidence interval; *LL* Lower limit; *UL* Upper limit.

Normality of residuals and homoscedasticity were confirmed. A small (Cohen, 1988) significant interaction effect of Time by Group ($F_{1,178} = 7.40$, $p = .007$, Cohen's $f = 0.20$) was found. Based on the principle of marginality (Fox 2016), the lower order terms marginal to the interaction are not to be tested. This significant interaction reflects a differential effect depending on the group. This result confirms the intervention effect, answering RQ1 with respect to writing quality in the affirmative. Contrasts between pretest and posttest marginal means for Text quality are reported in Table 4 and plotted in Fig. 1. The model summary is provided in Table 5.

RQ2: development of reading comprehension intervention effect

For RQ2, Reading test scores were taken as the dependent variable. The same configuration as above for writing was applied for fixed effects, the interaction effect and random effects. Thus, the final model specification for reading was as follows:

Model 2: *Reading ~ 1 + (1|Class) + (1|Class:Participant) + Time + Group + Time:Group.*
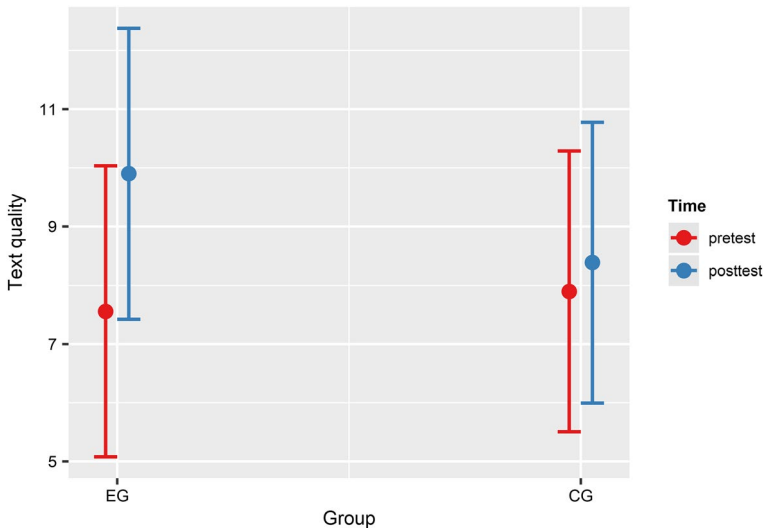
Fig. 1 Estimated marginal means for Text quality from Model 1 (RQ1) by Group at pretest and posttest.

Normality of residuals and homoscedasticity were confirmed. In contrast to the model for Text quality, a significant main effect of Time was found alone ($F_{1,178}$ = 112.2, $p < .001$). Neither Group ($F_{1,16}$ = 0.48, $p = .50$, ns), nor the interaction between Time and Group ($F_{1,16}$ = 0.48, $p = .73$, ns) were found to be significant. This result reflects differences in reading scores with time in both groups, but without a differential effect depending on the group, meaning that Reading scores did not change as a function of Group alone. These results do not confirm an intervention effect, answering RQ2 with respect to reading comprehension in the negative.

RQ3: moderators of intervention effect

First, the interaction between Linguistic Background and the intervention effect for Text quality (confirmed in answering RQ1) was tested. Given that the interaction term of interest is therefore Time:Group:Linguistic background, applying the principle of marginality requires the inclusion of all lower order terms that are marginal to it. Thus, the model specification was:

Model 3: *Text quality ~ 1 + (1|Class) + (1|Class:Participant) + Time * Group * Linguistic background.*

Normality of residuals and homoscedasticity were confirmed. Also following the principle of marginality, only the highest order interaction is tested in this way. No significant interaction between Linguistic background and the intervention effect was found ($F_{2,174}$ = 0.07, $p = .93$, ns). This result reflects the presence of the intervention effect occurring similarly in all three linguistic background categories considered.

Table 6 Contrasts in estimated marginal means for Text quality from Model 5, from pretest to posttest, by Training profile (EBA and EFZ) and Group (EG = experimental group, CG = control group) with t-statistics and p values. SE = standard error, df = degrees of freedom

Contrast	Training profile	Group	Difference	SE	df	t	p
Pretest-posttest	EBA	EG	3.68	0.73	176	− 5.01	0.00
Pretest-posttest	EFZ	EG	1.37	0.63	176	− 2.17	0.03
Pretest-posttest	EBA	CG	0.42	0.75	176	− 0.55	0.58
Pretest-posttest	EFZ	CG	0.54	0.62	176	− 0.87	0.38

Second, the interaction between Reading comprehension and the intervention effect for Text quality (confirmed in answering RQ1) was tested by means of the model specification:

Model 4: *Text quality ~ 1 + (1|Class) + (1|Class:Participant) + Time * Group * Reading*.

Normality of residuals and homoscedasticity were confirmed. No significant interaction between Reading comprehension and the intervention effect was found ($F_{1,241}$ = 0.008, $p = .93$, ns). This result reflects the presence of the intervention effect occurring independently of reading scores.

Finally, the interaction between Training profile and the intervention effect for Text quality (confirmed in answering RQ1) was tested with the model specification of:

Model 5: *Text quality ~ 1 + (1|Class) + (1|Class:Participant) + Time * Group * Training profile*.

Normality of residuals and homoscedasticity were confirmed. The interaction between Training profile and the intervention effect was not found to be significant at $p = .05$, but nearly so ($F_{1,176}$ = 3.16, $p = .08$, ns). However, inspection of the contrasts between pretest and posttest marginal means for Text quality reported in Table 6 and plotted in Fig. 2, reveal a strong difference between the improvement in the two Training profiles (EBA and EFZ) within the experimental group. An ad hoc comparison of the change (between pretest and posttest) in predicted mean values of Text quality revealed a significant difference, between EBA (M = 2.1, SD = 1.6) and EFZ (M = 0.94, SD = 0.42), t(80) = 25.9, $p < .001$. These results reflect that, while the intervention effect found in the experimental group extends to both Training profiles, those in the EBA display greater improvement overall.

Discussion

This study investigated the effects of scenario-based reading and writing education in a VET context. The results confirm Hypothesis 1 regarding the beneficial effect of scenario-based reading and writing education on text quality compared to
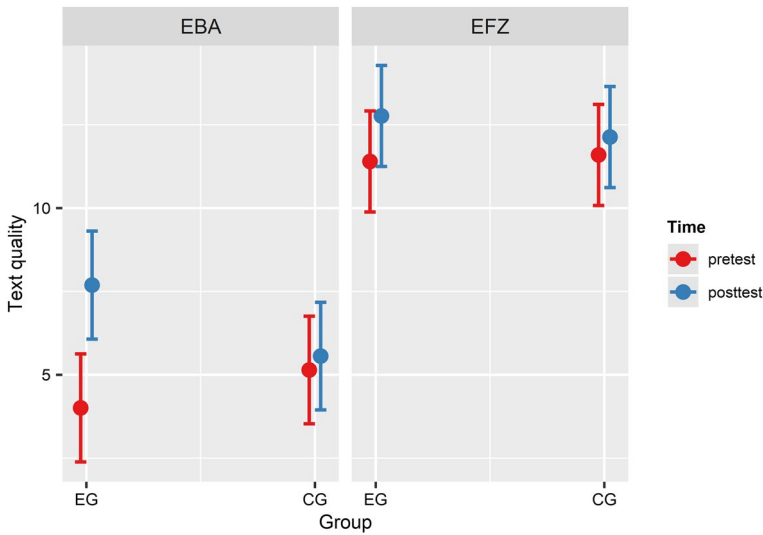
Fig. 2 Estimated marginal means for Text quality from Model 5 by Training profile (EBA and EFZ) and Group at pretest and posttest.

the regular language and communication teaching in VET schools. As expected, the EG showed significant development in terms of writing competence over time and compared to the CG. This positive effect was achieved despite the short time span of the intervention (12–14 lessons). To examine the sustainability of the identified intervention effect in more detail, it is recommended that future studies target longer-term interventions, and also include a delayed follow-up test. Results from a prior study with a similar process-oriented and language sensitive approach found the beneficial effects to be sustained three months after the intervention (Hoefele & Konstantinidou, 2016, 2018).

The results of this study also confirmed Hypothesis 2, regarding the effects of the intervention on reading comprehension ability. Given that reading instruction (including reading strategies instruction) took place in the regular curriculum, we did not expect greater improvements in reading in isolation in the experimental group compared to the control group. To further explore the question of the extent to which and the manner in which scenario-based reading and writing education can contribute to literacy development, not only through reading-to-write, but also through writing-to-read (Graham & Harris, 2017), follow-up studies will be needed. These studies should not only include the writing-to-read approach in their intervention design, but also in the test design. Considering the importance of the alignment between teaching and testing, the current study used the scenario-based reading-to-write-approach also for testing students' ability. In this sense, to test the effects of a combined reading-to-write and writing-to-read approach, we propose to additionally include measuring students' functional reading competences after the writing phase.

With respect to the moderation effects of learner-related variables and additionally reading ability, this third hypothesis was confirmed, although only for

training profile. In terms of training profile, as expected, students in the academically weaker EBA classes (EG-EBA) benefitted more from the intervention than their peers in the EFZ classes (EG-EFZ). Scenario-based reading and writing education, with its strong focus on reading and writing support at all stages of the writing process, can thus be assumed to meet the needs of (academically) weaker students particularly well, while at the same time not constituting a disadvantage for academically stronger students. This is evident in Fig. 2, where improvement from the intervention is also seen in the academically stronger group (EG-EFZ).

Linguistic background, in contrast, did not give rise to significant differential effects with respect to the experimental intervention. This implies that students respond similarly well to the experimental treatment, independently of their linguistic background. This, in turn, suggests that the scenario- and source-based approach meets the needs of linguistically heterogeneous classrooms, which are the norm in VET contexts. Nevertheless, follow-up studies with larger cohorts are needed to confirm these results, and to explore more closely distributions and potential interactions between complex linguistic background configurations and other learner-related variables within the two academic streams in VET contexts.

It is worth noting that a considerable number of students did not indicate their linguistic background (EG: 10.4%; CG: 20.5%), While this had no effect on the analyses and results presented here (largely within the dropout group), it does raise questions about possible stigma regarding languages students identify with as a potential cause for the reluctance to provide language information. For instance, L1s such as English or French have potentially higher social status than other L1s, such as Arabic or Albanian, which are often negatively portrayed socially (see Shohamy 1998). This area requires further studies.

Importantly, overall, all subgroups in our sample, including the L1 and 2L1 students in the (academically more challenging) EFZ programs, still achieved only modest text quality ratings, even though the first-year VET curriculum explicitly includes the competences tested and rated in this study as learning objectives. This raises the more general question of whether language and communication skills training must be reinforced in VET contexts, and if so, how this is likely to be achieved, given that literacy competence, in the broad sense is increasingly relevant to life-long learning, and to participation in society in general and in the labor market in particular (European Commission, 2018).

To conclude, we note a methodological limitation concerning the monitoring of students' actual behavior during writing tasks. In our study, we did not collect quantitative behavioral data investigating whether, or to what extent students actually integrated reading and writing activities (e.g., input logging, cf. Vandermeulen et al., 2020). Although, in our approach, scenario-based tasks can be inherently considered as stimuli for reading-to-write because functional reading before and during writing is an essential part of the problem-solving activity arising from the scenario, future studies may fruitfully explore the effects of students' actual behavioral patterns (i.e., their level of integrating reading and writing within tasks) on the quality of their writing output.

References

Ammon, U., Bickel, H., & Ebner, J. (2004). *Variantenwörterbuch des Deutschen: Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*. Berlin: Walter de Gruyter

Applebee, A. N., & Langer, J. A. (1983). Instructional scaffolding: Reading and writing as natural language activities. *Language Arts, 60*(2), 168–175.

Ascención Delaney, Y. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes*, *7*, 140–150

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2021). *Lme4: Linear mixed-effects models using eigen and S4*. Retrieved 1 Aug 2021, from https://github.com/lme4/lme4/

Baumann, K. (2014). *Man muss schon ein bisschen mit dem Schreiben zurechtkommen! Eine Studie zu den Schreibfähigkeiten von Auszubildenden im unteren beruflichen Ausbildungssegment im Kontext von Ausbildungsreife*. Paderborn: Eusl Verlagsgesellschaft.

Ben-Shachar, M. S., Makowski, D., & Lüdecke, D. (2021). *Effectsize: Indices of effect size and standardized parameters*. Retrieved 1 Aug 2021, from https://easystats.github.io/effectsize/

Becker-Mrotzek, M., Kusch, E., & Wehnert, B. (2006). *Leseförderung in der Berufsbildung*. Duisburg: Gilles & Francke Verlag

Becker-Mrotzek, M., & Böttcher, I. (2012). *Schreibkompetenz entwickeln und beurteilen* (4th ed.). Berlin: Cornelsen Verlag

Champely, S. (2020). *Pwr: Basic functions for power analysis*. Retrieved 1 Aug 2021, from https://github.com/heliosdrm/pwr

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates

Conrad, S., Pfeiffer, T. J., & Lamb, K. (2018). Improving student writing with research-based instruction: Results from the Civil Engineering Writing Project. Paper presented at *2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah*. Retrieved 1 Mar 2021, https://www.semanticscholar.org/paper/Board-27%3A-Improving-Student-Writing-with-Results-Conrad-Pfeiffer/caae9dd71bf5e0d0982a5db085fb0732261ea781

Corbeil, G. (2000). Exploring the effects of first- and second-language proficiency on summarizing in French as a second language. *Canadian Journal of Applied Linguistics*, *3*, 35–62

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Crawley, M. J. (2007). *The R Book*. Wiley.

Cumming, A., Rebuffot, J., & Ledwell, M. (1989). Reading and summarizing challenging texts in first and second languages. Reading and Writing. *An Interdisciplinary Journal*, *2*, 201–219

Dürscheid, C., Elspass, S., & Ziegler, A. (2011). Grammatische variabilität im gebrauchsstandard – das projekt "Variantengrammatik des Standarddeutschen." In M. Konopka, J. Kubaczak, C. Mair, S. Frantisec, & U. Waßner (Eds.), *Grammar & corpora/grammatik und korpora 2009* (pp. 123–140). Tübingen: Narr.

Efing, C. (2008). „Aber was halt schon schwer war, war, wo wir es selber schreiben sollten." Defizite und Förderbedarf in der Schreibkompetenz hessischer Berufsschüler. In E. M. Jakobs, & K. Lehnen (Eds.), *Berufliches Schreiben. Ausbildung, Training, Coaching* (pp. 17–34). Frankfurt am Main: Peter Lang

Efing, C., & Kiefer, K. H. (Eds.). (2018). *Sprache und Kommunikation in der beruflichen Aus- und Weiterbildung. Ein interdisziplinäres Handbuch*. Tübingen: Narr

Commission, European. (2018). *Annex to the proposal for a council recommendation on key competences for lifelong learning*. Brussels: ETF: European Training Foundation.

Feilke, H., Lehnen, K., Rezat, S., & Steinmetz, M. (2016). *Materialgestütztes schreiben lernen*. Braunschweig: Schroedel.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models*. SAGE Publications.

Gibbons, P. (2015). *Scaffolding language, scaffolding learning. Teaching second language learning in the mainstream classroom* (2nd ed.). Portsmouth: Heinemann.

Gogolin, I. (2008). *Der monolinguale habitus der multilingualen schule*. Münster & New York: Waxmann.

Graham, S., & Harris, K. R. (2017). Reading and writing connections: How writing can build better readers (and Vice Versa). In C. Ng & B. Bartlett (Eds.), *Improving reading and reading engagement in the 21st century* (pp. 333–350). Singapore: Springer. https://doi.org/10.1007/978-981-10-4331-4_15

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243–284. https://doi.org/10.3102/0034654317746927

Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: the contribution of scaffolding in articulating ESL education. *Prospect An Australian Journal of TESOL, 20*(1), 6–30

Harsch, C. (2014). General language proficiency revisited: Current and future issues. *Language Assessment Quarterly, 11*(2), 152–169.

Harsch, C., Neumann, A., Lehmann, R., & Schröder, K. (2007). Schreibfähigkeit. In B. Beck & E. Klieme (Eds.), *Sprachliche kompetenzen. Konzepte und messung, DESI studie* (pp. 42–62). Weinheim & Basel: Beltz.

Hirvela, A. R. (2016). *Connecting reading & writing in second language writing instruction* (2nd ed.). Ann Arbor: University of Michigan Press.

Hoefele, J., & Konstantinidou, L. (2016). Förderung der allgemeinen Schreibkompetenz an Berufsfachschulen: Prozessorientierte Schreibdidaktik zwischen Deutsch als Muttersprache (DaM) und Deutsch als Zweitsprache (DaZ). In R. Kreyer, S. Schaub, & B. Güldenring (Eds.), *Angewandte linguistik in schule und hochschule: Neue wege für sprachunterricht und ausbildung* (pp. 135–164). Frankfurt am Main: Peter Lang. https://doi.org/10.3726/978-3-653-05953-3 (Forum Angewandte Linguistik 58).

Hoefele, J., & Konstantinidou, L. (2018). Förderung der allgemeinen Schreibkompetenz im Bereich der beruflichen Bildung. In C. Efing, & K. H. Kiefer (Eds.), *Sprache und Kommunikation in der beruflichen Aus- und Weiterbildung. Ein interdisziplinäres Handbuch* (pp. 339–348). Tübingen: Narr

Hoefele, J., Konstantinidou, L., Kovacs, M., Sigges, S., & Vacaretu, A. S. (2018). *Integrated reading and writing support in vocational education. Framework*. Retrieved 1 Mar 2021, from https://rewrv et.de/wp-content/uploads/2018/10/O2_Framework_english.pdf

Hoefele, J., Konstantinidou, L., & Weber, C. (2017). Zweitsprachendidaktische Konzepte in der Förderung der allgemeinen Schreibkompetenz an Berufsfachschulen. In K. H. Kiefer & C. Efing (Eds.), *Sprachbezogene curricula und aufgaben in der beruflichen bildung* (pp. 131–156). Frankfurt am Main: Peter Lang. (Wissen - Kompetenz - Text 12).

Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning, 63*, 87–106.

Jarvis, S. (2017). Grounding lexical diversity in human judgments. *Language Testing*, *34*(4), 537–553

Johns, A., & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics*, *11*, 253–271

Kennedy, M. (1985). The composing process of college students writing from sources. *Written Communication*, *2*, 434–456

Krekeler, C. (2002). "Das Ende der Zweitsprachendidaktik - Zweitsprachendidaktik am Ende? Grenzüberschreitungen m berufsvorbereitenden deutschunterricht zwischen Muttersprachendidaktik und Zweitsprachendidaktik." In P. Josting & A. Peyer (Eds.), *Deutschdidaktik und berufliche Bildung* (pp. 62–78). Schneider: Baltmannsweiler.

Konstantinidou, L., Hoefele, J., & Kruse, O. (2016). Assessing writing in vocational education and training schools: Results from an intervention study. In S. Göpferich & I. Neumann (Eds.), *Developing and assessing academic and professional writing skills* (pp. 73–101). Frankfurt am Main: Peter Lang.

Kline, P. (2000). *The Handbook of Psychological Testing*. Routledge.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2020). *lmerTest: Tests in linear mixed effects models*. Retrieved 1 Aug 2021, from https://github.com/runehaubo/lmerTestR

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science*. https://strengejacke.github.io/sjPlot/

Mangiafico, S. (2016). *Summary and analysis of extension program evaluation in R (Version 1.11.1)*. New Brunswick, NJ: Rutgers Cooperative Extension

Milke, M. W., Upton, C., Koorey, G. F., O'Sullivan, A. D., & Comer, K. (2013). Improving the writing of engineering students through portfolios. In *Conference Proceedings of the 120th ASEE Annual Conference and Exposition*. American Society for Engineering Education, 1–12. Retrieved 1 Mar 2021, from https://www.asee.org/public/conferences/20/papers/8074/view

OECD. (2017). *International migration outlook 2017*. Paris: OECD Publishing.

Piepho, E. (2003). *Lernaktivierung im fremdsprachenunterricht: "Szenarien" in Theorie und Praxis*. Hannover: Schroedel.

Plakans, L. (2009). The role of reading strategies in integrated L2 writing tasks. *Journal of English for Academic Purposes*, *8*, 252–266

Risemberg, R. (1996). Reading-to-write: Self-regulated learning strategies when writing essays from sources. *Reading Research and Instruction, 35*, 365–383.

Robinson, D. W. (2000). Building consensus scoring on the scoring of students' writing: A comparison of teacher scores versus native informants' scores. *French Review*, *73*(4), 667–688

Roche, J. (2017). Herleitung von Grundlagen der handlungsorientierten Sprachvermittlung an beruflichen Schulen. In E. Terrasi-Haufe, & A. Börsel (Eds.), *Sprache und Sprachbildung in der beruflichen Bildung* (pp. 187–200). Münster & New York: Waxmann

Schneider, H., Becker-Mrotzek, M., Sturm, A., Jambor-Fahlen, S., Neugebauer, U., Efing, C., & Kernen, N. (2013). *Expertise zur Wirksamkeit von Sprachförderung*. Aarau & Köln: Pädagogische Hochschule FHNW, Universität zu Köln, Mercator Institut

SERI State Secretariat for Education, Research and Innovation. (2022). *Vocational and professional education and training in Switzerland – Facts and Figures*. Biel: Gassmann.

Shohamy, E. (1998). Critical language testing and beyond. *Studies in Educational Evaluation, 24*(4), 331–345.

Spivey, N. N., & King, J. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, *24*, 7–26

Sturm, A. (2017). Texte zu texten verfassen – vom schulischen Genre zum fachlichen Schreiben. *Leseforum. ch*, *2*, 2–15. Retrieved 1 Mar 2021, from https://www.leseforum.ch/sysModules/obxLeseforum/Artikel/603/2017_2_Sturm.pdf

Vandermeulen, N., De Maeyer, S., Van Steendam, E., Lesterhuis, M., van den Bergh, H., & Rijlaarsdam, G. (2020). Mapping synthesis writing in various levels of Dutch upper-secondary education: A national baseline study on text quality, writing process and students' perspectives on writing. *Pädagogische Studien*, *97*, 187–236

Watanabe, Y. (2001). Read-to-write tasks for the assessment of second language academic writing skills: investigating text features and rater reactions. Unpublished doctoral dissertation, The University of Hawaii

Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Ziegler, B., Balkenhol, A., Keimes, C., & Rexing, V. (2012). Diagnostik „funktionaler Lesekompetenz". In T. Tramm, S. Seeber & H.-H. Kremer (Eds.), *Funktionen und Erträge pädagogischer Diagnostik im wirtschafts- und berufspädagogischen Bereich*. bwp@ Ausgabe Nr. 22, 1–19. Retrieved 1 Mar 2021, from http://www.bwpat.de/ausgabe22/ziegler_etal_bwpat22.pdf

BBT Bundesamt für Berufsbildung und Technologie (Eds.) *Berufliche grundbildung: Rahmenlehrplan für den allgemeinbildenden Unterricht*. Bern:BBT

BBW Berufsbildungsschule Winterthur (Eds.) (2017). *Schullehrplan. Allgemeinbildender Unterricht für die 3-jährige und 4-jährige Grundbildung (EFZ)*. Winterthur: BBW Berufsbildungsschule Winterthur. Retrieved 1 Mar 2021, from https://www.bbw.ch/files/content/1%20Angebot/Berufsfachschule/Allgemeinbildung/BBW%20SLP%20ABU%208-2017.pdf

Federal Statistical Office (2021). Vocational education and training (VET) - Schools. Retrieved 21 Feb 2021, from https://www.bfs.admin.ch/bfs/en/home/statistics/education-science/pupils-students/upper-secondary/vocational-training.html

NAEP National Assessment of Educational Progress (2008). *NAEP scoring. Writing scoring specifications*. Retrieved 1 Mar 2021, https://nces.ed.gov/nationsreportcard/tdw/scoring/training_scorers_guide_writing.aspx

OECD (2016). *PISA 2018 draft analytical framework May 2016*. Paris: OECD Publishing. Retrieved 1 Mar 2021, from https://www.oecd.org/pisa/data/PISA-2018-draft-frameworks.pdf

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

SERI State Secretariat for Education, Research and Innovation (Eds.) (2014). Verordnung des SBFI über mindestvorschriften für die Allgemeinbildung in der beruflichen Grundbildung vom 27 April 2006. Retrieved 1 Mar 2021, from https://www.fedlex.admin.ch/eli/cc/2006/510/de?print=true

UNESCO Institute for Statistics (2011). *International standard classification of education ISCED 2011*. Retrieved 1 Mar 2021, from http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Liana Konstantinidou[1] · **Karin Madlener-Charpentier**[1] · **Aleksandra Opacic**[1] · **Curtis Gautschi**[1] · **Joachim Hoefele**[1]

Karin Madlener-Charpentier
karin.madlener-charpentier@zhaw.ch

Aleksandra Opacic
aleksandra.opacic@zhaw.ch

Curtis Gautschi
curtis.gautschi@zhaw.ch

Joachim Hoefele
joachim.hoefele@zhaw.ch

[1] School of Applied Linguistics, Institute of Language Competence, Zurich University of Applied Sciences, Theaterstrasse 17, 8401 Winterthur, Switzerland