Metacognition and Learning
https://doi.org/10.1007/s11409-022-09299-7

Supporting reflection to improve learning from self-generated concept maps

Elise H. Eshuis[1] · **Judith ter Vrugte**[1] · **Ton de Jong**[1]

Received: 26 March 2021 / Accepted: 30 March 2022

Abstract
Creating concept maps is considered to be a powerful means for learning. It requires students to systematically organize and integrate their knowledge, which can foster meaningful learning. However, students scarcely spontaneously engage in the (meta)cognitive processes necessary for effective knowledge integration, such as reflection, which can hamper the effectiveness of concept mapping for learning. This study explores the effect of additional support that stimulates reflection by means of expert examples, reflection prompts, and classroom discussion. First-year technical vocational students ($N = 144$, $M_{age} = 17.5$ years) studied electricity-related topics in an online learning environment and created concept maps about their knowledge. Students' concept maps were supplemented with either (1) no support (control condition), (2) an expert example (a combined concept map, containing their own and an expert example concept map, with differences highlighted), (3) an expert example and reflection prompts students had to process individually, (4) an expert example and reflection prompts students had to process individually after a teacher-guided classroom discussion. Students in the classroom discussion condition showed higher learning gains compared to all other conditions. This can be explained by the quality of their reflection, which proved to be a significant predictor of learning gain.

Keywords Concept maps · Expert examples · Question prompts · Classroom discussion · Reflection

Introduction

Concept mapping is an often adopted approach in education as it is seen as a successful means to foster meaningful learning (Schroeder et al., 2018). Meaningful learning takes place when newly acquired knowledge is deliberately integrated into already existing knowledge structures (Novak, 2002). This anchors the information in a network of understanding, which makes it more deeply rooted and fosters retention (Romero et al.,

✉ Elise H. Eshuis
e.h.eshuis@utwente.nl

[1] Department of Instructional Technology, Faculty of Behavioural, Management, and Social Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Published online: 04 May 2022

2017). Despite the positive results of concept mapping for learning that have generally been found (Nesbit & Adesope, 2006; Schroeder et al., 2018), many studies have reported that merely having students create concept maps does not automatically harness the full instructional potential of such mapping (e.g., Hilbert & Renkl, 2009; Kinchin, 2014; Lim et al., 2009). To foster students' processing of the concept maps, and consequently their knowledge integration, students should engage in both cognitive and metacognitive processing.

Cognitive processing refers to the organization and elaboration of knowledge (Berthold et al., 2007). Metacognitive processing concerns the ability to evaluate one's cognitive performance and regulate one's cognitive activities (Kauffman, 2004), which, among others, requires the ability to monitor one's knowledge (Berthold et al., 2007). To facilitate learning from concept maps, students should use a concept map not only to externalize their knowledge, by identifying main idea's, structuring their knowledge, and creating links between (new) information (and prior knowledge) (i.e., cognitive processing; Berthold et al., 2007), but also carefully select what to focus on, monitor their knowledge by identifying and acknowledging what they do and do not know, think about what this means for them given the situation, and plan for future actions (i.e., metacognitive processing; Berthold et al., 2007; Winne, 2021). Both metacognitive and cognitive processes can take place in a process of reflection, or as Davis (2003, p. 92) puts it: "I use the term reflection to refer to both metacognition and sense-making. For example, reflection can focus on goals or on one's own thinking (metacognition), or on content itself (sense-making)".

In the context of concept mapping reflection can be triggered by providing students with a frame of reference to which they can compare and contrast their knowledge (Schwendimann & Linn, 2016), like an example concept map. Having said that, research shows that students are very reluctant in employing (meta)cognitive processes related to comparing and contrasting their knowledge on their own and if they do, they often employ them suboptimally (Renkl, 2014b). Hence, more guidance is likely neccesary. A common aid used to support students' reflection is reflection prompts (e.g., Chi et al., 1994; Davis, 2003). In this paper we explain why an expert example could foster students' learning with concept maps and how reflection prompts may influence the effectiveness of students' learning with concept maps. But we also posit that research shows that the effectiveness of reflection prompts is highly dependent on student characteristics (e.g., prior knowledge, (meta) cognitive skills) and their motivation (Bannert & Mengelkamp, 2013) and that discussion between peers may compensate for the (negative) impact of such individual characteristics (Kim & Lee, 2002; Schwendimann & Linn, 2016; Xun & Land, 2004). Therefore, the current study not only explores the benefit of an example and the added benefit of reflection prompts, but also whether individual or joint processing of these prompts influences their effectiveness.

Supporting knowledge integration by providing students with an example

Creating a concept map entails interconnecting and structuring one's knowledge by forming propositions (i.e., labelled, logical connections between two or more pieces of conceptual information) to create a meaningful schematic representation of this knowledge (Nesbit & Adesope, 2006; Schroeder et al., 2018). When students create a concept map, they ideally engage in knowledge integration, which requires them to actively expand their ideas, distinguish between ideas, make links between them, and identify weaknesses in their current knowledge (Linn, 1995, as cited in Davis, 2003, p. 93). This identification of weaknesses is essential; Davis (2003) argued that when students are able to identify

weaknesses in their knowledge, it is more likely that they will engage in the other processes of knowledge integration. Providing students with an (expert) example can help them to identify these potential weaknesses (Schwendimann & Linn, 2016).

In the first place, students might learn from just examining an example (ter Vrugte et al., 2017; van Merriënboer & Kester, 2005). On top of that, an (expert) example can also serve as a source against which students can compare their own knowledge. Using an example concept map as a comparison source can help students, not only identifying gaps and misunderstandings in their knowledge (Schwendimann & Linn, 2016; Wu et al., 2012), but also stimulate them to think more deeply about and reflect upon their knowledge (Lin et al., 1999). Within the context of creating concept maps, a comparison source can be provided in the form of an example concept map. In a study by Eshuis et al. (2022), students had to examine a combined concept map, containing their self-generated concept map and an expert example, with the differences between the two highlighted to draw their attention towards possibly interesting points to focus on. Results of this study indicated that the more often students consulted the combined concept map, the higher their domain-related learning gains. However, average learning gains were not found to be significantly higher for students who were offered the combined concept map compared to students who received no combined concept map. This suggests that providing students with an example in combination with their self-generated concept map can contribute to students' knowledge, but its effectiveness likely depends on the extent to which students are able or see the need to actively process and make sense of the information provided.

To benefit from an example concept map as comparison source, students should decide whether and what information to integrate into their existing knowledge structure by comparing, contrasting, evaluating, and reflecting (Kao et al., 2008). Here students need to take into account contextual factors. Although these processes can be triggered just by offering the example, many students do not spontaneously engage in such processes (Bannert, 2006; Renkl, 1997, 2014a). This might be due to either a deficiency in students' prerequisite knowledge or students' passivity in processing (Renkl, 2014b); students might not have sufficient knowledge to properly perform the necessary processes and make use of the example, they might not know how to do so, or might not engage on the level necessary for substantive change. Prompting can stimulate and help students to actively engage in the (meta)cognitive processes necessary to support knowledge integration (Davis, 2003; Renkl & Schwonke, 2013; Renkl, 2014b).

Reflection prompts to overcome suboptimal processing

Reflection prompts are commonly used to foster students' (meta)cognitive processing; not only can prompts activate them to engage in particular processes, but prompts can also scaffold these processes by indicating what to do or where to focus on (i.e., providing support for how to perform a task) (Bannert, 2006; Bannert & Mengelkamp, 2013; Davis, 2003; Kori et al., 2014; Lin & Lehman, 1999). In the context of concept mapping, prompting students to critically evaluate their own and an example concept map has been shown to be helpful in stimulating them to reflect upon their knowledge and also to revise their knowledge structure when needed (Schwendimann & Linn, 2016).

Although providing students with reflection prompts ideally stimulates them to actively engage in the intended processes, this may still not be enough to benefit their knowledge acquisition (Eshuis et al., 2022). Research has demonstrated the intricate balance between students' ability and the level of structure that prompts offer. While the level of structure provided by the

prompt is ideally kept to a minimum to foster deep processing through active engagement (ter Vrugte et al., 2015; Wylie & Chi, 2014), research findings have demonstrated that the optimal level of structure calls for careful tailoring to accommodate differences in student characteristics; offering too little structure puts students at risk of not processing the prompts at the desired level, while over-structuring the prompts increases the risk of unreflective processing and/or low engagement (as also discussed by ter Vrugte & de Jong, 2017; Wylie & Chi, 2014).

Although modifying the structure of the prompt can affect the extent to which students are triggered to engage in active processing, there is reason to assume that this may not be adequate to address individual student characteristics and needs. Research has shown that when students are offered prompts individually, they are not always motivated to use them (Belland et al., 2013). Students often simply ignore them, or provide shallow or ill-considered answers (Ge & Land, 2003; Lin et al., 1999; ter Vrugte et al., 2015; Xun & Land, 2004). The unique affordances of discussion among peers might help overcome these motivational issues (e.g., Ge & Land, 2003; Kim & Lee, 2002).

Classroom discussion to motivate students

Discussion among peers can promote student motivation because it may enhance their expectancy for success (i.e., students' expectations that they'll be able to succeed at a task) and their perceived subjective task value (i.e., students' perception of the usefulness of a task); two key factors of motivation, also addressed in expectancy-value theory of motivation (Wigfield & Eccles, 2000). More specific to the situation addressed in the current study, joint processing and discussion of information presented in an example can prepare students for subsequent reflection prompts. Discussion stimulates students to clarify, correct, or reorganize their knowledge when they come across new or conflicting information, or to justify their own insights to others (Ge & Land, 2003; Kao et al., 2008; Kim & Lee, 2002; Lin et al., 1999). Comparing and contrasting their ideas with peers, can help them to engage in deeper thinking than they would do individually. This enriches the (meta)cognitive processes necessary for successful knowledge integration (Lin et al., 1999). Consequently, students are more likely to feel equipped to successfully respond to the reflection prompts that require similar levels of (meta)cognitive processing and thus might feel more motivated to do so. Moreover, peers can serve as role models when contributing to the discussion. This can impact students' perceived task value (Fan, 2011) and consequently also contribute to students' motivation.

Classroom discussions about concept maps can act as a catalyst for the desired peer interactions (Schwendimann & Linn, 2016). Researchers recommend that classroom discussions are guided by a teacher (Lin et al., 1999; Schwendimann & Linn, 2016). Teachers can serve as an external facilitator to motivate students to engage in this reflective activity, increase the extent to which students value reflection, and can guide students' discussions by posing relevant questions (Belland et al., 2013; Ge & Land, 2003; Lin et al., 1999). To maintain the student-centered nature of the task, rather than transferring knowledge to their students, teachers can facilitate students' learning process by asking questions that stimulate the students to reflect on their knowledge (Hmelo-Silver & Barrows, 2006), similar to what the previously discussed reflection prompts would do. Beyond just providing reflection prompts, a teacher's assistance during discussions allows for dynamic and adaptive guidance, meaning that teachers are able to promptly read, interpret, and respond to the students' needs (Belland et al., 2013). Meeting different needs of individual students more closely, can in turn affect the way they value the task at hand (e.g., Fan, 2011; Wigfield & Eccles, 2000).

Current study

The current study set out to explore how to increase the effectiveness of creating concept maps for learning through facilitating students' knowledge integration, by triggering and supporting reflection, and thereby optimizing their knowledge acquisition. The study employed a digital tool that enabled students to combine their own, self-generated concept map with an expert example concept map as comparison source for their knowledge. The added value of such an example is likely to depend on how students process it. The effectiveness of providing an example depends on whether students successfully process the information by comparing and contrasting it with their own knowledge. Hence, to further stimulate students to examine the combined concept map critically, the tool was supplemented with reflection prompts. In the current study, it was hypothesized that individual processing of these reflection prompts depends on students' motivation (i.e., their perceived ability and value of completing the task) and that students might benefit from jointly processing the prompts by means of a classroom discussion, prior to responding to the prompts individually.

In this study, secondary vocational students worked in an online learning environment in which they independently and individually learned about topics related to electricity and electric power transmission and created concept maps about their knowledge. The learning environment was the same for all students, but differed in terms of the offered support related to their concept mapping task. The different modes of support are inseparable and build upon each other, which is also translated in the cumulative design of the employed conditions. To investigate the effect of the interventions on students' knowledge acquisition, students' concept maps were supplemented with either (1) no support (control condition), (2) an example concept map (a combined concept map containing their own and an expert example concept map, with differences being highlighted), (3) an example concept map and reflection prompts that students had to answer individually, (4) an example concept map and reflection prompts that students individually answered after a teacher-guided classroom discussion (see Table 1). The conditions were identical in terms of learning material.

Although it is plausible that offering an example alongside students' self-generated concept maps can have added value, it is interesting to investigate whether this translates to the current target group (i.e., secondary vocational students). As these students generally have difficulty verbalizing their knowledge (Slaats et al., 1999), it might well be that the example alone (as embedded in the combined concept map) would be insufficient to trigger the (meta)cognitive processes needed to benefit from it. The expectation was, therefore, that additional support could be of added value to help these students to engage in and carry out these processes. It was therefore expected that the addition of reflection prompts

Table 1 Overview of conditions

Conditions	Treatment
Control	Concept mapping
Example	Concept mapping; Combined concept map (containing a student's concept map and an example concept map)
Example + prompts	Concept mapping; Combined concept map; Individual reflection prompts
Example + discussion + prompts	Concept mapping; Combined concept map; Classroom discussion; Individual reflection prompts

would enhance the effect of the combined concept map. It was also hypothesized that the addition of a teacher-guided classroom discussion would lead to deeper reflection and, consequently, higher learning gains compared to solely individual reflection.

Method

Participants and design

An initial total of 253 secondary vocational education students (246 males, 7 females) participated in this study. Participants were first-year students from 11 classes divided over four schools for secondary vocational education (in Dutch: MBO) in the Netherlands. These schools offer four-year programs in which they prepare students for a vocational career (e.g., as a hairdresser, car mechanic, electrical engineer). All participants in the current study were enrolled in technical training programs that include electrical engineering as a fundamental part of the curriculum. The majority of students enrolled in these types of studies are male, therefore this sample mainly consisted of male participants.

This study utilized a pretest (session one) – intervention (sessions two–three) – posttest (session four) design. Students were assigned to one of the four conditions as presented in Table 1. After careful consideration of practical concerns (e.g., the classroom discussion was a whole-class activity and could affect peers present in the same room, while the other interventions were covert and had no risk of affecting peers not involved in that condition) and theoretical concerns (e.g., familiarity with peers affects the dynamics of a discussion; Janssen et al., 2009), a (partial) quasi-experimental approach was opted. This means that where possible an experimental approach was adopted and an exception was made for the condition with discussion. Each of the four participating schools were represented in all conditions. For the example + discussion + prompts condition one intact class was randomly selected from each school. Students within each of the remaining seven classes were ranked within their class on their average pretest score and alternately assigned to one of the three other conditions (the control condition, the example condition, or the example + prompts condition). Students who did not complete the pretest were randomly distributed over the three conditions. This resulted in the following distribution of students among the four conditions: 57 in the control condition, 57 in the example condition, 56 in the example + prompts condition, and 83 in the example + discussion + prompts condition.

Criteria for inclusion of participants' data in the final sample were based on students' attendance: a total of 109 students were excluded from the analysis because they missed one of the sessions (i.e., 44.1% missed one session; 34.9% missed two sessions; 9.2% missed three sessions; 11.9% missed four sessions). Therefore, the final sample consisted of 144 students (140 males, 4 females) with a mean age of 17.5 years ($SD = 0.93$): 34 in the control condition, 36 in the example condition, 27 in the example + prompts condition, and 47 in the example + discussion + prompts condition.

At first glance, this attrition rate might be interpreted as relatively high. However, it should be taken into account that students who participated in this study have generally passed the age at which they are obliged to attend school in the Netherlands. Instead, these students have an 'obligation to qualify' (in Dutch: kwalificatieplicht). This is an extended form of compulsory schooling, yet students have more freedom when it comes to school attendance, which consequently implies that irregular attendance is not an exception (Onderwijsconsument, 2019).

Online learning environments

The online learning environment of the current study consisted of two (in terms of domain content) successive environments, similar to the learning environments designed and used by Eshuis et al. (2022); during session two, students individually worked in the first learning environment, during session three in the second learning environment. Both contained the same two online labs, a series of assignments, and instructive multimedia material related to the topic of electricity and electric power transmission; they differed in the complexity of the domain principles offered. They were designed with the Go-Lab ecosystem (de Jong et al., 2021) and included nine assignments each. Both environments were structured by means of tabs at the top of the screen (see Fig. 1). The first tab opened an introduction in which the purpose and use of the learning environment were briefly explained. The next tab opened the first assignment, in which students had to depict their prior domain knowledge in a concept map by using the provided concept mapping tool (see next section). To help students determine relevant prior knowledge, they were given a set of general learning goals for each learning environment.

The next seven tabs in each of the learning environments opened a series of assignments that revolved around two labs: the Electricity Lab and the Electric Power Transmission Lab. In the Electricity Lab (see Fig. 1), students could create electrical circuits based on direct or alternating current, perform measurements on them, and view measurement outcomes. In the Electric Power Transmission Lab (see Fig. 2), students could design a transmission network by choosing different power plants and cities, and by varying different components within the network (e.g., properties of the power line and the voltage). The assignments in the first environment addressed basic principles of electricity (e.g., current and voltage) and basic elements of electric power transmission (e.g., efficiency,

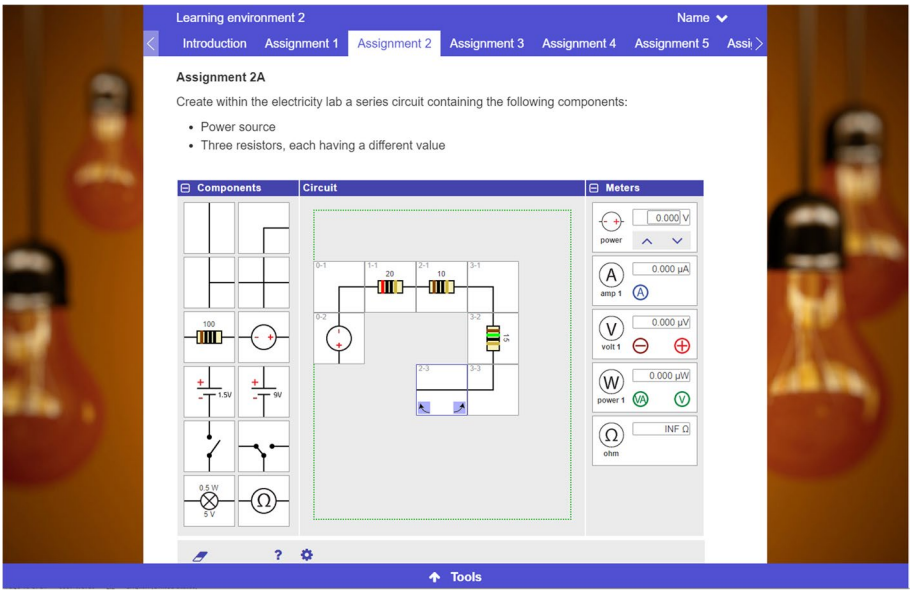

Fig. 1 Screenshot of the online learning environment (translated from Dutch) with the Electricity Lab in the center

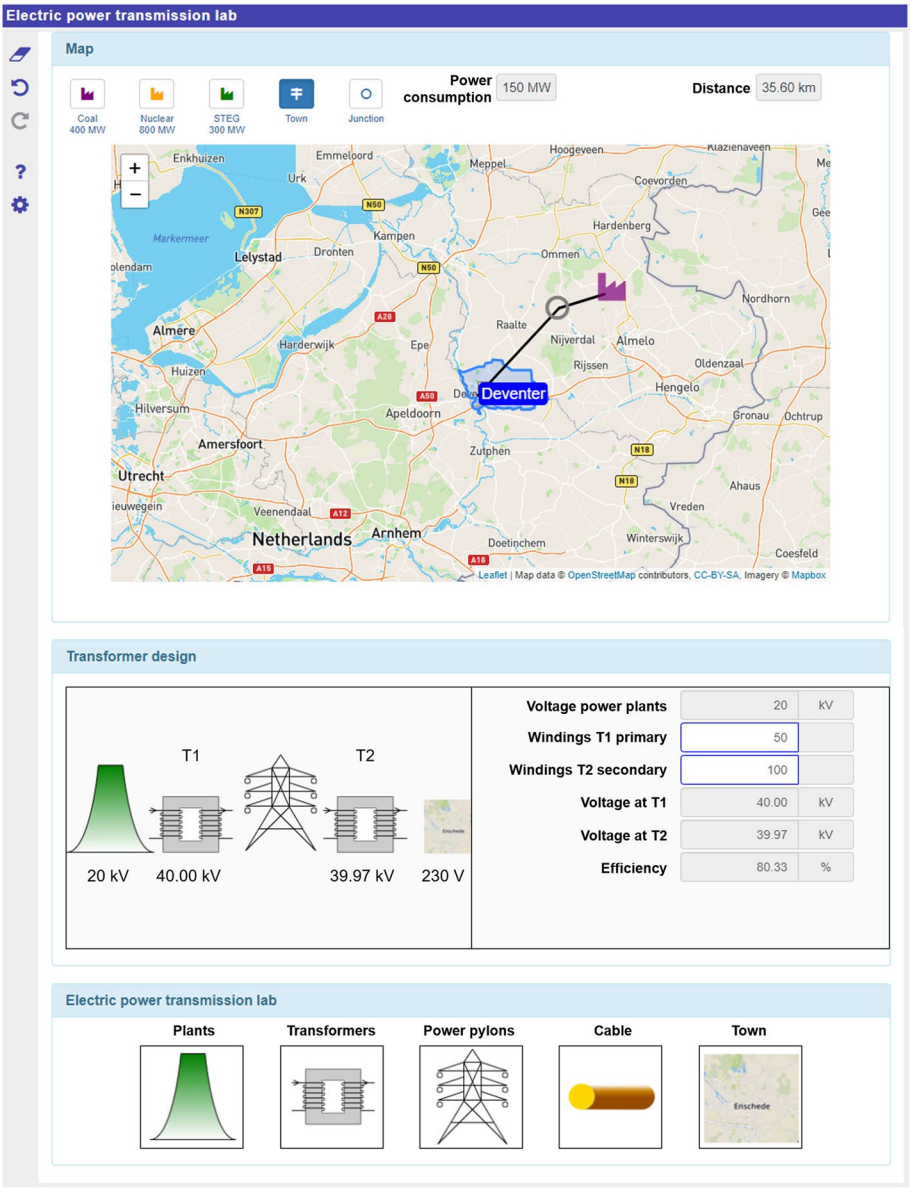

Fig. 2 The Electric Power Transmission Lab (translated from Dutch)

transformers, and cable resistance). The assignments in the second environment expanded further on the basic principles of electricity (e.g., equivalent resistance) and electric power transmission (e.g., cable design, costs, high current).

In the final tab in each learning environment, students were again provided with the concept mapping tool, which now showed their initial concept map. They were then asked to update it to match their current knowledge and understanding (i.e., their final concept map).

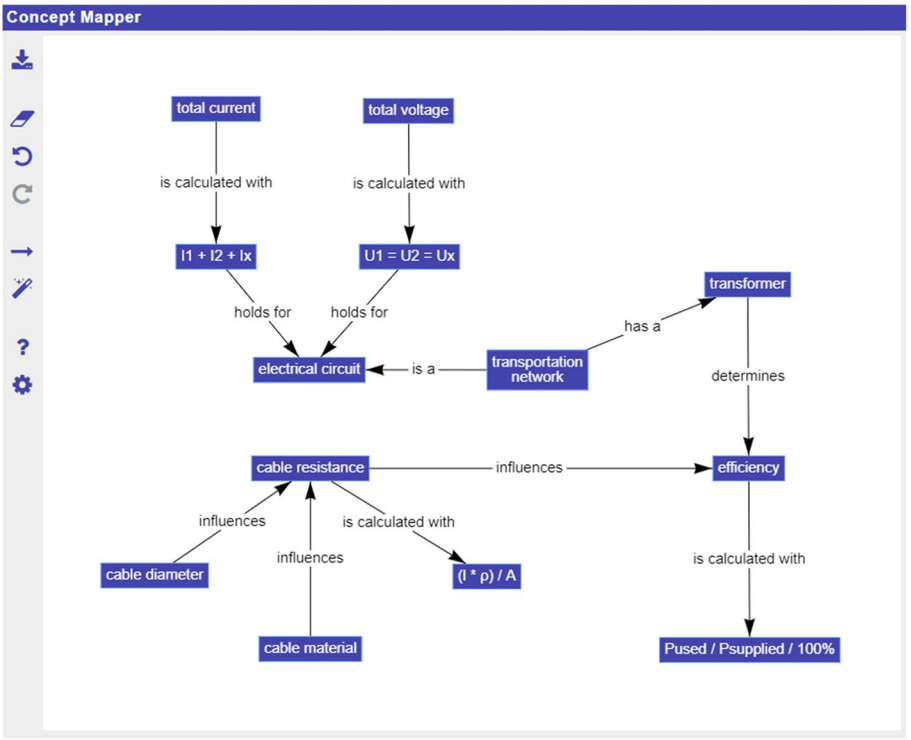

Fig. 3 Concept mapping tool containing a hypothetical student concept map

Concept mapping tool

The concept mapping tool enabled students to create their own concept maps (see Fig. 3). Students could insert concepts and could interrelate them by drawing a line between concepts and labeling the line. The tool scaffolded the concept-mapping process with predefined names for concepts and labels for links. Students could also insert self-defined concepts, which allowed them to use their own terminology and extend the concept map according to their own insight. The labels for links, on the other hand, could only be selected from a drop-down menu. Providing predefined linking labels not only supports students in creating their propositions, it also allows for better comparison with the example concept map (see next section), as their propositions would follow the same structure (Yin et al., 2005). Students in all conditions used this tool in the first and last assignments in both learning environments.

Combined concept map The concept mapping tool was supplemented with a combined concept map feature, which mapped students' concept map onto an expert example. The expert example was created by the researcher with the help of teachers who were experts within the subject area. This expert example represented one way, but not the only way, to display relevant key concepts of the domain and their underlying relations; for students, this concept map was described as 'the example concept map'. In the combined map, differences and commonalities between the expert map and the student's concept map were

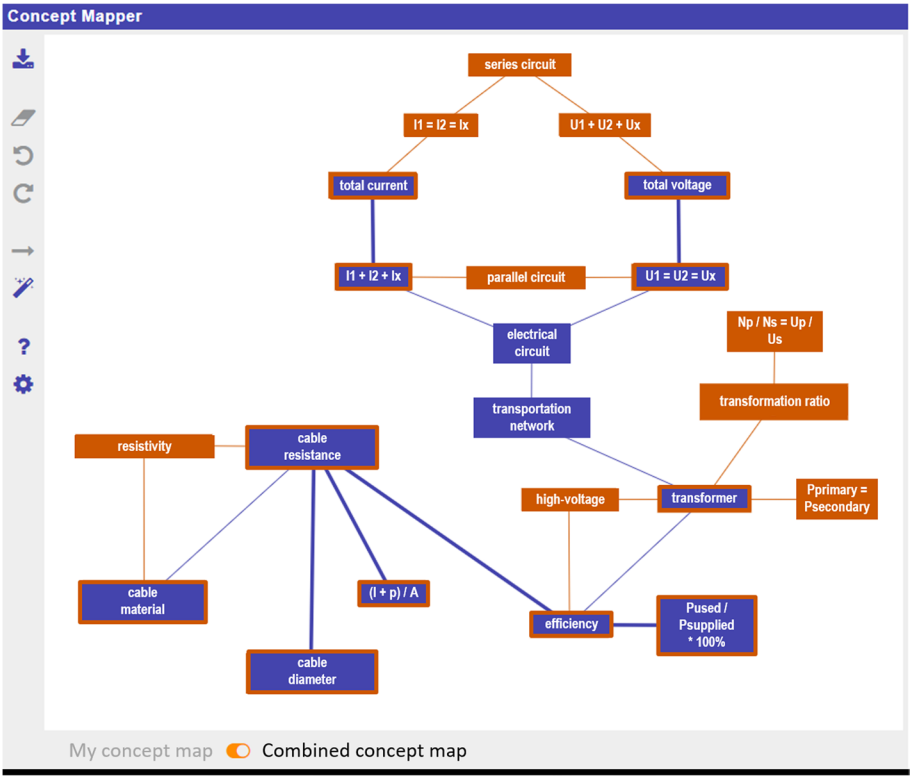

Fig. 4 Concept mapping tool with combined concept map activated (containing the expert example (translated from Dutch) and a hypothetical student concept map)

indicated with colors and line weights (see Fig. 4): concepts and lines that were unique to the expert concept map were displayed in orange, those that were unique to the student's concept map were displayed in purple; concepts that were present in both concept maps were shown as a purple box with an orange border, while links that were present in both concepts maps were represented by a thick (i.e., thicker than the unique links) purple line.

Students in all conditions other than the control condition used the combined concept map feature at the end of both learning environments. As part of the final assignment, students were instructed to activate and examine the combined concept map after they had finished updating their own concept map. The combined concept map could be (de)activated by clicking a button, which allowed them to switch between the combined concept map and their self-generated concept map.

Concept map training A classroom concept map training was used to support students' concept map creation. Training students in creating concept maps is especially advised when concept maps are new to students, due to the likelihood that when students are more familiar with creating them, they can focus more on their learning and knowledge that should be represented in their maps (Hilbert & Renkl, 2008).

The training on concept maps was designed to explain to students what concept maps are, what they can be used for and how to create them. The training consisted of three parts and

took about 20–25 min in total. The first part started with a brief explanation to the class of what concept maps are, what their components are, and why concept maps can be beneficial for students in a learning context. This explanation was followed by a demonstration of the concept mapping tool, during which an example concept map (about a school-related topic that was not related to the domain of the current study) was created; students were shown how to create a concept map and how to use the functionalities of the tool. For the second part, which was hands-on, students were instructed to extend the example concept map, working individually on their own device. After students explored the concept mapping tool themselves, the instruction ended with a final, interactive classroom part during which students' experiences with creating a concept map were discussed, whether they faced difficulties and if they understood what is entailed in creating a concept map with meaningful propositions.

Students in all conditions participated in the concept map training prior to working in the learning environment. The training was delivered by the students' own teacher.

Reflection prompts

The reflection prompts were designed to stimulate students to process and make sense of the differences between their own and the example concept map, as highlighted in the combined concept map, by asking them to evaluate both concept maps and providing them with suggested aspects to focus on when doing so. For students in both reflection prompts conditions, the prompts were displayed once the combined concept map was activated.

Students in the example + prompts condition were presented with a set of three prompts (see Fig. 5). They were asked to describe the differences between their own and the example concept map (1), after which they had to indicate for both the example concept map (2) and their own concept map (3) how they would improve them and *why*. For students who participated in the classroom discussion prior to responding to the prompts individually, only prompts 2 and 3 (see Fig. 5) were displayed. Instead of the first prompt (describing the differences between the example and their own concept map), at the start of the classroom discussion (see next section), the same assignment was given by the teacher (i.e., students were instructed to have a look at the differences between their own and the example concept map as highlighted in the combined concept map in their own learning environment).

Students' answers to the final two reflection prompts were coded using the coding scheme in Table 2. This means that four answers in total were coded for each student (two for each learning environment). Each answer received one of these mutually exclusive codes. A second coder coded 20% of the answers independently, which resulted in an inter-rater reliability (Cohen's Kappa) of 0.84.

Classroom discussion The classroom discussion aimed to deepen the reflective thinking process that students ideally would engage in when examining the combined concept map and filling out the reflection prompts. In preparation for the discussion, students first had a look at the highlighted differences between their own and the example concept map. Then, in order to help students to answer their first reflection prompt (i.e., number 2 of Fig. 5), students were stimulated to discuss how they would change or what they would add to the example concept map and why. To do this, the teacher asked guiding questions to elicit explanations and higher order thinking and therefore would benefit students' knowledge construction (Xun & Land, 2004). Examples of such questions that teachers asked their students were: why would you change/add this information, can you explain why this information is relevant/important,

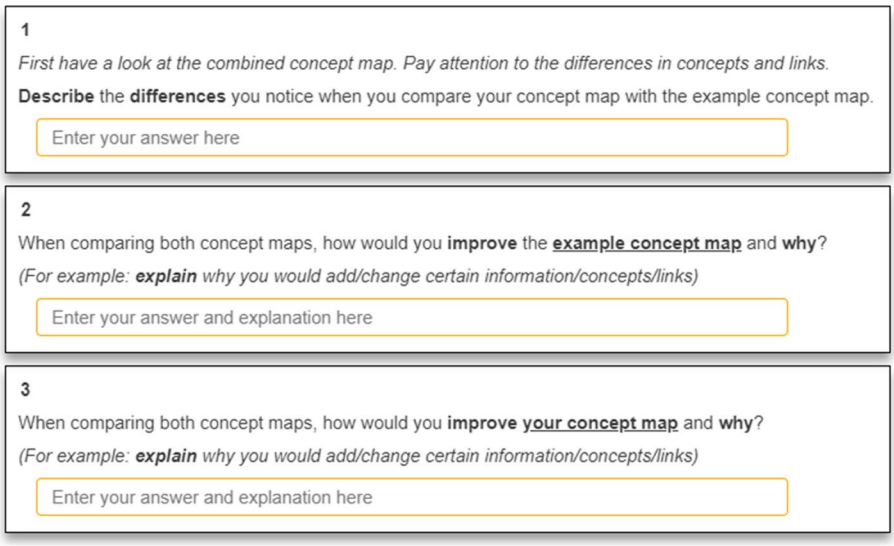

Fig. 5 Reflection prompts (translated from Dutch)

which concepts/relations would you use to represent this information, and why, according to you, is this information better represented in your own concept map? Other types of questions that teachers asked were meant to trigger discussions among students and to stimulate them to discuss alternative ideas or information. Examples of these questions were: who would (dis)agree and why, who would use other concepts/relations to represent that information and why, and who has other ideas to improve the example? After this first part of the discussion, students were instructed to individually respond to the first prompt that was presented below their combined concept map (i.e., number 2 of Fig. 5).

After all students had answered the first prompt, the focus of the classroom discussion shifted to students' own concept map, following the same procedure as with the example concept map. Unlike the discussion revolving around the example concept map, the source

Table 2 Overview of codes for answers to reflection prompts

Codes	Explanation
No answer to the prompt	
Empty	No answer is given
N/a	Answer not related to the question
Answer to the prompt	
No suggestion	Answer that indicates that no improvements or adaptations of the (example) concept map are needed or possible
Concept map lay-out	Answer, not content-related, that contains a suggestion related to improving/changing the structure or layout of the (example) concept map
Concept map content	Answer that contains a suggestion related to improving/changing or adding something to the content of the (example) concept map or related to the process of improving the content of the concept map, but not containing domain-related terms or information
Domain	Answer that contains domain-related terms/information or that demonstrates any reasoning about one's own knowledge

for this part of the discussion (i.e., students' own concept map) differed between the students. However, as the overarching topic was still the same, students could benefit from hearing each other's perspectives and explanations and from explaining their insights to others.

Only students in the example + discussion + prompts condition participated in the classroom discussion. The discussion was guided by their own teacher, and occurred at the end of individually working in each learning environment (sessions two and three), after students had completed their final, updated concept map.

Domain knowledge tests

Two parallel paper-and-pencil tests were used to measure students' domain knowledge before and after the intervention. Both domain knowledge tests were similar to the tests developed by Eshuis et al. (2022) and consisted of 11 open-ended questions each. Each question contained one sub-question that assessed knowledge at the conceptual level (e.g., recalling a definition or formula) and one sub-question that assessed knowledge at the application level (e.g., applying a formula or explaining a particular principle). A rubric was used to score the tests. Per test, a maximum of 22 points could be earned (11 points for items on the conceptual level; 11 points for items on the application level). A second coder coded 16% of the tests independently, which resulted in an interrater reliability (Cohen's Kappa) of 0.89 for the pretest and 0.85 for the posttest.

The parallel items assessed similar knowledge but differed from each other in context or formulation. For example, in both tests students were asked to calculate the equivalence resistance based on a provided electrical circuit including three separate resistors with predefined resistor values. Here the difference between the parallel items concerned the predefined values of the resistors. Or, students in both tests were asked to calculate the efficiency of a particular device. Here the difference between the parallel items concerned the described device (i.e., a laptop or a smartphone) and the provided power values needed to calculate the efficiency. Counterbalancing was used to prevent order effects. That is, approximately 50% of the students of each condition received version A as pretest and version B as posttest, and the remaining students received version B as pretest and version A as posttest. A univariate ANCOVA with pretest scores as dependent variable, school and prior training trajectory as covariates, and pretest version as independent variable revealed that pretest scores did not significantly differ between pretest versions, $F(1, 138) = 1.97$, $p = 0.163$, $\eta_p^2 = 0.014$, indicating that version A and B can be considered equally difficult. Reliability analysis yielded a Cronbach's alpha of 0.67 on the pretest and 0.76 on the posttest.

Procedure

Teacher training To prepare teachers for their role in the current study (i.e., providing concept map training and possibly guiding reflective classroom discussions), they all participated in a preparatory training session given by the researcher that took 60 min. The training at each participating school consisted of three parts. *First*, the aim of the study and the training were briefly explained, after which the various components of the study were discussed in more detail (i.e., the use of concept maps for learning, the combined concept map, the reflection prompts, the classroom discussion, and the online learning environments). *Second*, the teachers were provided with a hand-out containing all information on the fully scripted concept map training

and classroom discussion, after which the full concept map training was demonstrated by the researcher (i.e., the instruction with demonstration, hands-on, and interactive parts). During the hands-on part, teachers were asked to come up with and create a concept map about a school-related topic that had no relation to the domain of the current study, which they would also use as an example when training their own students. *Third*, based on the guidelines provided in the hand-out, the aim (i.e., guiding a classroom discussion), procedure and various possible scenarios for the classroom discussion were discussed (e.g., how to start and continue the discussion, how to engage students, what type of questions to pose to guide the discussion, and what to avoid when guiding the discussion). The teachers were explicitly told not to share any domain-related information, but only to stimulate and guide the discussion between the students.

The experiment The experiment took place in a real school setting during regular school hours. Prior to the experiment, a letter was sent to the students giving information about the purpose and procedure of the upcoming experiment that their school was participating in. They were given the option to indicate any objections regarding the use of their data before the start of the experiment.

During all sessions in all conditions both a researcher and a (to the students familiar) teacher was present. The teacher's presence was not only for the reason of providing the concept map training and—if applicable—guiding the classroom discussion, but also for classroom management purposes. The experiment involved four sessions; the first three sessions took around 90 min each, while the last session lasted for a maximum of 60 min. The first session started with a brief introduction during which students were informed about the upcoming lessons. Subsequently, students were given the domain knowledge pretest, which they had to complete within 40 min. After the students completed the pretest, the concept map training was given by their teacher, which took 20–25 min. At the start of the second session, the online learning environments with the labs were briefly introduced. After this introduction, students received a piece of paper with the URL of the learning environment, a login code, and brief instructions on the online labs and concept mapping tool. For students in the three example conditions, the written instruction about the concept mapping tool was extended with explanations about the combined concept map; students in both prompt conditions also received additional information about when to answer reflection prompts. Thereafter, students started working in the first learning environment. For the duration of the session, students had access to the complete first online learning environment, but they were instructed to proceed through the assignments in the learning environment in consecutive order. After 60 min, students were told to stop working on their current assignment and move on to the final tab (i.e., updating their initial concept map). The third session, during which students only had access to and worked in the second learning environment, followed basically the same procedure as session two. The follow-up procedure per condition for the second and third session is displayed in Table 3. In the fourth session students completed the domain knowledge posttest, within a 40 min time limit.

Results

The analyses presented below are based on data of 144 students (see Participants and design section for inclusion criteria). Univariate ANOVAs indicated no significant differences between conditions regarding students' age, $F(3, 138) = 0.16$, $p = 0.924$, $\eta_p^2 = 0.003$ and their pretest scores, $F(3, 140) = 1.82$, $p = 0.146$, $\eta_p^2 = 0.038$. Also, no significant differences were found for pretest scores when broken down by test item categories (conceptual

Table 3 Overview of intervention (sessions two and three) per condition

Time sequence *(total time in min.)*	Actions per condition			
	Control	Example	Example + prompts	Example + discussion + prompts
0-60 *(60 min.)*	Students work individually in the online learning environment			
60-70 *(10 min.)*	Students individually update their concept map of their knowledge			
		Students process example concept map		
70-85 *(15 min.)*	Students continue updating their concept map	Students continue updating their concept map and processing the **example concept map**	Students process **example concept map** with the help of **reflection prompts** that they answer **individually**	Students process **example concept map** with the help of **reflection prompts** that they answer individually **after discussing them as a class**

level: $F(3, 140) = 1.62$, $p = 0.188$, $\eta_p^2 = 0.033$; application level: $F(3, 140) = 1.94$, $p = 0.126$, $\eta_p^2 = 0.040$). This indicates that the four conditions in this final dataset were comparable – even after drop-out – in terms of age and prior knowledge.

Domain knowledge tests

Table 4 presents an overview of the mean pretest scores, posttest scores, and learning gains (posttest scores minus pretest scores) for each condition. A mixed MANOVA with Time (from pretest to posttest) as within-subject factor and Condition as between-subject factor was conducted. For the domain knowledge pre- and posttest, the two different test item categories (i.e., conceptual level and application level) were inserted as dependent variables. Results indicated a significant multivariate main effect of Time, $F(2, 139) = 36.53$, $p < 0.001$; Wilk's $\Lambda = 0.655$, $\eta_p^2 = 0.345$, a significant main effect of Condition, $F(6, 278) = 3.24$, $p = 0.004$; Wilk's $\Lambda = 0.874$, $\eta_p^2 = 0.065$, and a significant interaction effect of Time*Condition, $F(6, 278) = 2.56$, $p = 0.020$; Wilk's $\Lambda = 0.898$, $\eta_p^2 = 0.052$. Results of the univariate analyses are reported in Table 5. These results indicate significant effects of Time, Condition, and Time*Condition for items at both the conceptual and the application level.

Analyses on the multivariate simple effects of Time (Bonferroni controlled) indicated that, on average, students in all conditions significantly learned from pre- to posttest (control condition: $F(2, 139) = 5.62$, $p = 0.005$; Wilk's $\Lambda = 0.925$, $\eta_p^2 = 0.075$; example condition: $F(2, 139) = 4.72$, $p = 0.010$; Wilk's $\Lambda = 0.936$, $\eta_p^2 = 0.064$; example + prompts condition: $F(2, 139) = 4.76$, $p = 0.010$; Wilk's $\Lambda = 0.936$, $\eta_p^2 = 0.064$; example + discussion + prompts condition: $F(2, 139) = 36.09$, $p < 0.001$; Wilk's $\Lambda = 0.658$, $\eta_p^2 = 0.342$). Post-hoc comparisons (Bonferroni controlled) on these learning gains (i.e., pretest score to posttest score) showed that the found multivariate effect of Condition could be explained from significantly higher learning gains on the application level in the example + discussion + prompts condition compared to all other conditions (control condition $p = 0.026$;

Table 4 Mean pretest scores, posttest scores, and learning gains (max = 22) by condition, separated for conceptual level (max = 11) and application level (max = 11)

	Condition							
	Control (n = 34)		Example (n = 36)		Example + prompts (n = 27)		Example + discussion + prompts (n = 47)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pretest	4.75	3.41	4.91	3.81	4.84	3.38	6.19	2.64
Conceptual level	2.64	1.76	2.60	2.03	2.85	1.97	3.35	1.45
Application level	2.11	1.80	2.31	1.94	1.99	1.62	2.84	1.42
Posttest	6.15	3.49	6.06	4.43	6.24	3.39	9.21	3.67
Conceptual level	3.44	2.15	3.37	2.36	3.73	2.07	5.10	2.03
Application level	2.71	1.61	2.69	2.29	2.51	1.50	4.11	1.83
Learning gain	1.40	2.26	1.16	2.67	1.40	2.13	3.01	2.57
Conceptual level	.81	1.44	.77	1.65	.88	1.46	1.74	1.50
Application level	.59	1.38	.38	1.50	.52	1.18	1.27	1.43

example condition $p = 0.046$; example + prompts condition $p = 0.014$). For the learning gains on the conceptual level, those of the example + discussion + prompts condition were significantly higher compared to the control condition ($p = 0.027$) and the example condition ($p = 0.016$). All other comparisons were found not to be significant.

Answers to reflection prompts

In Table 6, frequencies (and percentages) of students' answer codes for the reflection prompts (see Table 2) are cross-tabulated by the two reflection prompt conditions. A chi-square test was performed to assess whether there was an association between condition and the types of answers students provided to the reflection prompts. This association was found to be significant: χ^2 (5, $N = 296$) = 25.49, $p < 0.001$, $V = 0.29$. Post-hoc column comparisons (Bonferroni controlled) revealed that students in the example + discussion + prompts condition answered the reflection prompts significantly more often than students in the example + prompts condition ($p < 0.001$). Students in the example + discussion + prompts condition also provided answers containing domain-related

Table 5 Univariate tests of the mixed MANOVA for Time and Time*Condition

Effect	Dependent variable	df	*F*	*p*	η_p^2
Time	Conceptual level	1, 140	66.34	< .001	.322
	Application level	1, 140	34.22	< .001	.196
Condition	Conceptual level	3, 140	4.25	.007	.083
	Application level	3, 140	4.71	.004	.092
Time*Condition	Conceptual level	3, 140	3.98	.009	.079
	Application level	3, 140	3.41	.019	.068

Table 6 Frequencies and percentages per answer code by condition

Code	Condition					
	Example + prompts (n = 27)		Example + discussion + prompts (n = 47)		Total	
	Count	%	Count	%	Count	%
Empty	22	20.4	9	4.8	31	10.5
N/a	11	10.2	16	8.5	27	9.1
No suggestion	13	12.0	24	12.8	37	12.5
Concept map lay-out	14	13.0	24	12.8	38	12.8
Concept map content	46	42.6	92	48.9	138	46.6
Domain	2	1.9	23	12.2	25	8.4
Total	108	100	188	100	296	100

information significantly more often compared to students in the example + prompts condition ($p = 0.002$).

To explore whether the answers students provided to the reflection prompts affected their knowledge acquisition, a backwards stepwise regression analysis was performed, with the frequency of each answer code for the reflection prompts (empty, n/a, no suggestion, concept map lay-out, concept map content, and domain) and condition (example + prompts = 0; example + discussion + prompts = 1) as possible predictors and students' learning gain as outcome. In the first (full) model (adjusted $R^2 = 0.159$), $F(6, 67) = 3.30$, $p = 0.007$, 'concept map content' was excluded because it reached the minimum tolerance level (< 0.001). To check the assumption of no multicollinearity, tolerance statistics of the remaining variables in the model were consulted. These values indicated no cause for concern (i.e., values ranged from 0.62 to 0.79). Correlations are reported in Table 7. In the subsequent models 'empty', 'no suggestion', 'n/a', 'concept map lay-out', and 'condition' were excluded consecutively. The final model (adjusted $R^2 = 0.129$), $F(1, 72) = 11.79$, $p = 0.001$, only included 'domain-related answers to reflection prompts' ($\beta = 0.375$) as a significant predictor of learning gain.

Table 7 Correlations between learning gain, condition, and answer code frequencies (n = 74)

	1	2	3	4	5	6	7	8
1. Learning gain	-							
2. Condition	.309**	-						
3. Empty	-.185	-.358**	-					
4. N/a	-.226	-.044	-.206	-				
5. No suggestion	.028	.019	-.207	-.088	-			
6. Concept map lay-out	.138	-.005	-.206	-.236*	-.236*	-		
7. Concept map content	-.054	.135	-.299**	-.171	-.282*	-.288	-	
8. Domain	.375**	.311**	-.162	-.235*	-.099	-.077	-.224	-

*$p < .05$, **$p < .01$

Discussion and conclusion

The main finding of our study is that students at the level of secondary vocational education can be stimulated to reflect on their knowledge, and in this way improve their learning results from concept mapping. However, these students may need more support than normally is reported in the literature. More specifically, in contrast to research findings that suggest the added value of providing an example concept map in combination with students creating their own concept maps (Kao et al., 2008; Schwendimann & Linn, 2016), the current study demonstrated that merely adding an example does not add to the value of students creating their own concept map (in terms of knowledge acquisition as measured by a pre- and posttest). For the population in this study, introducing a classroom discussion before answering reflection prompts proved to be a successful intervention.

The finding that providing an example was not enough to significantly improve students' knowledge acquisition can be explained by the fact that successful knowledge integration is the driving force in students' knowledge acquisition when learning from concept maps. This implies that students should engage in several cognitive and metacognitive processes – such as organizing, elaborating on, and monitoring their knowledge – which can take place in a process of reflection. The literature is clear on the challenges these (meta)cognitive processes present to students (Quintana et al., 2005). With a student population from secondary vocational education (MBO in Dutch) these challenges were likely to be only more prevalent in this study; the level of education of our students has lower admission requirements than other secondary programs (i.e., university or university of applied sciences) and since (meta)cognitive skills are closely aligned with academic achievement (Ohtani & Hisasaka, 2018) secondary vocational students are more likely to struggle with (meta)cognitive skills.

Support: structure alone is not enough

The current study addressed a rarely studied population, secondary vocational students, and took their unique characteristics into account in the design of the support. In addition to assessing whether generally found effects of examples would also apply to this population, the current study focused on how to support these students in reflection processes required to benefit from the presented examples. More specifically, two modalities of additional support were investigated: reflection prompts (individually processed), and reflection prompts combined with classroom discussion.

In relation to these modalities of additional support, the results of the current study demonstrated that individual processing of reflection prompts was not sufficient to foster knowledge acquisition, but that reflection prompts do have added value when combined with a teacher-guided classroom discussion in which students jointly reflect and discuss. This is in contrast with research that shows that offering prompts can provide adequate support to promote students' knowledge acquisition (e.g., Bannert & Mengelkamp, 2013; Chen et al., 2009; Ifenthaler, 2012). Nonetheless, it is recognized that there is no uniform conclusion as to which type of prompt is generally the most effective (Kori et al., 2014). It can be argued that differences in outcome are related to the type of prompts that are offered; these different types can, for example, be distinguished based on the amount of structure and direction they provide (Wylie & Chi, 2014). Although our reflection prompts did include instruction regarding where to focus on, they can generally be considered

rather open (i.e., no limits in the formulation and type of answers). One might argue that our students in particular need more structure to benefit from reflection prompts. However, previous research (Eshuis et al., 2022) has shown that offering rather specific prompts (i.e., mostly check the box items) has not led to an increase in students' learning gains with this target group either.

A more likely explanation of the outcome lies in the notion that other studies involving reflection prompts in general involved students from a different level of education (mostly higher or university education). The findings of this study suggest that, in order for secondary vocational students to engage in adequate reflection, they need more extensive support; only structuring the process, as reflection prompts do, was not enough. These students need to be prepared to process and answer these prompts to benefit from them.

The results of this study indicates that classroom discussion provides an accessible and easily adaptable means to prepare students for the (meta)cognitive processes required. Strong positive effects of classroom discussion on students' comprehension of text were already clear from a meta-analysis by Murphy et al. (2009), and the current study suggests that the importance of classroom discussion extends beyond that situation and that a classroom discussion may also influence the effectiveness of other instructional aids.

The results met our expectation that students who participated in classroom discussion prior to responding to reflection prompts would outperform students in all other conditions regarding their learning gain (i.e., learning gains in this condition were, on average, significantly higher for items on the application level compared to all other conditions; for items on the conceptual level learning gains were, on average, significantly higher compared to all other conditions except for the example + prompts condition). The hypothesis underlying this expectation was that successful knowledge integration depends on how students process the reflection prompts. Students' level of processing depends on their motivation, which, in turn, is a product of their perceived ability to complete the task and the perceived value of completing the task (Wigfield & Eccles, 2000). Our findings support the credibility of this line of reasoning; in comparison to students who processed the reflection prompts completely individually, students whose class discussed the prompts left the reflection prompts unanswered less often and provided more domain-related answers.

Specifically, the domain-relatedness of answers was a predictor of students' learning gain. In this respect it should be noted that students whose class discussed the prompts provided more domain-related answers and thus learned more, but also that within this condition there were still a considerable number of answers that were not domain-related. This could imply that there are specific interactions during a classroom discussion that could affect its impact. Because the teachers in the current study were trained to structure the classroom discussion in a similar fashion, therefore minimizing the impact of differences in teacher support, it is most likely that differences in student contributions and interaction affected the outcome.

Now that we have an indication that classroom discussion may be a pivotal component to make reflection prompts successful in their mission, follow-up research is suggested to gain insight into the actual working mechanisms that drive this impact. More specifically, research should aim to identify factors that impact students' follow up reflections and knowledge acquisition. We suggest two directions for future research. First, to substantiate our assumption that the benefit of jointly processing reflection prompts is due to student motivation, follow-up research could include assessments of student motivation for the task at hand. Second, to identify whether specific interactions and cognitive processes during classroom discussion are essential, follow-up research could capture the content and dynamics of the group discussions. Knowing exactly how various discussions unfold can provide valuable insight into possible cognitive processes, such as the extent to which students' elaborate upon their

knowledge. Students engaging in a discussion on a metacognitive level might yield different effects than students engaging in discussion on a cognitive level in terms of the quality of their response to reflection prompts and their knowledge acquisition. Indications of possible effects resulting from metacognitive processes can be students' explicit revisions regarding shortcomings and misconceptions in their knowledge due to adequate knowledge monitoring. These follow-up directions would provide valuable information that can be used to ensure that a classroom discussion is designed for optimal affordances for learning.

Conclusion

The fact that many students, even more experienced and university students, struggle to reflect on their knowledge (Quintana et al., 2005) may be resulting from a combination of a lack of students' metacognitive skills (i.e., the ability to evaluate their cognitive performance and regulate their cognitive activities) and a lack of their cognitive skills (i.e., the ability to organize and elaborate upon their knowledge). These metacognitive skills and their related cognitive processes are an essential component in knowledge construction (Azevedo, 2005; Ohtani & Hisasaka, 2018). Hence, finding support to optimize these processes is highly sought after, but results are ambiguous.

In line with other studies, the current study showed that students who typically experience difficulty in expressing themselves verbally and who are in general used to having a more practical than theoretical focus, such as secondary vocational students (Slaats et al., 1999), require tailored support. For our students, introducing a classroom discussion before answering reflection prompts was found to be an effective add-on to the support offered. The success of classroom discussion for these kinds of students was substantiated by the findings of Murphy et al. (2009). Their meta-analysis showed that the effects of classroom discussion are stronger for students with lower academic abilities compared to average and above-average students. Our findings corroborate these results.

Though the findings of this study imply that support (such as examples and reflection prompts) that has proven to be of value for other students does not necessarily work for students in secondary vocational education, we should be wary of accepting the reverse as true. It could well be that classroom discussion as implemented in the current setting can be also be beneficial to students in other levels of education. In addition to this, we would like to point out that the current study adopted an approach where the classroom discussion was tied to the reflection prompts. The comparison between the conditions with reflection prompts only allows us to attribute the results to the discussion; therefore, it remains unclear whether a classroom discussion without the reflection prompts (which students individually answered after discussion) would be equally effective. Furthermore, it could be noted that contextual factors may impact classroom discussion. In the present study, it was a deliberate choice to resemble a usual classroom situation as much as possible, by using intact classes. Thus, maintaining a familiar environment for the students. (Non)familiarity among students can impact the effectiveness of the intervention; not being familiar or being less familiar with fellow students impacts the way students communicate and collaborate with each other (Janssen et al., 2009) and can therefore limit the effect of such a classroom discussion. Future research should contribute to our understanding of how classroom discussion should be structured and facilitated.

To conclude, the current study fostered secondary vocational students' knowledge integration by offering them an aid that supported the processing of information from an expert example concept map combined with their self-generated concept map. A combination of

reflection prompts and classroom discussion proved to be effective in this respect. From what we understood from the teachers in the current study, whole-class discussion on such an overarching, metacognitive, level is, in general, not a strategy that they practice. Making teachers more aware of these kinds of interventions and providing tools and training to assist them in implementing these in their classroom can benefit educational practice in secondary vocational education.

Acknowledgements The authors wish to thank Anjo Anjewierden (†), Jakob Sikken, and Hagop Jamkojian for the technical support and Emily Fox for her helpful comments and edits.

Author contribution Not applicable.

Funding This work was supported by the Netherlands Organisation for Scientific Research (NWO) [project number 409–15-209], TechYourFuture, and Thales Nederland BV.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Materials and/or Code availability Not applicable.

Declarations

Ethics approval The procedures used in this study were in accordance with the ethical standards of and approved by the BMS Ethics Committee of the University of Twente.

Consent Passive consent was obtained from all individual participants in this study.

Conflicts of interest The authors declare that they have no conflict of interest.

References

Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist, 40*(4), 193–197. https://doi.org/10.1207/s15326985ep4004_1

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35*(4), 359–375. https://doi.org/10.2190/94v6-r58h-3367-g388

Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (Vol. 28, pp. 171–195). Springer Science.

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist, 48*(4), 243–270. https://doi.org/10.1080/00461520.2013.838920

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Chen, N.-S., Wei, C.-W., Wu, K.-T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education, 52*(2), 283–291. https://doi.org/10.1016/j.compedu.2008.08.007

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439–477. https://doi.org/10.1016/0364-0213(94)90016-7

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12*(1), 91–142. https://doi.org/10.1207/S15327809JLS1201_4

de Jong, T., Gillet, D., Rodríguez-Triana, M. J., Hovardas, T., Dikke, D., Doran, R., Dziabenko, O., Koslowsky, J., Korventausta, M., Law, E., Pedaste, M., Tasiopoulou, E., Vidal, G., & Zacharia, Z. C. (2021). Understanding teacher design practices for digital inquiry–based science learning: The case of Go-Lab. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-020-09904-z

Eshuis, E. H., ter Vrugte, J., Anjewierden, A., & de Jong, T. (2022). Expert examples and prompted reflection in learning with self-generated concept maps. *Journal of Computer Assisted Learning*, *38*(2), 350–365. https://doi.org/10.1111/jcal.12615.

Fan, W. (2011). Social influences, school motivation and gender differences: An application of the expectancy-value theory. *Educational Psychology, 31*(2), 157–175. https://doi.org/10.1080/01443410.2010.536525

Ge, X., & Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development, 51*(1), 21–38. https://doi.org/10.1007/bf02504515

Hilbert, T. S., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instructional Science, 36*(1), 53–73. https://doi.org/10.1007/s11251-007-9022-9

Hilbert, T. S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Computers in Human Behavior, 25*(2), 267–274. https://doi.org/10.1016/j.chb.2008.12.006

Hmelo-Silver, C. E., & Barrows, H. S. (2006). Goals and strategies of a problem-based learning facilitator. *Interdisciplinary Journal of Problem-Based Learning, 1*(1), 4. https://doi.org/10.7771/1541-5015.1004

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Journal of Educational Technology & Society, 15*(1), 38–32.

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2009). Influence of group member familiarity on online collaborative learning. *Computers in Human Behavior, 25*(1), 161–170. https://doi.org/10.1016/j.chb.2008.08.010

Kao, G. Y. M., Lin, S. S. J., & Sun, C. T. (2008). Breaking concept boundaries to enhance creative potential: Using integrated concept maps for conceptual self-awareness. *Computers & Education, 51*(4), 1718–1728. https://doi.org/10.1016/j.compedu.2008.05.003

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research, 30*(1–2), 139–161. https://doi.org/10.2190/AX2D-Y9VM-V7PX-0TAD

Kim, D., & Lee, S. (2002). Designing collaborative reflection supporting tools in e-project-based learning environments. *Journal of Interactive Learning Research, 13*(4), 375–392.

Kinchin, I. M. (2014). Concept mapping as a learning tool in higher education: A critical analysis of recent reviews. *The Journal of Continuing Higher Education, 62*(1), 39–49. https://doi.org/10.1080/07377363.2014.872011

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review, 11*, 45–55. https://doi.org/10.1016/j.edurev.2013.11.003

Lim, K. Y., Lee, H. W., & Grabowski, B. (2009). Does concept-mapping strategy work for everyone? The levels of generativity and learners' self-regulated learning skills. *British Journal of Educational Technology, 40*(4), 606–618. https://doi.org/10.1111/j.1467-8535.2008.00872.x

Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*(7), 837–858. https://doi.org/10.1007/BF02214052

Lin, X., Hmelo, C., Kinzer, C. K., & Secules, T. J. (1999). Designing technology to support reflection. *Educational Technology Research and Development, 47*(3), 43–62.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101*(3), 740–764. https://doi.org/10.1037/a0015576

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413–448. https://doi.org/10.3102/00346543076003413

Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education, 86*(4), 548–571. https://doi.org/10.1002/sce.10032

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*(2), 179–212. https://doi.org/10.1007/s11409-018-9183-8

Onderwijsconsument. (2019). *Wat betekent kwalificatieplicht* (What does 'obligation to qualify' mean)? Retrieved from https://www.onderwijsconsument.nl/kwalificatieplicht

Quintana, C., Zhang, M., & Krajcik, J. (2005). A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding. *Educational Psychologist, 40*(4), 235–244. https://doi.org/10.1207/s15326985ep4004_5

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*(1), 1–29. https://doi.org/10.1016/S0364-0213(99)80017-2

Renkl, A. (2014a). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*(1), 1–37. https://doi.org/10.1111/cogs.12086

Renkl, A. (2014b). The worked examples principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 391–412). Cambridge University Press.

Renkl, A., & Schwonke, R. (2013). Static visual displays for deeper understanding: How to support learners to make use of them. In G. Schraw, M. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 165–185). Information Age.

Romero, C., Cazorla, M., & Buzón, O. (2017). Meaningful learning using concept maps as a learning strategy. *Journal of Technology and Science Education, 7*(3), 313–332. https://doi.org/10.3926/jotse.276

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455. https://doi.org/10.1007/s10648-017-9403-9

Schwendimann, B. A., & Linn, M. C. (2016). Comparing two forms of concept map critique activities to facilitate knowledge integration processes in evolution education. *Journal of Research in Science Teaching, 53*(1), 70–94. https://doi.org/10.1002/tea.21244

Slaats, A., Lodewijks, H. G. L. C., & van der Sanden, J. M. M. (1999). Learning styles in secondary vocational education: Disciplinary differences. *Learning and Instruction, 9*(5), 475–492. https://doi.org/10.1016/S0959-4752(99)00007-9

ter Vrugte, J., & de Jong, T. (2017). Self-explanations in game-based learning: From tacit to transferable knowledge. In P. Wouters & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games* (pp. 141–159). Springer.

ter Vrugte, J., de Jong, T., Wouters, P., Vandercruysse, S., Elen, J., & van Oostendorp, H. (2015). When a game supports prevocational math education but integrated reflection does not. *Journal of Computer Assisted Learning, 31*(5), 462–480. https://doi.org/10.1111/jcal.12104

ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., & Elen, J. (2017). Computer game-based mathematics education: Embedded faded worked examples facilitate knowledge acquisition. *Learning and Instruction, 50*, 44–53. https://doi.org/10.1016/j.learninstruc.2016.11.007

van Merriënboer, J. J. G., & Kester, L. (2005). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 71–93). Cambridge University Press.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Winne, P. H. (2021). Cognition, metacognition and self-regulated learning. In G. Noblit (Ed.), *Oxford research encyclopedia of education.* Oxford University Press.

Wu, P.-H., Hwang, G.-J., Milrad, M., Ke, H.-R., & Huang, Y.-M. (2012). An innovative concept map approach for improving students' learning performance with an instant feedback mechanism. *British Journal of Educational Technology, 43*(2), 217–232. https://doi.org/10.1111/j.1467-8535.2010.01167.x

Wylie, R., & Chi, M. T. H. (2014). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 413–432). Cambridge University Press.

Xun, G. E., & Land, S. M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development, 52*(2), 5–22. https://doi.org/10.1007/BF02504836

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching, 42*(2), 166–184. https://doi.org/10.1002/tea.20049