Chapter 5

General discussion

Introduction

This dissertation focused on secondary vocational technical education, where students are trained to become professional practitioners. More specifically, the topic of this dissertation aligns with the need not only to train these secondary vocational technical students to become domain specialists within their field, but also to have them acquire a broader professional orientation. On the one hand, becoming a professional practitioner requires a thorough specialized knowledge base. On the other hand, it is of growing importance that these future technical professionals possess a range of more generalized skills. These are, for example, skills that would help them to be able to work collaboratively in increasingly complex teams and continually develop their knowledge. Thus, as mentioned in the General introduction, our focus was on a specific set of interpersonal and intrapersonal 21st-century skills. More concretely, we developed and evaluated support that aims to help secondary vocational technical students develop skills related to their collaboration and knowledge monitoring.

The support consisted of two digital tools: one to promote students' collaboration skills by fostering the development of relevant communication activities and the other to aid students' knowledge monitoring skills by stimulating students to externalize and evaluate their knowledge. These tools had in common that they incorporated elements that in theory foster reflection: *comparative feedback, reflection prompts*, and *interaction with peers.* Previous research indicated that through fostering reflection, these elements potentially promote students' knowledge and skills development (e.g., Kori et al., 2014; Radović, Firssova, Hummel, & Vermeulen, 2021). This dissertation reported three studies that explored the effect of the digital tools, which were complemented with teacher-coordinated classroom activities (i.e., specific instruction and plenary classroom discussion). Different implementations of each tool were studied and compared to a control group and each other. Each implementation contained some variation of the before-mentioned elements.

The tools were offered as part of an online learning environment in which the students actively learned about topics – electricity and electrical power transmission – connected with the curriculum of their program, by interacting with online labs operating on a smaller (i.e., electrical circuits) and larger (i.e., power networks) scale.

This final chapter summarizes the main findings of the three studies, provides an overarching discussion, highlights a few practical issues, and gives some suggestions for future research.

Main findings

Study 1 (Chapter 2) focused on improving students' collaboration skills by supporting communication activities that are seen as essential for effective collaboration (i.e., the RIDE rules; Saab et al., 2007) and explored the effect on students' collaborative behavior and knowledge acquisition. Just providing instruction about the RIDE rules to the students was not enough to induce an effect on either their collaborative behavior or their knowledge acquisition. Instruction combined with iterative use of the developed tool, which incorporated *comparative feedback* (i.e., students received an overview of scores derived from self- and peer assessment about the RIDE rules), *reflection prompts,* and *interaction with peers* (i.e., students were prompted to reflect together upon their collaboration and set goals for improvement), positively affected students' collaborative behavior (compared to both instruction alone and no instruction) and knowledge acquisition (compared to no instruction).

Studies 2 (Chapter 3) and 3 (Chapter 4) involved fostering students' knowledge monitoring skills by stimulating their externalization and evaluation of knowledge. In Study 2, a concept mapping tool with additional supportive features was used to enable students to externalize their knowledge. Neither the addition of *comparative feedback* (i.e., a combined concept map that mapped students' self-generated concept map onto an expert example concept map and highlighted differences and commonalities between the two) nor the use of *reflection prompts* (i.e., highly directed prompts that stimulated students to indicate missing information in their concept map) benefited students' knowledge acquisition over creating a concept map without this additional support.

For Study 2, we speculated that the lack of an effect could have been caused by the fact that the structure of students' concept maps was often very different from the expert example concept map, which could have made a comparison more difficult. It was conjectured that the potential of the comparative feedback could be enhanced when the student's concept map and the example concept map are as similar as possible in terms of structure. We reasoned that training students in creating concept maps and providing them with a more restricted concept mapping facility (in terms of predefined concepts and linking words) could help reduce the occurrence of structural misalignments between the expert and student concept maps. The fact that, contrary to our expectations, no effect of the reflection prompts was found could possibly be attributed to the relatively high level of specificity employed in these prompts (they were mainly "check-the-box items"), which might have prevented students from active reasoning (Wylie & Chi, 2014). Another potential factor inhibiting the active processing of the information could be the phrasing of the prompts. The prompts we used could possibly be improved: instead of focusing on possible gaps in their knowledge, asking students to reason about how their knowledge can be improved may foster their reflections (e.g., Dekker et al., 2013; Pitt & Norton, 2017). Optimizing the reflection prompts by formulating them more openly and phrasing them in a positive manner could potentially stimulate students' active processing.

The suggested changes described above were incorporated in the design of the intervention implemented in Study 3 (i.e., concept map training, more restricted concept mapping facility, and compared to Study 2, more open-ended and positively phrased reflection prompts). In addition, we argued that the effectiveness of individual processing of the reflection prompts is likely to depend on students' motivation (i.e., a product of their perceived ability to respond to the prompts and the perceived value of responding to the prompts). Therefore, we extended the intervention by incorporating a teacher-guided classroom discussion, thereby facilitating *interaction with peers*, before students (individually) responded to the reflection prompts. Despite the adjustments made, in Study 3 as well no effect of *comparative feedback* and *reflection prompts* was found on students' knowledge acquisition. However, the addition of the teacher-guided classroom discussion (i.e., facilitating *interaction with peers*) positively affected students' knowledge acquisition.

Type of support: effects of the employed design elements

Although Study 1 and Studies 2 and 3 differed in terms of the skills that were the target of the intervention, comparison based on the design elements incorporated in the support that was provided (i.e., *comparative feedback, reflection prompts,* and *interaction with peers*) may further contribute to our understanding of what type of support does or does not benefit secondary vocational technical students. The similarities between the studies in terms of these elements allow for speculation about what the outcomes of various studies have in common from an overarching perspective, which in turn provides input for discussing how support regarding fostering more generalized, trans-disciplinary skills, such as collaboration and knowledge monitoring, can be designed to benefit our target group.

Comparative feedback

Comparative feedback was incorporated in the support provided in all three studies. In Study 1, students received as part of the tool comparative feedback consisting of an overview of scores on the RIDE rules derived from self- and peer assessment. This overview included the students' own scores, average scores of peers, and the group's average score. In Studies 2 and 3 the comparative feedback consisted of a combined concept map, which mapped each student's self-generated concept map onto an expert example concept map and highlighted differences and commonalities between the two.

In Study 1, students in the condition in which the tool was part of the intervention outperformed students in other conditions regarding their collaborative behavior and domain knowledge acquisition. In Study 2, consultation frequency of the combined concept map was found to predict knowledge gain. Although descriptive data for the average knowledge gains suggested that students who received the combined concept map demonstrated higher learning gains than students who did not receive a combined concept map, statistical comparison of learning gains did not result in significant differences between conditions. The inclusion of a concept map training and a more restricted concept mapping facility in Study 3 was expected to optimize the comparative feedback and reinforce the observed trend in learning gains. However, the opposite turned out to be the case, as the trend was no longer evident in Study 3; learning gains of students were found to be comparable, regardless of whether they received the combined concept map or not.

When considering differences between the way the comparative feedback was offered in Study 1 compared to Studies 2 and 3, two things in particular stand out. *First*, in Study 1 student performance was contrasted with the performance of peers (i.e., social comparison feedback), whereas in Studies 2 and 3 comparison with an expert example was offered (i.e., norm-referenced or criterion-based feedback). Results of related research are, however, inconclusive when it comes to which type of comparative feedback is in general most beneficial for students (e.g., Labuhn, Zimmerman, & Hasselhorn, 2010). Comparison with peers, on the one hand, can trigger competition (e.g., Nebel, Beege, Schneider, & Rey, 2016; Nebel, Schneider, Beege, & Rey, 2017), which can potentially positively affect engagement and learning, especially when students compare themselves with others who perform better (Dijkstra, Kuyper, van der Werf, Buunk, & van der Zee, 2008). However, comparison with peers is not always found to positively affect engagement and learning; it depends on individual and contextual factors (Schneider, Beege, Nebel, Schnaubert, & Rey, 2021). In some situations, social comparison can even have a negative effect on students' self-concept, although, according to Dijkstra et al. (2008), this does not outweigh possible positive effects.

Comparison with input from an expert, on the other hand, might provide a more objective norm, which students could perceive as more reliable than input from peers (e.g., Tsui & Ng, 2000). However, in a study by Kollöffel and de Jong (2016), who targeted a population similar to ours, students who were provided with social comparison feedback considerably outperformed students who received criterion-based feedback. Although students in the study by Tsui and Ng (2000) favored teacher comments over those of peers, results indicated that the effectiveness of peer feedback was enhanced when students could subsequently explain and discuss their feedback orally. Our students in Study 1 had a similar opportunity in the phase that followed the comparative feedback phase, in which they discussed what went (not so) well and what, how, and by whom things could be improved, based on the feedback (see also next section). This might have strengthened the possible effect of the comparative feedback in this study.

Second, in Study 1, unlike Studies 2 and 3, students themselves have actively contributed to the comparative feedback by assessing their own performance and the performances of their peers. This possibly contributed to the effect that was found. Related research has shown that students can benefit from both assessing their own performance and assessing the performance of their peers (e.g., Dmoshinskaia, Gijlers, & de Jong, 2021; Ion, Sánchez Martí, & Agud Morell, 2019; Li, Xiong, Hunter, Guo, & Tywoniw, 2020); this assessment encourages students to think actively about their own or others' performances, for example, in relation to set criteria or to each other, which can deepen their learning and increase their engagement. So, it may well be that not only the (social) comparative feedback, but particularly students' active *input* (i.e., the assessments) associated with the comparative feedback was an essential aspect in this situation.

All in all, we may wonder about the effect of comparative feedback, as employed in the current studies, as separate element. We can speculate that the social comparison variant (Study 1) may be preferable over the norm-referenced version (Studies 2 and 3). Although the overall intervention in Study 1, of which the comparative feedback was part, was found to be effective, we have reasons to assume that prior activities (i.e., active input associated with the comparative feedback) and subsequent activities (i.e., discussion of the comparative feedback) also contributed to the observed effect.

Reflection prompts and interaction with peers

Reflection prompts were another element of support in all three studies. However, whether these prompts were *individually* processed or in *interaction with peers* (i.e., *jointly* processed) differed.

In Study 1, prompts were *jointly* processed: students were prompted to discuss what went (not so) well regarding their collaboration and to set goals for improving their collaboration. In Studies 2 and 3, the reflection prompts were designed to stimulate processing of differences presented in the combined concept map. In Study 2 all students responded to the prompts *individually*, whereas in Study 3 the prompts were *individually* processed in one condition, and *jointly* processed in another condition, during a brief teacher-guided classroom discussion before students (individually) formulated their responses to the prompts.

As described above, students in Study 1 who used the reflective support, which included the *jointly* processed reflection prompts, as part of the intervention, outperformed students in other conditions. In Study 3, students who *jointly* processed the prompts during the classroom discussion showed significantly higher learning gains compared to students who only processed them *individually.* Furthermore, students who only processed the prompts *individually* did not perform better, on average, than students who received no prompts at all. These findings were in line with the results of Study 2, in which no differences were found regarding students' learning gains with respect to whether they received (*individually* processed) prompts or not.

When considering the type of prompts that were employed across the various studies, one way to differentiate between them is to categorize the prompts based on the amount of structure and direction they provide, as distinguished by Wylie and Chi (2014): from open-ended (i.e., no limits on the type of answer) to menu-based (i.e., answers selected from a list). Regarding this characterization, the *reflection prompts* in Study 1 were formulated in an open-ended and generic way so that they would allow for discussion among the students. In Study 2 it was a deliberate choice to present the prompts in a more specific and directive way (i.e., mainly "check-the-box items"), not only because related research has argued that this type of prompts (compared to more open-ended prompts) are preferable when comparing and contrasting information sources (Gadgil et al., 2012; Wylie & Chi, 2014), but also because it was expected to suit our students best (i.e., many secondary vocational students struggle with written language and their language skills generally tend to fall behind those of peers from other educational levels; Groot, Houtkoop, Steehouder, & Buisman, 2015). However, with the absence of an effect in Study 2, our presumption was that the focus on missing information (instead of focusing on improvement) and the highly directive formulation may have inhibited students' critical thinking and reasoning; we suspected that the directive prompts led to just checking the boxes rather than active processing. This made us design prompts that were phrased more openly and in a more positive manner in Study 3. Either way, no effect was found of the *individually* processed prompts in either way of formulating them.

Although there is no uniform conclusion regarding which type of prompts is generally most recommended and their effectiveness may be context-dependent (Belland, 2014; Kori et al., 2014), the fact that (*individually* processed) prompts did not have a direct effect in our situation may be seen as surprising. In a related context, solving ill-structured problems, strong effects were found of providing students with prompts (Ge & Land, 2003). A difference may be that the prompts in the Ge and Land study focused on the structure of the problem itself, whereas our prompts had a more general, domain-independent character. An example prompt from the Ge and Land study was, "Have I discussed both the technical components and the issues with use, for example, usability and effectiveness?". In another study, Ge, Chen, and Davis (2005), supporting instructional design students, gave prompts such as: "Is this design project in response to a problem or a need?", "Is there really a need for Web-based instruction?", "How do you know?", "How can you determine it?". In our situation, a typical prompt was: "When comparing both concept maps, how would you improve your concept map and why?". Although findings vary with respect to the effect of more context-specific prompts (e.g., Belland, 2014; McNeill & Krajcik, 2009), it may be that, especially for our type of students, these more context-independent prompts gave them too little grip to perform well-informed reflection.

What could substantiate this line of reasoning is that during the classroom discussion in Study 3, the situation in which an effect was found, the teachers stimulated students to elaborate on the content presented in the concept maps without themselves providing domain-related information, for example: "Why would you change/add this information?", "Can you explain why this information is relevant/important?", "Which concepts/relations would you use to represent this information?". These prompts can be regarded as more context-specific, which might have been beneficial to our students.

Besides elaborating on the predefined prompts, another potential reason why the classroom discussions contributed to the effect found in Study 3 is that they might have featured a more adaptive situation better tailored to students' individual needs, in addition to the predefined prompts. Prompts in general have not been found to be equally beneficial for all students and not every student makes use of the opportunity to reflect by means of the provided prompts (e.g., Davis, 2003; Kim & Lee, 2002; Kori et al., 2014). The classroom discussion could have played a positive role in this respect: from the results of Study 3, we could see that students who *jointly* processed the reflection prompts gave both more and relatively higher quality answers when compared to students who processed the prompts individually. When also considering the effectiveness of the tool in Study 1, where students also *jointly* processed the reflection prompts, it is most likely that our students in general benefit from *interaction with peers* when processing prompts.

Overall, taking into account the possibility that there is no one-size-fits-all design when it comes to reflection prompts and our rather diverse group of students, processing the prompts in *interaction with peers* might have made them feel more capable and motivated to perform the suggested actions. For example, in Studies 2 and 3 students needed to be able to actively process differences between their own concept map and the expert concept map. Although this may seem straightforward at first sight, maybe for our type of students to perform this on their own was a bridge too far. Performing this together with peers or in a classroom discussion with the support of the teacher may have given them just the extra information and motivation needed to perform the actions (see, e.g., Radović et al., 2021; Trede & Jackson, 2019).

Based on the overview above, it can be conjectured that support (i.e., *comparative feedback, reflection prompts)* that is found to be effective in many other educational contexts does not necessarily benefit secondary vocational technical students' behavior or learning. It stands out that these students benefited most when *interaction with peers* was part of the support. Overall, it seems that, for this group of students, interaction with peers is an essential aspect of successful learning.

Putting things in perspective

Research within the context of secondary vocational education as conducted in the current dissertation is relatively scarce, even though the student population is quite large. Over 500,000 students followed a secondary vocational educational track in the Netherlands in 2020 (Centraal Bureau voor de Statistiek, 2021b) and almost 30% of the Dutch population has completed such a track as their highest educational level (Centraal Bureau voor de Statistiek, 2021a), and their needs are vastly different from other student populations (e.g., Groot et al., 2015). Hence, research with secondary vocational students can be considered very valuable. However, it should be noted that this research context involves some practical challenges.

Contextual factors

Students in our population have an "obligation to qualify" (in Dutch: kwalificatieplicht). This implies that they have a certain degree of freedom when it comes to obligatory school attendance (see also Chapter 1). As a result, the students' decision about whether or not to attend a particular lesson seems often related to their perceptions of the usefulness or relevance of a particular lesson; they should either see the direct added value ("Do I like to do this?"; "Is there something else to do that I like better or that is of higher relevance to me?") or there should be an incentive present ("Do I have to do this?"; "Do I need this?"). The present studies were carried out in different classes at four different schools, which means that although the content of the lessons was established in accordance with teachers and generally fitted within the curriculum of the students, the content of the experimental lessons was not exactly tailored to the various classes. Thereby, due to the experimental set-up within each school (i.e., students in particular conditions might have an advantage over students in other conditions), it was not possible to connect a grade with the posttests that were used. Not being able to fully tailor the subject matter and the lack of a grade as an incentive may have contributed to reducing perceived usefulness and relevance for students, which might have affected two aspects in the studies conducted: attrition and learning gains.

The *attrition* in all three studies can be considered relatively high (43-56%). Given the degree of freedom regarding obligatory school attendance that comes with the "obligation to qualify", irregular school attendance is not exceptional. However, this is something that should be anticipated when conducting multi-session experimental research within this context. Even though it would have been preferable to include all initial participants in the final analysis, we argue that this relatively high attrition did not resulted in a substantially different sample or in a final sample that was substantially different from the group of students who dropped out. Generally, it tended to be the case that different students missed one or more random sessions and therefore were excluded from further analyses, while it rarely occurred that students were present during the first session and subsequently absent from all remaining sessions.

Considering ecological validity in relation to attrition, the current studies were conducted in a realistic school setting as much as possible. This implies, among other things, that the experimental sessions were scheduled during regular school hours and took place as much as possible during already-scheduled lessons, with the presence of a teacher. Our impression was that, especially with our students, the presence of a teacher, preferably their own, and leaving the existing classes intact whenever possible is essential. Not doing so would likely have further increased the attrition rate and possibly could have affected the level of classroom disorder.

Regarding *learning gains* in the current studies, although overall posttest scores can be considered relatively low (mean of 6.39 out of 25 points in Study 1; means of 8.61 and 7.14 out of 22 points in Studies 2 and 3, respectively), students in all studies and conditions, on average, learned significantly from pretest to posttest (average learning gain of 1.89 points in Study 1, 2.19 points in Study 2, and 1.87 points in Study 3). This indicates that students, on average, did learn, but had not yet fully mastered the domains addressed by the various studies. Despite the relatively low posttest scores, students' significant learning gains that were found in all conditions can be regarded as a satisfactory result when taking into account the absence of an incentive, which might have influenced students' motivation to complete the test to the best of their ability. Additionally, it should be noted that the time students spent in the online learning environment (i.e., opportunity to learn) can be considered relatively brief (depending on the study, 75 to 90 minutes around 2 or 3 times). Typically, in secondary vocational education learning materials are offered repeatedly in different ways, over a longer period of time, and students are usually given the opportunity to study the learning material prior to the test.

Type of education

As in many educational settings, in secondary vocational education a shift from teacher-centered to more student-centered instruction is taking place (Christoffels & Baay, 2016b), meaning that students are given a higher degree of freedom and are expected to take the initiative themselves when it comes to their learning, compared to a situation in which the teacher is mostly taking the lead in students' learning (e.g., by transferring knowledge and/or providing a step-by-step procedure as to what students should be able to do; Trilling & Fadel, 2009). However, this transition is typically limited to particular moments in most secondary vocational settings, especially in the first school year (e.g., in projects where students are gradually given more freedom and are expected to take more responsibility). Although a certain degree of autonomy is expected from the students, it can be seen that these (especially first-year) students are generally used to being guided and supervised by their teacher (Elffers, 2014). In addition, we have also noticed that the teacher played an important role when it came to student attendance and how students dealt with the learning materials. We can argue that the teacher could act as a kind of external incentive. On top of the fact that the research sessions already took place during regular classes and that the domain-related themes were connected to the students' curriculum, the presence of a teacher in addition to attendance of a researcher could contribute to the research sessions being taken more seriously and to the feeling that the sessions are part of students' regular educational practices.

Taking the above into consideration during experimental research in a similar context may raise some practical challenges, as the schedules of both students and teachers in secondary vocational education are generally quite tight and teacher availability is rather limited (i.e., a set-up with several experimental conditions, as in our case, can require more teachers than would have been the case in a regular setting). Taking into account students' schedules and assuring teacher attendance might seem to be small factors, but they can have a major impact.

Looking ahead

Considering the most prominent overarching finding, namely, that our students benefited most when *interaction with peers* was part of the support, future research can build on the current findings and further our understanding by getting a better grip on this element and investigating *how* interaction with peers is beneficial to our target group. Suggestions for possible future directions are discussed below.

In Study 1, *interaction with peers* was intertwined throughout the intervention and interacted with the other elements that were used (i.e., *comparative feedback* and *reflection prompts*). Although the other single elements were not found to be of added value in an individual setting, as deployed in Studies 2 and 3, we have also reasoned that the elements might have strengthened each other in Study 3, which is in line with Belland's (2014) suggestion to combine scaffolding types. Insight into the unique contribution of these various elements in a similar setting to Study 1 can provide insight into whether all elements are essential for the currently observed effectiveness and whether and how possible additional support should be offered.

To investigate the way *interaction with peers* impacts the results, follow-up research could study, for example, the quality of the interaction during the classroom discussion in a setting similar to Study 3. Quality of group interaction can impact the quality of students' (joint) reflection (Daniel & Jordan, 2017). Therefore, capturing what students contribute to the discussion and in what way and investigating how this relates to the quality of their reflection can help optimize the organization and support of *interaction with peers*. In addition, and related to the point regarding the role of the teacher as discussed above, it would be relevant as a follow-up to investigate the added value of the teacher during such an interaction: was it mainly the *interaction with peers* that caused the effect in our situation or was the teacher a significant factor in guiding and scaffolding the interaction? The fact that no teacher was involved in the intervention in Study 1 suggests that *interaction with peers* alone may also work. Insight into this point can improve the effectiveness and efficiency of future support.

Reflection prompts that were *jointly* processed (i.e., in *interaction with peers*) were found to be effective in our studies. However, students' written responses to those prompts (i.e., in Studies 1 and 3), still indicate room for improvement. Although we know, based on the results of Study 3, that certain written responses predicted students' learning gain, an interesting direction to explore is whether only discussing responses to prompts, instead of (also) writing them down, would be effective as well. This would not only make the support more efficient, but probably also fit our target group even better, among whom there are students who struggle with written language, or who are not always inclined to write down their response to such prompts.

As a final suggestion, if we were to implement or extend *interaction with peers* to a greater extent in the design of knowledge monitoring support in a setting similar to Studies 2 and 3, instead of the currently employed norm-referenced feedback, utilizing social comparison feedback (as was done in Study 1) could be a promising approach, for example, by having students compare their concept map with those of (a couple of) peers, combined with a subsequent discussion.