Instructional Science (2025) 53:1691–1714
https://doi.org/10.1007/s11251-024-09694-4

ORIGINAL RESEARCH

# Promoting future teachers' pedagogical knowledge: The role of self-generated vs. provided illustrative examples after instruction

Theresa Krause-Wichmann[1] · Eric Klopp[1] · Lisa Stark[1] · Robin Stark[1]

Received: 25 September 2023 / Accepted: 26 November 2024 / Published online: 17 April 2025


## Abstract
Illustrative examples demonstrate how abstract information can be applied in real-world. In the context of advancing evidence-informed teaching practice, the current intervention study investigated to what extent student teachers should be supported in learning educational theories and findings by different example-based approaches. Conducting a $1 \times 3$-factorial design, $N = 105$ student teachers were randomly assigned to three experimental groups: After a pre-test, all groups received the same learning instruction on the topic of cooperative learning. Then, (1) $n = 35$ students were prompted to generate own examples for the instructional text, (2) $n = 35$ students received examples along with the text, and (3) $n = 35$ students studied the text only, without any prompts or examples. In a post-questionnaire, it was retrospectively assessed how students perceived their learning control in engaging with the material; in a post-test, knowledge retention and knowledge transfer were measured. As assumed, findings revealed that generating examples enhanced perceived learning control and learning outcomes compared to studying provided examples. Students who learned with the instructional text only achieved lowest learning outcomes; but contrary to the expectations, these students perceived their learning control comparably high as those who generated examples. Mediation analyses indicated that for students who received illustrative examples or the instructional text only, a greater learning control perception was positively associated with knowledge retention, subsequently enhancing knowledge transfer. The study underscores the benefits of illustrative examples in teacher education, particularly when students engage in generating them. It suggests further examination of how and why example generation facilitates learning.

**Keywords** Teacher education · Evidence-informed teaching · Illustrative examples · Generative learning

✉ Theresa Krause-Wichmann
theresa.krause-wichmann@uni-saarland.de

[1] Department of Education, Saarland University, Saarbrucken, Germany

# Introduction

## Background and relevance

Teachers are increasingly asked to include evidence from educational research into lesson planning, analysis, and reflection with the aim of optimizing teaching and learning (*evidence-informed teaching practice*; Ferguson, 2021). The term *evidence* encompasses both empirical findings and general bodies of knowledge such as theories, concepts, or principles (Renkl, 2022; Wilkes & Stark, 2023). For example, when planning group work, it is advisable to consider theoretically well-founded elements of cooperative learning and findings on scaffolding. However, several studies have indicated that pre- and in-service teachers rarely draw on evidence in practical questions (Dagenais et al., 2012; Kiemer & Kollar, 2021). Besides unfavorable boundary conditions (e.g., lack of resources; Thomm et al., 2021) or dispositional prerequisites (e.g., beliefs; Voss, 2022), a lack of applicable pedagogical knowledge is regarded as one of the main reasons for this lack of evidence use (Neuweg, 2014). Thus, it is essential that student teachers acquire a robust foundation of pedagogical knowledge that can be transferred to diverse pedagogical questions or problems. Supporting student teachers in acquiring applicable knowledge requires instructional approaches in teacher education that go beyond knowledge retention and foster enduring understanding. To promote deep engagement with learning materials, it is important that learners experience a sense of control during learning (Corbalan et al., 2009; Ryan & Deci, 2000). Besides providing student teachers with choices in what or how they learn (Patall et al., 2008), control perception could also be enhanced when they recognize the real-world applications of the abstract concepts they are studying. A promising approach to bridge the gap between theory and practice in teacher education is *instruction followed by illustrative examples* (Rawson et al., 2015).

*Instruction followed by illustrative examples* is an example-based learning approach that comprises two different phases: In the initial instruction phase, learners are explicitly informed about the new, abstract learning content (e.g., theory and empirical findings). In studying the instructional text, learners can gain a first understanding by integrating the information with existing prior knowledge (Anderson & Krathwohl, 2001). In the subsequent illustration phase, the new content is connected with an example that concretely instantiates how the content can be mapped on real-world situations (Rawson et al., 2015). Although the use of illustrative examples has a long-standing tradition in formal education, such as in university textbooks (Rawson et al., 2015), there is still insufficient and inconsistent evidence of how they affect knowledge retention and transfer. Particularly, it is unclear whether it might be more beneficial for learning, if students are asked to generate own examples based on their own experience instead of being provided with pre-defined examples (e.g., by a teacher or a text; Zamary & Rawson, 2018). As the role of perceived learning control for knowledge retention and transfer has been insufficiently considered in research on both examples and teacher education, there is also a need to explore the relationships between knowledge retention, knowledge transfer, and perceived learning control in different instructional approaches.

Therefore, the present study investigates to what extent learning outcomes (i.e., *knowledge retention, knowledge transfer*), *perceived learning control*, and their relationships differ among student teachers who generate their own examples (i.e., *instruction followed by self-generated examples*), those who study provided examples (i.e., *instruction followed by provided examples*), and those who do not engage with examples during learning at all (i.e., *instruction only*).

## Fostering retention and transfer of pedagogical knowledge

Pedagogical knowledge encompasses declarative and procedural knowledge facets (Voss et al., 2011). As defined by Anderson and Krathwohl (2001) in their two-dimensional taxonomy of educational objectives and student learning, declarative knowledge refers to specific content elements such as terms and facts (i.e., *factual knowledge*; e.g., definition of different group effects) as well as theories, concepts, and principles (i.e., *conceptual knowledge*; e.g., social learning theory). Procedural knowledge includes both knowledge of skills, techniques, or methods and the criteria used to determine when to apply a procedure (i.e., *conditionalized knowledge;* e.g., managing group work). In accordance with the hierarchical nature of Anderson's and Krathwohl's (2001) taxonomy, the ability to retrieve relevant knowledge from memory (i.e., *knowledge retention*) is regarded as a necessary precondition for applying knowledge in answering new questions or solving problems (i.e., *knowledge transfer*). This hierarchy is based on the assumption that transfer involves higher-leveled, complex cognitive processes (e.g., understanding, applying, analyzing, evaluating, and creating) that cannot be managed without an adequate knowledge base. This theoretical hierarchy is supported by empirical findings from different fields (e.g., Butler, 2010; Khasawneh et al., 2024) and is further discussed by Roediger and Butler (2011), who emphasize the critical role of retrieval practice in fostering knowledge retention and transfer.

However, although retrievable knowledge can be considered a necessary precondition for transfer, it cannot be regarded as sufficient by itself. Adapting retained knowledge to diverse practical situations requires a deep understanding. Particularly, deficits in knowledge structure, i.e., compartmentalization of acquired knowledge and own experiences can cause knowledge to remain *inert* (Renkl et al., 1996). To prevent the development of inert knowledge and to foster both retention and transfer of pedagogical knowledge, instruction must extend beyond simply presenting abstract information. Student teachers should be encouraged to encode and interconnect new information with application possibilities, enhancing elaboration. By generating their own examples or studying provided examples after instruction, learners' knowledge becomes *conditionalized*, reducing the risk of it remaining inert (Renkl, 2017; Renkl et al., 1996).

## Benefits of self-generated and provided illustrative examples

By helping learners to connect abstract information to practical applications, instruction followed by provided illustrative examples might more effectively support the retention and transfer of knowledge, compared to studying instruction only. The advantages of provided illustrative examples for learning have also been empirically demonstrated (Balch, 2005; Micallef & Newton, 2022; Rawson et al., 2015; Wissman et al., 2023). However, provided examples entail inherent challenges that could reduce the depth of engagement. For example, learners might divert their focus to less pertinent example details (Harp & Mayer, 1998) or overemphasize the example itself over the instructional text (LeFevre & Dixon, 1986). Against this background, it might be promising to incorporate more active learning strategies.

Generative learning theory (Wittrock, 1989) and the interactive-constructive-active-passive (ICAP) framework (Chi & Wylie, 2014) both suggest advantages of active and generative engagement compared to rather passive activities such as reading. Grounded in a constructivist view of learning, both frameworks share the key idea that active engagement facilitates learning and transfer. Generative learning theory suggests that generative activities foster understanding and sense-making, as learners are encouraged to actively integrate the new information with prior knowledge and experience and to deeply process the information at a meaningful level (Fiorella, 2023; Wittrock, 1989). The ICAP framework gives a more granular classification based on the level of cognitive engagement: Learning will increase as learners become more engaged with the material, from a *passive* (e.g., reading a text) to an *active* (e.g., highlighting a text) to a *constructive* (e.g., mapping) engagement (Chi & Wylie, 2014)[1]. Compared to the rather passive task of (re-)studying instruction (with or without examples), generating examples mandates learners to identify core features of the new content and incorporate them into specific instances (Anderson & Krathwohl, 2001). Since learners need to form meaningful links with prior knowledge or experience and reflect on their understanding, the risk of knowledge compartmentalization is reduced (Renkl et al., 1996). The involved activities of aligning, evaluating, and organizing can further enhance the depth of processing and improve learning outcomes (Fiorella, 2023; Markant et al., 2016).

Despite the theoretical advantages of generative activities, research on the effects of instruction followed by self-generated vs. provided examples shows mixed and contradictory results. A pioneer in research on illustrative examples is Hamilton, who conducted several studies in the field of educational psychology (Hamilton, 1989, 1990, 1997, 1999, 2004). In Hamilton's seminal study (1989), participants studied an instructional text that defined the concepts of reinforcement and punishment, followed by a set of four examples, further explanations, and multiple-choice questions with immediate feedback. Compared to the control group, the example generation group was additionally asked to self-generate further examples. The latter yielded better performance in applying the new concepts, but the groups did not differ in recall performance. However, the results must be interpreted with caution, as Hamilton's material already contained pre-defined examples, independently from condition (for a discussion cf. Rawson & Dunlosky, 2016). In Hamilton's subsequent studies (1990, 1997, 1999, 2004), conditions requiring example generation were outperformed by other learning techniques, including studying additional provided examples. Other studies investigating self-generated examples in (educational) psychology have not demonstrated significant advantages for learning over provided examples (Dornisch et al., 2011; Griffin, 1993; Steininger et al., 2022). In a study by Zamary and Rawson (2018), generating examples proved less beneficial for learning compared to studying examples, while Rawson and Dunlosky (2016) showed that generating examples was at least more beneficial than (re-)reading an instructional text only.

From here, the question arises why the theoretical benefits of example generation over studying are hardly reflected in empirical findings. Besides methodological limitations of Hamilton's studies, the specificity and practical applicability of the learning content could also be a factor. The learning contents used in the aforementioned intervention studies, such as test score interpretation (Griffin, 1993) or standardized testing and normal distribution (Dornisch et al., 2011), are undoubtedly fundamental to (educational) psychology and

[1] *Note*. *Generative learning* is not closely related to the *generation effect* that concerns enhancing long-term memory by self-generating information (Slamecka & Graf, 1978; cf. Waldeyer et al., 2020).

(pedagogical) practice. However, they are quite abstract and less observable in everyday practice compared to topics such as motivational theories or different learning approaches (e.g., cooperative learning). Limited direct applicability—possibly paired with a lack of relevant prior knowledge and experience—could make it challenging for students to come up with own real-world instances, which in turn could impede learning from instruction followed by self-generated examples (Rawson & Dunlosky, 2016). Therefore, example generation might be more effective with topics that are directly related to students' everyday practice.

Given the inconsistent findings in previous research, it is necessary to explore when and why learning from instruction followed by self-generated vs. provided illustrative examples works. The sense of control that students perceive in their learning process is worth considering in this question, as it might be decisive in shaping motivation and effort allocation (Corbalan et al., 2009; Ryan & Deci, 2000).

## The role of perceived learning control for learning

The cognitive processes that are crucial for learning from instruction can be initiated either by instructors (e.g., teachers) or by the students themselves. Several theories that focus on how to structure learning emphasize the importance of enabling learners to control their learning process, such as the elaboration theory of instruction (Reigeluth et al., 1980) and the 4C/ID model (van Merriënboer et al., 2002). In addition to these cognitive-focused theories, motivational theories, such as the self-determination theory (Ryan & Deci, 2000) and the control-value theory (Pekrun, 2006), highlight the relevance of learning control from an affective-motivational perspective. However, the control granted to learners can only enhance learning if it is *perceived* as such by the learners themselves: Only if they *feel* a strong connection between their activities and their learning success, their sense of agency, self-endorsement and autonomy is strengthened; consequently, they are more likely to engage in the learning activity, invest more effort, and allocate their cognitive resources effectively (Corbalan et al., 2009; Perry et al., 2005; Stark et al., 2018). Thus, perceived control could promote knowledge retention and transfer; in turn, a lack of perceived control could decrease involvement and learning success.

Perceived control might also be a crucial factor for explaining the benefits of instruction followed by (provided) illustrative examples. Although reading per se is an autonomous activity (Shuell, 1988), reading a text with examples might enhance learners' perception of control compared to reading it without examples. When learners are directly enabled to link a theoretical concept with its practical applications, and understanding is deepened, they might rather feel a strong connection between the reading-activity and their learning outcomes. However, it is difficult to predict whether this also applies to self-generated examples. On the one hand, learners who self-develop own ideas based on individual experience might feel to have the ownership over the material. From this perspective, self-generated examples might lead to higher perceived control than provided examples, which do not account for individual experience. On the other hand, the opposite is also conceivable: Especially in early stages of knowledge acquisition, the high responsibility in generative activities might overwhelm learners (Steininger et al., 2022), resulting in a loss of control perception.

Based on theoretical considerations it could not be argued, to what extent self-generating examples vs. studying provided examples vs. studying an instructional text only affects perceived learning control. Even if it can be assumed that perceived control has an influence on learning outcomes in terms of knowledge retention and transfer, no statement can be made as to whether these relations differ across the different instructional approaches, respectively.

## Research questions and hypotheses

To draw conclusions on whether instruction in teacher education should be enriched by prompts that encourage student teachers to generate own illustrative examples or by provided illustrative examples, student teachers studied different versions of learning material about the topic cooperative learning. Firstly, we analyzed the effects of *instruction followed by self-generated examples* vs. *instruction followed by provided examples* vs. *instruction only* on learning outcomes (i.e., *knowledge retention, knowledge transfer*; research question 1) and *perceived learning control* (research question 2). Secondly, we aimed to explore the role of perceived learning control in enhancing learning outcomes across these three instructional approaches (research questions 3 and 4). To address these research questions, the following hypotheses were formulated:

*Research question 1: Effect of instruction on learning outcomes.* In line with the theoretical considerations of the ICAP framework (Chi & Wylie, 2014) and benefits of generative activities in general (Fiorella, 2023), a constructive mode of engagement through generating examples should promote learning outcomes more effectively than passively studying provided examples or (re-)reading the text only. Thus, we hypothesized that students who generate own examples should achieve higher learning outcomes in terms of both *knowledge retention* and *transfer* than students who study provided examples or the instructional text only. Against the background of the advantages of examples in general (Rawson & Dunlosky, 2016; Rawson et al., 2015), we expected that learning with *self-generated* as well as learning with *provided examples* lead to higher *knowledge retention* and *transfer,* respectively, than *instruction only*.

*Research question 2: Effect of instruction on perceived learning control.* Although reading a text (with or without examples) is an autonomous activity per se (Shuell, 1988)*,* it seems plausible to expect that the theory–practice connection given by *instruction followed by provided examples* could increase *perceived learning control* compared to *instruction only*. However, this does not necessarily apply to self-generated examples, which place high demands on the learners. On the one hand, it could be argued that students who generate examples might perceive the highest control, as they have the autonomy to create ideas based on individual experience. On the other hand, the required cognitive demands might overwhelm students, potentially leading to a decreased control perception. Thus, it was assumed that there is a difference in the effects of *instruction followed by self-generated examples* on *perceived control* compared to *instruction followed by provided examples* and to *instruction only*, respectively. Given the conflicting possibilities of effects, no specific directional hypotheses were established, here.

*Research question 3: Relations between learning outcomes and perceived learning control.* Given the importance of control perception for the motivation to invest effort in acquiring knowledge (Corbalan et al., 2009), we (1) hypothesized that perceived learning control should be positively related to learning outcomes in terms of both knowledge retention and transfer. Put differently, the ability to remember knowledge is considered to be important to perform more complex activities (Anderson & Krathwohl, 2001). Thus, we expected that knowledge retention should be positively related to transfer. Given that (a) perceived control is crucial for learning and (b) knowledge retention is a prerequisite for

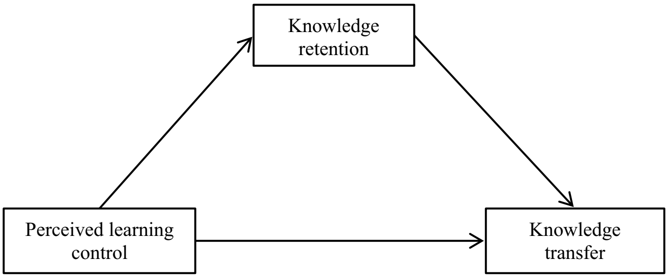


**Fig. 1** Regression model assumed for each of the three conditions (hypothesis 3)

transfer, we (2) hypothesized that knowledge retention will serve as a mediator with respect to the relation of perceived control and knowledge transfer (see the regression model in Fig. 1).

*Research question 4: Effect of instruction on the relations*. The *relations* assumed under hypothesis 3 between perceived learning control, knowledge retention, and knowledge transfer may vary across conditions (i.e., *self-generated examples* vs. *provided examples* vs. *instruction only*). However, based on the current state of research, no specific hypotheses could be established. Therefore, we sought to investigate the differences in these relations exploratorily.

## Method

### Participants and design

$N = 105$ student teachers ($M_{\text{Age}} = 22.34$, $SD = 2.96$; $M_{\text{Sem}} = 5.93$, $SD = 4.28$; 72% female) participated as part of their university courses. It was mandatory to take part in the training elements, but it was voluntary to participate in the data collection. In an experimental $1 \times 3$-factorial between-subjects design, the factor *example after instruction* was varied, resulting in three experimental conditions with $n = 35$ participants each (Table 1). Participants were randomly assigned to one of the three conditions. All participants studied the same instructional text, while it was varied whether they were prompted to generate illustrative examples (i.e., *instruction followed by self-generated examples*) vs. studied provided illustrative examples (i.e., *instruction followed by provided examples*) vs. (re-)studied the text without any further prompts or examples (i.e., *instruction only*).

### Procedure and material

The study comprised two sessions of 45 min that were conducted independently and individually with time limit. Before the first session started, participants answered a pre-questionnaire on socio-demographic data and a pre-test. After the pre-test, they had 45 min time to engage with the learning material in a self-regulated manner. All participants received the same 3000-word instructional text on cooperative learning, which contained the following sections: (1) *What is cooperative learning compared to other forms of learning?* (2) *To what extent is cooperative learning effective—and why?* (3) *When does cooperative learning not work—and why?* (4) *How can cooperative learning be fostered?* (5) *How can*

**Table 1** Exemplary illustration of the learning material (Sect. 1) for the example-based conditions

| Text section | Instruction followed by self-generated examples (translated) | Instruction followed by provided examples (translated and abbreviated) |
|---|---|---|
| (1) What is cooperative learning compared to other forms of learning? | Please develop a short fictional example of a learner with a competitive attitude and of a learner with an individualistic attitude that has a negative impact on a group work | *Example of a competitive and of an individualistic attitude with a negative effect on a group work:* Thomas wants to be the best in class. Thus, he pretends that he cannot answer the questions of his group members and keeps his knowledge to himself, so that the others cannot benefit from his knowledge (*competitive*). Lea prefers to work on her own. Thus, she prepares everything for the project at home, so that there is hardly anything left for the group to do in the next few lessons (*individualistic*) |

The instruction-only condition comprised neither prompts nor examples, which is why it is not presented here

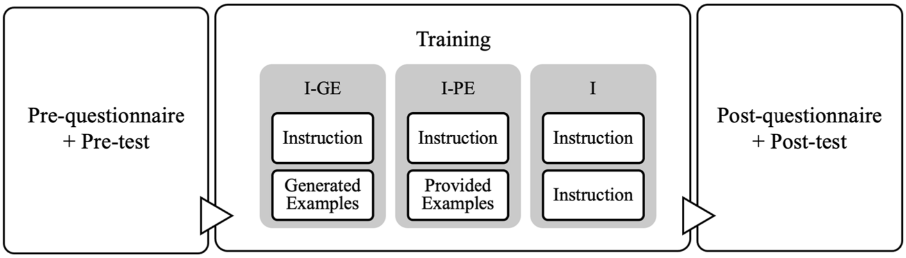


**Fig. 2** Design and procedure. *I-GE* Instruction followed by self-generated examples; *I-PE* Instruction followed by provided examples; *I* Instruction only

*cooperative learning be structured?* All students were instructed to study the material carefully and memorize as much as possible. In the end of each section, students received (a) prompts to generate examples (I-GE), or (b) provided examples (I-PE), or (c) no prompts or examples (I), depending on the condition (cf. *Experimental conditions;* Table 1). After the 45 min had passed, the material was collected to prevent further engagement with it until the next session. In the second session one week later, the material was laid out in front and participants could take it for further engagement based on their individual code. After 45 min, the material was collected, and a post-questionnaire that retrospectively assessed perceived learning control was handed out. Finally, the participants completed a 45-min post-test. An overview of the entire procedure is presented in Fig. 2.

## Experimental conditions

Participants learning with *instruction followed by self-generated examples* were asked to (1) engage with the content of each section and (2) generate an appropriate example that applies a particular content to classroom practice (Fig. 2: I-GE; Table 1, column 1). Similarly, students who learned with *instruction followed by provided examples* were asked to (1) engage with the content of each section and (2) study the example illustrating the content using a classroom scenario (Fig. 2: I-PE; Table 1, column 2). Where possible, the provided examples did not refer to any school subject; if a reference was necessary, it varied between different school subjects (e.g., language studies, social sciences, natural sciences, mathematics), to reduce a possible effect of the participants' subjects. Participants learning with *instruction only* were asked to (re-)study the material in both sessions; they did neither receive prompts to generate examples nor were they provided with pre-defined examples (Fig. 2: I).

## Measures

### Learning outcomes

The participants completed a self-designed learning performance test on cooperative learning. The test included seven recall tasks to assess *knowledge retention* (max. 36 points; e.g., "Please describe Johnson's & Johnson's basic elements of cooperative learning"; average split-half reliability = 0.72) and five case-based application tasks to assess *knowledge transfer* (max. 18 points; average split-half reliability = 0.72). The case-based application tasks required evaluating the realization of group work in written text vignettes ("Please discuss how group work is realized"). The vignettes varied in length and complexity: Four shorter (60–120 words) and less complex cases referred to different didactical-methodical implementations of group work in different subjects (mathematics, biology, philosophy, geography). The fifth, more complex case consisted of 500 words; it described not only the implementation of group work but also referred to different group phenomena by describing how four groups interacted with each other. One of the cases of less complexity was as follows (translated):

> "*Biology teacher Ms. Adam has her class conduct experiments in groups of three. Each group should test the activity of a water plant in relation to the factors light, carbon dioxide and temperature. Therefore, she handed out experiment instructions to each group. To ensure that everyone really experiments once, she decided that one student per group should focus on a specific factor. Later, everyone will present the factor he or she has worked on.*"

Two coders were trained to code the participants' answers. It was assessed whether the students discussed the realization of group work regarding functionality (1 point) and explained it by referring to relevant evidence from the instructional text (1 point). Accordingly, two points could be achieved with each of the four short cases. For the more complex scenario, a maximum of ten ($5 \times 2$) points was awarded. 20% of the sample were double-coded blind to conditions with sufficient inter-rater-reliability $\kappa > 0.80$. The total scores for *knowledge retention* and *knowledge transfer* were calculated by adding the achieved scores.

### Perceived learning control

To assess *perceived learning control*, eight items were adapted from Perry et al. (2001) to the specific context of the learning environment; the items were reformulated to capture the control that students perceived in learning retrospectively on a 5-point Likert scale ranging from 1 = not at all true to 5 = very much true. For example, the item "I have a great deal of control over my academic performance in my psychology course" was reformulated to "I had a great deal of learning control with the material"). Reliability was sufficient (*Cronbach's* $\alpha = 0.76$).

### Control measures

To check the randomization, participants completed a short version of the learning performance test with three of the recall tasks (*knowledge retention*; max. 13 points) and four of the case-based transfer tasks (*knowledge transfer*; max. 8 points). Above that, the participants' beliefs regarding the *usefulness of educational evidence* were measured by nine items adapted from Wagner et al., (2016); "Teachers should apply educational evidence when reflecting on their teaching."; 1 = not at all true, 6 = very much true; *Cronbach's* $\alpha = 0.89$). Participants also had to rate five items measuring their *academic self-concept* regarding their educational studies (adapted from Dickhäuser et al., 2002; e.g., "I hold my talent for the educational studies for 1 = low, 7 = high."; *Cronbach's* $\alpha = 0.75$).

### Example quality

As example quality might be related to learning outcomes (Rawson & Dunlosky, 2016), the *quality of generated examples* was assessed. In the condition *instruction followed by self-generated examples*, the generated examples were coded on whether students had applied the evidence appropriately in a real-world situation through a high-inferential rating (0 = no example; 1 = low quality example/instruction paraphrazed; 2 = moderate quality example; 3 = high quality example). Inter-rater-reliability was sufficient (20% double-coded; $\kappa > 0.80$).

### Analytic strategy

For all calculations, R (R Core Team, 2020, version 4.02) was used with the lavaan package (Rosseel, 2012, version 0.6.15). A multi-group approach was applied, i.e., a mediation model was set up for each experimental condition (see Fig. 1). The model is saturated and can exactly estimate the empirical means and variances of each variable in each experimental condition while at the same time providing the mediation model's regression coefficients separately in each condition. The model parameters, i.e., regression coefficients, means, and variances, for all three groups are estimated simultaneously within a single objective function, aiming to minimize the discrepancy between the observed data and the model predictions. The multi-group approach additionally allows assessing the model parameters between the conditions using Wald tests, which test the restriction that the respective parameter is equal. Thus, the multi-group approach allows a comparison of the means and regression coefficients across the conditions by testing the restriction that the difference between the parameters equals zero.

Regarding research questions 1 and 2, the mean differences between the conditions were calculated and investigated with Wald tests and planned contrasts to compare the means between the conditions. The following contrasts were used in line with our hypotheses: I-GE vs. I, I-PE vs. I, and I-GE vs. I-PE. We used Cohen's *d* to indicate the effect sizes of the between-group comparisons. For research question 3, we considered the direct effects (path coefficients) and computed the indirect and total effects using defined parameters. Concerning research question 4, we explored whether the direct effects differed between the conditions using Wald tests and the same planned contrasts as mentioned above. For all tests, the nominal significance was $\alpha = .05$. As some variables were not normally distributed within the conditions, we used the MLM estimator that yields robust standard errors (cf., Finney & DiStefano, 2013).

Regarding the power, we conducted a Monte Carlo simulation with 2000 repetitions. In this case, power is defined as the proportion of repetitions for which the null hypothesis is rejected for a given parameter with $\alpha = .05$, assuming that the current sample provided the quantity of the population parameters (cf., Beaujean, 2014). Power is calculated for each model parameter. We considered a minimum level of 0.50 with an ideal level of 0.80 (Kyriazos, 2018) as sufficient.

## Results

### Preliminary analyses

There were no a-priori differences between the experimental conditions regarding the control measures knowledge retention (pre-test), $F(2,102) = 0.045$, $p = .956$, knowledge transfer (pre-test), $F(2,102) = 2.07$, $p = .081$, usefulness of educational evidence, $F(2,102) = 0.610$, $p = .545$, and academic self-concept, $F(2,102) = 0.383$, $p = .683$ (see Table 2 for the descriptive statistics). The results regarding example quality revealed that the generated examples were of modest to high quality ($M = 2.67$; $SD = 0.45$). The quality of the generated examples was highly correlated with knowledge transfer ($r = .352$; $p = .038$) and perceived learning control ($r = .432$, $p = .010$), but not with knowledge retention ($r = .280$; $p = .103$).

### Effects of illustrative examples

The mean differences that were calculated to investigate *research questions 1* and *2* are presented in Table 3 (for means, standard deviations and correlations, see Table 2).

In terms of *knowledge retention* (KR), the contrasts revealed that the outcomes differed significantly between students who learned with self-generated examples (I-GE) and students who learned with provided examples (I-PE) with medium effect size. Students who generated examples achieved better results. The instruction-only condition (I) differed significantly from the two example-based conditions. The effects were medium-sized (I-PE/I) to large (I-GE/I). Students who were not provided with examples or prompts showed the lowest test performance.

Concerning *knowledge transfer* (KT), the contrasts indicated that the performance of students who learned with self-generated examples (I-GE) differed significantly from those who learned with provided examples (I-PE) and from those who restudied instruction only (I). The effects were medium-sized (I-GE/I-PE) to large (I-GE/I). Students who generated examples achieved the best results. The difference between the condition that provided students with examples (I-PE) and the instruction-only condition (I) was also significant with a medium effect size. The latter showed the lowest outcomes, again.

With respect to *perceived learning control* (PC), the contrasts indicated that results differed between the examples-based conditions (I-GE/I-PE) with medium effect size. Students who generated examples perceived their control to be higher than students who studied examples. Students who learned with instruction only (I) reported a higher control perception than students who were additionally provided with examples (I-PE). The result was significant with medium effect size. The difference between instruction followed by self-generated examples (I-GE) and instruction only (I) was not significant.

### Relations of learning outcomes and perceived learning control

Concerning *research question 3*, the coefficients of the direct, indirect, and total effects of the mediation model for each condition are presented in Table 4. Consideration should also be given to the results of the overall correlation analyses (Table 2), which indicated that perceived learning control was positively correlated with knowledge retention but not with transfer.

In the condition *instruction followed by self-generated examples* (I-GE), the results of the mediation analysis did not align with our assumptions: The direct, indirect, and total effects failed statistical significance. Neither perceived learning control nor knowledge retention were related to knowledge transfer.

In contrast, for students who learned with instruction followed by provided examples (I-PE), all direct paths were significant and positive. As expected, for students who were provided with illustrative examples, higher perceived control led to higher learning outcomes (PC $\rightarrow$ KR; PC $\rightarrow$ KT) and higher knowledge retention led to higher transfer

**Table 2** Descriptive statistics and correlations

| | I-GE | | I-PE | | I | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | (1) | (2) | (3) | (4) | (5) | (6) |
| 1. Knowledge retention (post-test) | 22.54 | 4.94 | 19.16 | 6.32 | 13.77 | 7.17 | – | | | | | |
| 2. Knowledge transfer (post-test) | 8.57 | 2.41 | 6.67 | 2.22 | 5.20 | 2.23 | .53** | – | | | | |
| 3. Perceived learning control | 4.07 | 0.50 | 3.71 | 0.53 | 4.16 | 0.52 | .26** | .10 | – | | | |
| 4. Knowledge retention (pre-test) | 3.85 | 1.37 | 3.74 | 1.84 | 3.83 | 1.51 | .29** | .33** | .15 | – | | |
| 5. Knowledge transfer (pre-test) | 2.46 | 0.80 | 2.29 | 0.85 | 1.98 | 1.03 | .30** | .34** | .03 | .27** | – | |
| 6. Usefulness of educational evidence | 4.10 | 0.91 | 4.04 | 0.93 | 3.87 | 0.94 | .34** | .23* | .25* | .06 | .22* | – |
| 7. Academic self-concept | 4.70 | 0.65 | 4.54 | 0.82 | 4.64 | 0.79 | .23* | .27** | .24* | .13 | .34** | .43** |

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

*I-GE* Instruction followed by self-generated examples; *I-PE* Instruction followed by provided examples; *I* Instruction only

**Table 3** Mean comparisons (planned contrasts; research questions 1 and 2)

| | Knowledge retention | | | | | | Knowledge transfer | | | | | | Perceived learning control | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Diff* | $\chi^2$ | *df* | *p* | *d* | *Power* | *Diff* | $\chi^2$ | *df* | *p* | *d* | *Power* | *Diff* | $\chi^2$ | *df* | *p* | *d* | *Power* |
| I-GE/I-PE | **3.39** | **6.41** | **1** | **.011** | **0.51** | .721 | **1.90** | **12.11** | **1** | **< .001** | **0.66** | .931 | **0.36** | **8.56** | **1** | **.003** | **0.57** | .835 |
| I-GE/I | **8.77** | **36.58** | **1** | **< .001** | **1.24** | 1.000 | **3.37** | **37.87** | **1** | **< .001** | **1.17** | 1.000 | − 0.10 | 0.59 | 1 | .444 | 0.15 | .107 |
| I-PE/I | **5.39** | **11.44** | **1** | **< .001** | **0.67** | .919 | **1.47** | **7.88** | **1** | **.005** | **0.54** | .804 | **− 0.45** | **13.11** | **1** | **< .001** | **0.70** | .950 |

*I-GE* Instruction followed by self-generated examples; *I-PE* Instruction followed by provided examples; *I* Instruction only; Significant results are printed in bold letters

**Table 4** Path coefficients of direct, indirect, and total effects in each experimental condition (research question 3)

| | PC→KR | | | | | PC→KT | | | | | KR→KT | | | | | Indirect effect | | | | | Total effect | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Est* | *SE* | *p* | *Std* | *Power* | *Est* | *SE* | *p* | *Std* | *Power* | *Est* | *SE* | *p* | *Std* | *Power* | *Est* | *SE* | *p* | *Std* | *Power* | *Est* | *SE* | *p* | *Std* | *Power* |
| I-GE | 0.95 | 1.37 | .487 | .10 | .145 | 0.31 | 0.85 | .714 | .07 | .111 | 0.04 | 0.04 | .362 | .08 | .126 | 0.04 | 0.05 | .508 | .01 | .003 | 0.35 | 0.85 | .684 | .07 | .116 |
| I-PE | **7.14** | **1.67** | **<.001** | .60 | .986 | **1.13** | **0.58** | **.049** | .27 | .453 | **0.16** | **0.06** | **.005** | .45 | .775 | **1.12** | **0.48** | **.019** | .27 | .680 | **2.25** | **0.51** | **<.001** | .54 | .954 |
| I | **5.57** | **2.42** | **.021** | .41 | .751 | − 0.21 | 0.48 | .668 | .05 | .095 | **0.13** | **0.05** | **.006** | .43 | .731 | **0.74** | **0.38** | **.049** | .17 | .367 | 0.53 | 0.58 | .360 | .13 | .153 |

*I-GE* Instruction followed by self-generated examples; *I-PE* Instruction followed by provided examples; *I* Instruction only; *PC* Perceived learning control; *KR* Knowledge retention; *KT* Knowledge transfer; *Est* Unstandardized estimate; *Std* Standardized estimate; Significant results are printed in bold letters

(KR → KT). The indirect effect and the total effect were significant, too. This implies that for these students, a higher perceived control resulted in superior learning outcomes. Knowledge retention acted as a mediator supporting our assumption that a higher perceived control led to higher knowledge transfer because of having a positive impact on knowledge retention.

For students who learned with instruction only (I), the direct paths from perceived control to knowledge retention (PC → KR) and from knowledge retention to transfer (KR → KT) were significant, but not the path from perceived control to knowledge transfer (PC → KT). The indirect effect was significant, but the total effect did not reach statistical significance. In line with our expectations, all significant paths were positive indicating that higher perceived control led to greater outcomes with knowledge retention acting as mediator. In other words, although a higher perceived control did not directly lead to higher knowledge transfer, it did so indirectly by enhancing knowledge retention.

Table 5 presents the comparisons of the direct, indirect, and total effects between the experimental conditions (*research question 4*). Concerning the direct effects of perceived control on knowledge retention (PC → KR), the contrasts revealed that the path coefficients differed significantly between students who learned with self-generated examples (I-GE) and students who learned with provided examples (I-PE), with the latter exhibiting a more substantial effect. Regarding the direct effects of perceived control on knowledge transfer (PC → KT) and of knowledge retention on transfer (KR → KT), all tests failed statistical significance.

The indirect effects differed across conditions. The contrasts showed a significant difference between the example-based conditions (I-GE/I-PE), with a more pronounced effect in the condition instruction followed by provided examples (I-PE). Thus, the mediating role of knowledge retention is more emphasized with provided than with generated examples. The other comparisons were not significant indicating no substantial differences regarding the mediating role of knowledge retention.

Lastly, a distinct difference emerged between the condition in which the instruction was paired with provided examples (I-PE) and the condition in which only the instruction was given (I). The total effect was more pronounced for students who learned with provided examples (I-PE) which indicated that the relationship of perceived control and knowledge transfer was more apparent.

## Discussion

Motivated by the aim to promote evidence-informed teaching practices, the current work systematically examined the efficacy of instruction coupled with self-generated and provided illustrative examples in comparison to mere instruction on the acquisition of applicable pedagogical knowledge in student teachers. To clarify when and why example-based instruction is effective, the study investigated the role of perceived learning control, considering knowledge retention as a critical prerequisite for transfer.

### Composite, example-based approaches are beneficial for learning

Consistent with established theoretical expectations (Renkl, 2017) and previous research (Rawson & Dunlosky, 2016; Rawson et al., 2015), our findings underscore the benefits of composite, example-based approaches compared to learning with mere instruction:

**Table 5** Planned contrasts of the direct, indirect, and total effects between the experimental conditions (research question 4)

| | PC→KR | | | | PC→KT | | | | KR→KT | | | | Indirect effect | | | | Total effect | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Diff* | $\chi^2$ | *p* | *Power* | *Diff* | $\chi^2$ | *p* | *Power* | *Diff* | $\chi^2$ | *p* | *Power* | *Diff* | $\chi^2$ | *p* | *Power* | *Diff* | $\chi^2$ | *p* | *Power* |
| I-GE/I-PE | **− 6.18** | **8.15** | **.004** | .771 | − 0.82 | 0.63 | .427 | .148 | − 0.12 | 2.99 | .084 | .269 | **− 1.08** | **5.10** | **.024** | .551 | − 1.90 | 3.66 | .056 | .506 |
| I-GE/I | − 4.62 | 2.76 | .097 | .447 | 0.52 | 0.28 | .597 | .117 | − 0.10 | 2.31 | .128 | .203 | − 0.70 | 3.44 | .064 | .284 | − 0.18 | 0.03 | .859 | .078 |
| I-PE/I | 1.56 | 0.28 | .595 | .123 | 1.34 | 3.17 | .075 | .325 | 0.02 | 0.11 | .742 | .083 | 0.38 | 0.39 | .530 | .105 | **1.72** | **4.94** | **.026** | .487 |

*I-GE* Instruction followed by self-generated examples; *I-PE* Instruction followed by provided examples; *I* Instruction only; *PC* Perceived learning control; *KR* Knowledge retention; *KT* Knowledge transfer; Significant results are printed in bold letters

Students who learned either with own or provided examples outperformed those who learned without, in terms of both knowledge retention and transfer. The alignment of content with practice obviously helped students contextualizing the information. Consequently, they were able to better remember and apply it.

#### Self-generated examples are more beneficial for learning than provided examples

When looking more closely at the effectiveness of the different example-based approaches, generating appeared superior to studying examples. While previous studies have reported ambiguous results, the pattern of our findings corresponds to the assumptions of ICAP framework (Chi & Wylie, 2014) and generative learning theory (Fiorella, 2023). The rather passive activity of studying examples might not have optimally activated students to engage in deeper, integrative processes, potentially leading to a more superficial processing, compared to generating examples. That provided examples were less beneficial than self-generated examples could also be explained by the reduced correspondence to students' individual experience or subject matter. A recent study by Steininger et al. (2024), for instance, indicated benefits of examples matching to the students' subject. Notably, in contrast to the study by Zamary and Rawson (2018), in the present study, the quality of generated examples was rather high and positively related to the ability to apply knowledge. Example quality might be a key of explaining the inconsistencies in previous research. Our findings further support the assumption that example generation unfolds its potential in dependence of the intuitive practice-applicability of the subject matter.

#### Self-generating examples and studying an instructional text only give a high sense of perceived control

Regarding the students' control perception, it is remarkable that learning with the mere instructional text led students perceive a level of control comparable to that of students who generated examples. Both the active and passive approach obviously led the students feel a strong link between the learning activity and outcomes (Perry et al., 2005). Thus, the assumption that the high demands of example generation could reduce control perception has not been confirmed. In contrast, the opportunity to develop own ideas based on individual experience and to determine the level of complexity obviously increased the students' feeling of being able to engage with the material in a self-determined way.

The unexpected lower control perception of the group who was provided with examples might be explained with a missing individuality of pre-defined examples, compared to self-generated examples. If the examples did not match the students' experiences, they may have felt a lower connection between the task and learning success (Perry et al., 2005). The high ratings of students who learned exclusively with the text might be explained by the fact that reading per se is an autonomous activity (Shuell, 1988). Students might also be used to this form of learning from university everyday life.

According to the so-called *feigned choice paradigm* (Schneider et al., 2018) providing learners with *seemingly* meaningful choices—even if the given options do not significantly alter the learning content or process—can enhance learning, too. Findings by Schneider et al. (2018) indicate that feigned choice can foster knowledge retention and transfer, with perceived decisional autonomy mediating the effect on knowledge retention. It is possible that in the instruction-only condition, estimation of perceived control reflected a rather superficial, feigned control, whereas in the self-generation group, it reflected actual control.

### Role of perceived control for learning outcomes and interplay of knowledge dimensions

The results of the mediation analyses revealed that only for students who were not engaged in a generative activity the level of perceived control influenced knowledge transfer via retention. Thus, the postulated relevance of control perception for learning (Corbalan et al., 2009) and the function of knowledge retention for transfer (Anderson & Krathwohl, 2001) could only be confirmed for students who were engaged in rather passive activities. However, it is striking that the self-generation group—who was granted the highest degree of autonomy and also felt this way—showed neither any relationship between perceived control and knowledge retention or transfer nor between the learning outcome variables themselves. It can only be speculated why there were these inconsistent patterns of results: Maybe, these students who generated examples were driven by the task itself to deeply process the information, so that the level of perceived control was of rather secondary importance, here. Moreover, the fact that perceived control was remarkably high in the generation group suggests a restriction of variance, which could make it difficult to detect any effects in the mediation model. However, since the instruction-only group also had high perceived control values, this restriction of variance might not totally explain the missing mediating relationship for the group who generated examples. The difficult-to-understand findings might also be explained by the complexity of (self-)assessing (perceived) control itself. The measurement could have led to biased ratings, especially since it was assessed retrospectively once after two sessions. The ratings could also be dependent on personal perception of time restriction.

## Limitations and future directions

Although the current study presents valuable insights into the benefits of instruction followed by self-generated and provided examples in teacher education, it is not without limitations.

### External validity

The ability to generate appropriate examples mainly relies on one's knowledge base and experience. Therefore, it needs to be clarified how much prior practical experience (such as school internship) is needed to be able to generate suitable (and high quality) examples. Future research could conduct aptitude-treatment interaction analyses to determine how students' level of prior knowledge might influence the effectiveness of different example-based instructional approaches. Future studies could also investigate the effects of fading procedures, which is a common approach in learning with worked examples (Atkinson et al., 2003). Fading might be implemented by first studying provided examples, followed by a combination of provided and self-generated examples or by complementing partially completed examples, and, finally, self-generated examples only (Zamary & Rawson, 2018). In this context, it must be noted that we did not investigate effects of more inductive forms of learning, such as classifying examples (cf. Brunmair & Richter, 2019 for an overview), which turned out to be more effective for learning than generating or studying examples in the study by Steininger et al. (2022). This inductive learning strategy might be a beneficial alternative to generating or studying examples, as it is still more active than studying provided examples (Chi & Wylie, 2014), but less demanding than generating examples. The fact that, in this study, self-generated examples demonstrated their potential specifically with the topic of cooperative learning suggests investigating whether the effectiveness of different example-based approaches varies depending on the practical applicability of the learning content.

### Internal validity and manipulation check

In general, internal validity can be assumed as secured, since the experimental conditions did not differ with respect to prior knowledge and possible other motivational/volitional confounding variables (i.e., *usefulness of educational evidence*, *academic self-concept*). However, it should be noted that although all participants had to spend 45 min per session on the material, we cannot make statements about how much time students actively spent on learning. As the uniform duration of the sessions provides a certain degree of control over time-on-task, it can be—albeit with some restrictions—assumed that the observed results are primarily due to the effectiveness of the instructional approaches and not to different amounts of time-on-task. Moreover, we did not conduct a manipulation check apart from assessing example quality in the condition *instruction followed by self-generated example*s. As manipulation checks bear the risk to amplify the intervention or trigger the effect of the intervention (Hauser et al., 2018), we refrained from asking students in the other two conditions, whether they have created own examples. Thus, it is uncertain whether students in these conditions might have spontaneously used self-generating as learning strategy, even if they were not instructed to do so. Further studies should explore the role of example quality, including effects of self-evaluation and feedback (Guerrero et al., 2024; Zamary et al., 2016). Here, a potential moderating role of example quality in the relationship between perceived control and learning outcomes should be considered.

### Presentation format of the learning material

The presentation format in the two example-based conditions might be another limiting factor itself. The learning material consisted of five sections, following the sequence *instruction-example* each (except for the instruction-only condition). Due to paper pencil format, students might have jumped through the material and might have not complied with the given sequence. Future studies that aim at comparing composite designs (for example instruction-example vs. example-instruction vs. interleaving; Rawson et al., 2015) should involve environments in that the adherence to sequence can be better controlled (such as online environments). Here, further process measures that ask for self-monitoring strategies, cognitive load (Sweller, 2020) or motivational/volitional factors should be incorporated.

### Conclusion and practical implications

The present study is closely aligned with the needs of teacher education, as it seeks to bridge the gap between theory and practice to promote evidence-informed teaching practices. In addition, it contributes to the ongoing debate in example-based learning research by examining the value of self-generated examples. Our findings offer strong implications for both research on composite instructional designs and the development of materials in teacher education that aim to foster the acquisition of applicable pedagogical knowledge.

First of all, the study highlights the critical importance of using examples to illustrate theoretical concepts in teacher education. Grounding pedagogical knowledge in real-world examples enables learners to connect abstract concepts with practical applications, thereby deepening their understanding. Specifically, our findings advocate for the integration of approaches that promote active learning strategies, such as generative activities. Moreover, the study underscores the importance of developing instructional strategies that enhance students' control perception, which is crucial for fostering effective learning. Beyond the benefits of example generation shown in the present study, it may also realize its potential in interactive and collaborative contexts. Nevertheless, the body of research remains limited, and this study raises further questions about when and why instruction followed by self-generated or provided examples is most beneficial for learning. In particular, the role of perceived learning control and other potential mechanisms—especially motivational/volitional factors—warrants further investigation. Given the relative paucity of studies on illustrative examples compared to worked examples, example-based learning research could benefit from a more integrative approach that connects these currently isolated strands of research. Overall, our work provides an empirical foundation for further studies on generative and example-based approaches and highlights the need for continued exploration of how these approaches can best be implemented in teacher education.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11251-024-09694-4.

**Funding** Open Access funding enabled and organized by Projekt DEAL.



## References

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*(4), 774–783. https://doi.org/10.1037/0022-0663.95.4.774

Balch, W. R. (2005). Elaborations of introductory psychology terms: Effects on test performance and subjective ratings. *Teaching of Psychology, 32*, 29–34. https://doi.org/10.1207/s15328023top3201_7

Beaujean, A. A. (2014). *Latent variable modeling using R. A step-by-step guide*. Routledge. https://doi.org/10.4324/9781315869780

Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin, 145*(11), 1029–1052. https://doi.org/10.1037/bul0000209

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118–1133. https://doi.org/10.1037/a0019902

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*, 219–243. https://doi.org/10.1080/00461520.2014.965823
Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2009). Combining shared control with variability over surface features: Effects on transfer test performance and task involvement. *Computers in Human Behavior, 25*(2), 290–298. https://doi.org/10.1016/j.chb.2008.12.009
Dagenais, C., Lysenko, L., Abrami, P. C., Bernard, R. M., Ramde, J., & Janosz, M. (2012). Use of research-based information by school practitioners and determinants of use: A review of empirical research. *Evidence & Policy, 8*(3), 285–309. https://doi.org/10.1332/174426412x654031
Dickhäuser, O., Schöne, C., Spinath, B., & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept [The academic self-concept scales]. *Zeitschrift Für Differentielle und Diagnostische Psychologie, 23*(4), 393–405. https://doi.org/10.1024//0170-1789.23.4.393
Dornisch, M., Sperling, R. A., & Zeruth, J. A. (2011). The effects of levels of elaboration on learners' strategic processing of text. *Instructional Science, 39*(1), 1–26. https://doi.org/10.1007/s11251-009-9111-z
Ferguson, L. E. (2021). Evidence-informed teaching and practice-informed research. *Zeitschrift Für Pädagogische Psychologie, 35*(2–3), 199–208. https://doi.org/10.1024/1010-0652/a000310
Finney, C., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modelling—a second course* (pp. 439–492). Information Age Publishing.
Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*(2), 50. https://doi.org/10.1007/s10648-023-09769-7
Griffin, M. M. (1993). Do student-generated rational sets of examples facilitate concept acquisition? *The Journal of Experimental Education, 61*(2), 104–115. https://doi.org/10.1080/00220973.1993.9943855
Guerrero, T. A., Griffin, T. D., & Wiley, J. (2024). The effects of generating examples on comprehension and metacomprehension. *Journal of Experimental Psychology: Applied, 30*(2), 318–330. https://doi.org/10.1037/xap0000490
Hamilton, R. (1989). The effects of learner-generated elaborations on concept learning from prose. *The Journal of Experimental Education, 57*(3), 205–217. https://doi.org/10.1080/00220973.1989.10806506
Hamilton, R. (1990). The effect of elaboration on the acquisition of conceptual problem-solving skills from prose. *The Journal of Experimental Education, 59*(1), 5–17. https://doi.org/10.1080/00220973.1990.10806547
Hamilton, R. J. (1997). Effects of three types of elaboration on learning concepts from text. *Contemporary Educational Psychology, 22*(3), 299–318. https://doi.org/10.1006/ceps.1997.0935
Hamilton, R. J. (1999). The role of elaboration within a text processing and text adjunct context. *British Journal of Educational Psychology, 69*(3), 363–376. https://doi.org/10.1348/000709999157770
Hamilton, R. J. (2004). Material appropriate processing and elaboration: The impact of balanced and complementary types of processing on learning concepts from text. *British Journal of Educational Psychology, 74*(Pt 2), 221–237. https://doi.org/10.1348/000709904773839851
Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414–434. https://doi.org/10.1037/0022-0663.90.3.414
Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.00998
Khasawneh, Y. J. A., Khasawneh, N., & Khasawneh, M. A. (2024). Exploring the long-term effects: Retention and transfer of skills in gamified learning environment. *International Journal of Data and Network Science., 8*(1), 195–200. https://doi.org/10.5267/j.ijdns.2023.10.004
Kiemer, K., & Kollar, I. (2021). Source selection and source use as a basis for evidence-informed teaching. *Zeitschrift Für Pädagogische Psychologie, 35*(2–3), 127–141. https://doi.org/10.1024/1010-0652/a000302
Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psych, 9*, 2207–2230. https://doi.org/10.4236/psych.2018.98126
LeFevre, J.-A., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction, 3*(1), 1–30. https://doi.org/10.1207/s1532690xci0301_1
Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education, 10*(3), 142–152. https://doi.org/10.1111/mbe.12117
Micallef, A., & Newton, P. M. (2022). The use of concrete examples enhances the learning of abstract concepts: A replication study. *Teaching of Psychology*. https://doi.org/10.1177/00986283211058069
Neuweg, G.H. (2014). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen [The knowledge of knowledge facilitators. Problems, findings, and

perspectives of research on teacher knowledge.]. In E. Terhart, H. Bennewitz & M. Rothland (Edts.), Handbuch der Forschung zum Lehrerberuf (pp. 583–614). Münster; New York: Waxmann.
Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134*(2), 270–300. https://doi.org/10.1037/0033-2909.134.2.270
Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9
Perry, R. P., Hladkyj, S., Pekrun, R. H., & Pelletier, S. T. (2001). Academic control and action control in the achievement of college students: A longitudinal field study. *Journal of Educational Psychology, 93*(4), 776–789. https://doi.org/10.1037/0022-0663.93.4.776
Perry, R. P., Hall, N. C., & Ruthig, J. C. (2005). Perceived (academic) control and scholastic attainment in higher education. *Higher education: Handbook of theory and research* (pp. 363–436). Berlin: Springer-Verlag.
R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Austria.
Rawson, K. A., & Dunlosky, J. (2016). How effective is example generation for learning declarative concepts? *Educational Psychology Review, 28*, 649–672. https://doi.org/10.1007/s10648-016-9377-z
Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review, 27*, 483–504. https://doi.org/10.1007/s10648-014-9273-3
Reigeluth, C. M., Merrill, M. D., Wilson, B. G., & Spiller, R. T. (1980). The elaboration theory of instruction: A model for sequencing and synthesizing instruction. *Instructional Science, 9*(3), 195–219. https://doi.org/10.1007/BF00177327
Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? *Zeitschrift Für Pädagogische Psychologie, 36*(4), 217–231. https://doi.org/10.1024/1010-0652/a000345
Renkl, A., Mandl, H., & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist, 31*(2), 115–121. https://doi.org/10.1207/s15326985ep3102_3
Renkl, A. (2017). Instruction based on examples. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 325–348, 2nd ed.). New York, NY: Routledge.
Roediger, H. L., III., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003
Rosseel, Y. (2012). lavaan: An r package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02
Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68
Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction, 58*, 161–172. https://doi.org/10.1016/j.learninstruc.2018.06.006
Shuell, T. J. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology, 13*(3), 276–295. https://doi.org/10.1016/0361-476X(88)90027-6
Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592
Stark, L., Malkmus, E., Stark, R., Brünken, R., & Park, B. (2018). Learning-related emotions in multimedia learning: An application of control-value theory. *Learning and Instruction, 58*, 42–52. https://doi.org/10.1016/j.learninstruc.2018.05.003
Steininger, T. M., Wittwer, J., & Voss, T. (2022). Classifying examples is more effective for learning relational categories than reading or generating examples. *Instructional Science, 50*, 771–788. https://doi.org/10.1007/s11251-022-09584-7
Steininger, T. M., Wittwer, J., & Voss, T. (2024). Learning relational categories: Benefits of blocking, classification, and subject-specific examples. *Psychology Learning & Teaching. Advance Online Publication.* https://doi.org/10.1177/14757257241248155
Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*(1), 1–16. https://doi.org/10.1007/s11423-019-09701-3
Thomm, E., Sälzer, C., Prenzel, M., & Bauer, J. (2021). Predictors of teachers' appreciation of evidence-based practice and educational research findings. *Zeitschrift Für Pädagogische Psychologie, 35*(2–3), 173–184. https://doi.org/10.1024/1010-0652/a000301

van Merriënboer, J. J. G., Clark, R. E., & de Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-Model. *Educational Technology Research and Development, 50*(2), 39–61. https://doi.org/10.1007/BF02504993
Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology, 103*(4), 952–969. https://doi.org/10.1037/a0025125
Voss, T. (2022). Not useful to inform teaching practice? Student teachers hold skeptical beliefs about evidence from education science. *Frontiers in Education*, *7*, Article 976791. https://doi.org/10.3389/feduc.2022.976791
Wagner, K., Klein, M., Klopp, E., & Stark, R. (2016). Förderung anwendbaren bildungswissenschaftlichen Wissens anhand kollaborativem Lernen aus Fehlern [Facilitation of applicable knowledge in teacher education by collaborative learning from instructional errors]. *Unterrichtswissenschaft, 44*, 373–390.
Waldeyer, J., Heitmann, S., Moning, J., & Roelle, J. (2020). Can generative learning tasks be optimized by incorporation of retrieval practice? *Journal of Applied Research in Memory and Cognition, 9*(3), 355–369. https://doi.org/10.1016/j.jarmac.2020.05.001
Wilkes, T., & Stark, R. (2023). Probleme evidenzorientierter Unterrichtspraxis. Anregungen und Lösungsvorschläge. [Problems of an evidence-oriented educational practice. Suggestions and solution approaches]. *Unterrichtswissenschaft, 51*, 289–313. https://doi.org/10.1007/s42010-022-00150-1
Wissman, K. T., Zamary, A., Rawson, K. A., & Dunlosky, J. (2023). Enhancing declarative concept application: The utility of examples as primary targets of learning. *Journal of Experimental Psychology: Applied, 29*(2), 341–357. https://doi.org/10.1037/xap0000432
Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345–376. https://doi.org/10.1207/s15326985ep2404_2
Zamary, A., & Rawson, K. A. (2018). Which Technique is most effective for learning declarative concepts—Provided examples, generated examples, or both? *Educational Psychology Review, 30*, 275–301. https://doi.org/10.1007/s10648-016-9396-9
Zamary, A., Rawson, K. A., & Dunlosky, J. (2016). How accurately can students evaluate the quality of self-generated examples of declarative concepts? Not well, and feedback does not help. *Learning and Instruction, 46*, 12–20. https://doi.org/10.1016/j.learninstruc.2016.08.002