



Promoting pre-service teachers' knowledge integration from multiple text sources across domains with instructional prompts

Inka Sara Hähnlein¹ · Pablo Pirnay-Dummer¹

Accepted: 19 February 2024
© The Author(s) 2024

Abstract

Multiple document comprehension and knowledge integration across domains are particularly important for pre-service teachers, as integrated professional knowledge forms the basis for teaching expertise and competence. This study examines the effects of instructional prompts and relevance prompts embedded in pre-service teachers' learning processes on the quality their knowledge integration in multiple document comprehension across domains. 109 pre-service teachers participated in an experimental study. They read four texts on "competencies" from different knowledge domains and wrote a text on a given scenario. Experimental group 1 was aided with instructional and relevance prompts, while experimental group 2 received only relevance prompts. The control group received no prompting. Perceived relevance of knowledge integration was assessed in a pre-post-test. Pre-service teachers' separative and integrative learning, epistemological beliefs, metacognition, study-specific self-concept, and post-experimental motivation were assessed as control variables. Participants' texts were analyzed concerning knowledge integration by raters and with computer linguistic measures. A key finding is that combined complex prompting enhances pre-service teachers perceived relevance of knowledge integration. This study found effects of prompting types on the pre-service teachers' semantic knowledge structures. Implications for transfer are discussed.

Keywords Knowledge integration · Multiple-document comprehension · Pre-service teacher education · Prompting

✉ Inka Sara Hähnlein
inka.haehnlein@paedagogik.uni-halle.de

Pablo Pirnay-Dummer
pablo.pirnay-dummer@paedagogik.uni-halle.de

¹ Faculty of Philosophy III, Institut of Pedagogy, Educational Psychology, Martin-Luther-University Halle-Wittenberg, Halle, Germany

Introduction

Multiple document comprehension and knowledge integration across domains are especially important for pre-service teachers, as they are necessary for forming integrated professional knowledge. This in turn is the basis for the professional competence of future teachers (Lehmann, 2020a, 2020b). Since pre-service teachers find it difficult to integrate knowledge, and teacher education is hardly designed to help in the integration of knowledge (Hudson & Zgaga, 2017), pre-service teachers are in need of support. Instructional strategies such as cognitive prompts have been found to be effective aids (Lehmann, Pirnay-Dummer, et al., 2019a, 2019b; Lehmann, Rott, et al., 2019a, 2019b). However, more empirical research is needed on how to promote cognitive knowledge integration across domains.

The present study aims to promote and facilitate pre-service teachers' knowledge integration from multiple text sources across domains with two kinds of cognitive instructional strategies. In our experimental study, pre-actional relevance prompts and pre-actional instructional prompts combined with relevance prompts are embedded at multiple time points in pre-service teachers' self-regulated learning processes. The learning process includes a reading phase and a writing task designed to integrate and transfer what has been read. We hypothesize that the types of prompts will influence the three criteria of pre-service teachers' text quality, knowledge structure, and perceived relevance of knowledge integration in various ways.

Theoretical background

Knowledge integration across disciplines as a precondition for teaching expertise

For pre-service teachers, pedagogical knowledge (PK), pedagogical content knowledge (PCK), and content knowledge (CK) form the core of their professional competence as future teachers (Shulman, 1986; Voss et al., 2011). Technological, organizational and counseling knowledge as well as self-regulation skills, motivational orientations, beliefs, and values are also part of teachers' professional competence (Model of Professional Competence of Teachers, Baumert & Kunter, 2006; Technological Pedagogical Content Knowledge, TPACK, Mishra & Koehler, 2006; Koehler & Mishra, 2008; Mishra, 2019; Krauskopf et al., 2020).

These aspects constitute the professional knowledge of teachers in training and are the basis for successfully teaching a specific subject (i.e., analyzing, planning, designing, developing, and evaluating instruction and instructional interactions; Seel et al., 2017). Teacher education includes study components from educational sciences, didactics, pedagogy, and the teaching subjects. Text is still by far the main source of information in academia (Pirnay-Dummer, 2020). For pre-service teachers, these texts come from different disciplines. Pre-service teachers are thus confronted with texts that take different perspectives on the same topics, and integrating these perspectives is not a simple straightforward task. Texts provide different domain-specific rationales for specific instructional decisions that interact with each other in a complex way (Lehmann, 2020a; Pirnay-Dummer, 2020).

Knowledge integration and multiple-document comprehension

According to Lehmann (2020a), knowledge integration is defined in two ways, as first- and second-order knowledge integration. First-order knowledge integration, as a form of constructive learning, is the active linking, merging, distinguishing, organizing, and structuring of knowledge structures into a coherent model (Lehmann, 2020a, 2020b; Linn, 2000; Schneider, 2012). Second-order knowledge integration, as a form of knowledge application, is the simultaneous transfer of knowledge from different domains with the goal of reaching a suitable problem solution (Graichen et al., 2019; Janssen & Lazonder, 2016; Lehmann, 2020a, 2020b).

Integrating knowledge from text across disciplines, a competency pre-service teachers are expected to possess is a complex process that relies on single-text comprehension (Construction-Integration Model; Kintsch, 1988, 1998; Trevors et al., 2016; van Dijk & Kintsch, 1983) but furthermore requires multiple document comprehension (Document Model Framework; Braasch et al., 2012; Britt & Rouet, 2012; Britt et al., 2018; Perfetti et al., 1999; Rouet, 2006).

The construction of a coherent integrated model of multiple texts requires readers to form an integrated mental model of the content of the texts (integrated mental model), including contradictions and their possible or impossible resolutions, as well as representations of the text sources and how these sources are related to each other as an *intertext model* (Bråten & Braasch, 2018; Britt & Rouet, 2012; Perfetti et al., 1999; Rouet, 2006).

Relevance of knowledge integration for teaching

Empirical evidence shows the importance of knowledge integration for pre-service teachers: The level of knowledge integration influences their degree of expertise and professional competence as teachers later on (Baumert & Kunter, 2006; Bromme, 2014; Graichen et al., 2019; Janssen & Lazonder, 2016; König, 2010; Lehmann, 2020a). Teachers' knowledge integration is also positively related to the learning performances of students (Hill et al., 2005).

For pre-service teachers and in-service teachers, successful knowledge integration means being able to apply integrated knowledge to solve complex (ill-defined) problems (Jonassen, 2012), such as instructional planning (Brunner et al., 2006; Dörner, 1976; Lehmann, 2020a; Norton et al., 2009), and to make informed instructional decisions, taking into account multiple perspectives and their complex interaction (Lehmann, 2020a).

Although successful knowledge integration is particularly important for prospective teachers, teacher education is hardly designed to link subject areas. Subject knowledge, didactics, and educational science are taught separately (Ball, 2000; Blömeke, 2009; Darling-Hammond, 2006; Hudson & Zgaga, 2017). As Zeeb et al. (2020, p. 202) point out, this way of knowledge acquisition increases the risk that students will develop structural deficits in their knowledge, in the sense of compartmentalization of knowledge (e.g., Whitehead, 1929), which in turn explains the fragmented, inert knowledge (Renkl et al., 1996). Accordingly, the professional knowledge present in pre-service teachers and in-service teachers tends to be fragmented and poorly integrated. There is hardly any systematic linking across subject areas and faculty boundaries. The principle of separating subject knowledge, subject didactics, educational science, and pedagogy prevails (Ball, 2000; Darling-Hammond, 2006; Graichen et al., 2019; Harr et al., 2014; Janssen & Lazonder, 2016).

Since integrating knowledge is an important but complex task, and teacher education is hardly designed to initiate and train integration of knowledge (Hudson & Zgaga, 2017), pre-service teachers are in need of support.

Supporting pre-service teachers' knowledge integration and multiple document comprehension

There are different approaches to the promotion of knowledge integration:

The MD-TRACE model ("Multiple-Document Task-Based Relevance Assessment and Content Extraction") describes how readers represent their goals by forming a task model in multiple-document comprehension (Rouet, 2006; Rouet & Britt, 2011). The reading context is vital for multiple-document processing. According to the RESOLV theory, readers initially construct a model of the reading context that influences the interpretation of the task at hand, for instance the interpretation of writing tasks (Rouet & Britt, 2011; Rouet et al., 2017). There is ample evidence to suggest that specific writing tasks can elicit changes in learning activities (Wiley & Voss, 1996, 1999; Bråten & Strømsø 2009, 2012; Lehmann, Rott, et al., 2019a, 2019b). Specific writing tasks can promote integrated understanding by stimulating elaborative processes of knowledge-transforming rather than simple knowledge-telling (Scardamalia & Bereiter, 1987, 1991). Writing a summary of multiple documents or answering overarching questions may also promote knowledge integration (Britt & Sommer, 2004; Wiley & Voss, 1999). Tasks asking pre-service teachers to combine information from multiple sources can foster the integrated application of knowledge (Graichen, et al., 2019; Harr et al., 2015; Lehmann, Rott, et al., 2019a, 2019b; Wäschle et al., 2015).

A growing body of empirical evidence shows that prompts are effective instructional strategies for supporting knowledge integration across domains (Lehmann, Pirnay-Dummer, et al., 2019a, 2019b; Lehmann, Rott, et al., 2019a, 2019b). Implemented as statements, focus questions, incomplete sentences, or relevance instructions, among other formats, they promote the cyclical process of self-regulated learning (SRL) (Schiefele & Pekrun, 1996; Zimmerman, 2002). Prompts can be embedded in the pre-actional, actional, or post-actional phase of SRL. They elicit the use of cognitive, metacognitive, and motivational learning strategies that promote learning at the respective level (Bannert, 2009; Ifenthaler, 2012; Lehmann et al., 2014; Lehmann, Rott, et al., 2019a, 2019b; Reigeluth & Stein, 1983; Zimmerman, 2002).

Implemented pre-actionally, that is, provided prior to learning, prompts can assist learners in constructing an appropriate task model (Ifenthaler & Lehmann, 2012; Lehmann et al., 2014). Empirical evidence demonstrates that pre-actional cognitive prompts promote an integrated deep understanding across core domains of professional knowledge of pre-service teachers (Lehmann, Rott, et al., 2019a, 2019b; Wäschle et al., 2015).

Relevance instruction enhances integrated knowledge in pre-service teachers by encouraging the use of integrative strategies (e.g. Zeeb et al., 2020). Relevance prompts can be implemented either specifically or independently. They can emphasize the relevance of the specific learning content (Cerdán & Vidal-Abarca, 2008) or learning task (Gil et al., 2010), or they can be implemented independently and refer to the relevance of knowledge integration in general (Zeeb, Biwer, et al., 2019; Zeeb et al., 2020). Zeeb et al. (2020) argue in favor of the independent implementation, since knowledge integration is important for teacher education across domains and not just within specific subjects or topics. Repeated

relevance instructions have been shown to be superior to one-time instructions (Zeeb et al., 2020).

The above described theoretical foundation shows the importance of multiple document comprehension and knowledge integration across domains for pre-service teachers and that prompts might be used to help them.

Objective

In our experiment, we investigated how to promote and facilitate pre-service teachers' knowledge integration from multiple text sources across domains. The design examined the effects of two different kinds of prompts embedded in the students' learning processes at three different points in an experimental design.

Research question

Do instructional and relevance prompts embedded in the learning process promote pre-service teachers' knowledge integration from multiple texts across domains?

The evidence described above indicates that providing students with *pre-actional instructional prompts* should lead to an increased quality of knowledge integration. The same is to be expected for repeated *pre-actional relevance prompts*. Our experiment combined the two types of prompts and embedded them at different stages in the learning process, in experimental contrast to providing relevance prompts only. A control group was provided with no support, as this is still often the standard in pre-service teacher training.

We hypothesized that the combined prompts would lead to higher rated text quality (dependent variable 1, DV1) than relevance prompts alone. We also hypothesized that support from relevance prompts alone would be better than no support at all.

Moreover, we hypothesized that there would be systematic differences between the three groups both in terms of knowledge structure and knowledge semantics (dependent variable 2, DV2) when compared with the knowledge model of the source material.

In addition, we hypothesized that *repeated pre-actional relevance prompts* would lead to an increase in perceived relevance (dependent variable 3, DV3).

We assessed pre-service teacher students' perceived integration and separation learning in teacher education, epistemological beliefs, metacognition, study-specific self-efficacy, and post-experimental motivation as control variables, because of their relevance for self-regulated knowledge integration (Barzilai & Strømsø, 2018; Lehmann, 2022).

Methods

Participants

The experiment was conducted with $N = 109$ of 119 pre-service teacher students (see Analysis and Results for explanation of dropout) who attended one of four courses on research methods and statistics in the 2021/22 fall term at a German university (81 females, 27 males, 1 n/a; age: $M = 22.46$, $SD = 3.23$). Of the 108 student teachers who specified their school type, most are studying elementary school teaching (45), while 36 are studying high school teaching, 17 secondary school teaching, 8 special education, and only 1

elementary and special education. The most common major subject is German (43), followed by mathematics (23), history (9), biology (6), English (5), sports (5) and others. The most chosen second subject are mathematics (29), German (24), biology (8), English (5). The participants have been studying their major subject for an average of about six semesters ($M = 5.9$, $SD = 2.3$).

Procedure and design

This experimental intervention study has a cross-sectional control group design. Figure 1 shows the experimental procedure and design. First, the pre-service teacher students were informed about the study at the beginning of the term (calendar week 40) in the introductory sessions of the courses on research methods. Participation was voluntary with no consequences for not participating. All students enrolled in the course chose to participate in this study giving their informed consent. The participants were anonymized by means of randomly assigned codes known only to themselves. Then, the participants were randomly assigned to three experimental conditions: experimental group 1 (EG1), experimental group 2 (EG2), and control group (CG). All assessments and the intervention took place at course time in the course room to make it as easy as possible for the students to participate. Students who dropped out of the course automatically terminated their participation in this study.

In weeks two and three of the term (calendar week 41/42), we collected the students' demographics, and as control variables assessed their epistemological beliefs (Students' Epistemological Beliefs; StEB; Hähnlein, 2018), metacognition in the learning process (MILP; Hähnlein & Pirnay-Dummer, 2019), study-specific self-efficacy (WIRKSTUD; Schwarzer & Jerusalem, 2003), as well as their self-reported separative and integrative learning in teacher education (SILTE; Lehmann et al.; 2020). All data collection was implemented online on Limesurvey.

To avoid test fatigue in the participants, the actual intervention was not conducted until several weeks after the control variables were collected. For all participants, the time

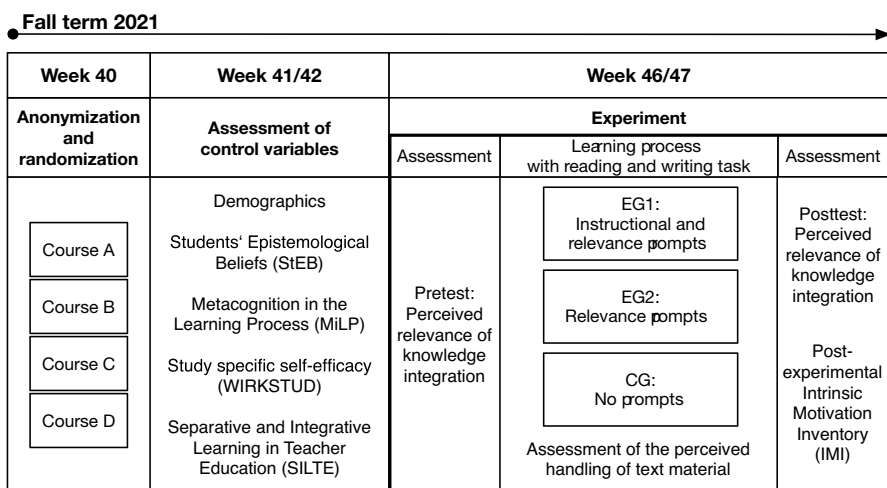


Fig. 1 Experimental procedure with three experimental conditions in an experimental intervention study with cross-sectional design

interval between pre-survey and experiment was the same. The experiment was conducted six to seven weeks (calendar week 46/47) into the term and took approximately 90 to 105 minutes for CG and EG2, but about 120 minutes for EG1 to account for the prolonged instruction.

The experiment started with a 3-minute pre-test regarding students' perceived relevance of knowledge integration (Figure 1). After that, to initiate a learning process, all participants received a learning task with an instructional part, a reading and a writing task (Figure 1). Participants were instructed to work through the learning process and remaining test procedure on their own time and were allowed to leave the experiment when they were finished, but not before 90 minutes had passed (normal course duration).

For all participants, the reading task consisted of four texts about "the concept of competence in teaching" from different subject domains of teacher education (8 pages in total). We recorded the time participants spent studying the text material including reading time ($M = 38$, $SD = 17$, $Min = 14$, $Max = 78$). Students' perceived handling of the text material was evaluated right after the reading phase (Text Material Questionnaire I; Deci & Ryan, n.d.; German adaption). This took just 2 minutes. The writing task was on a fictional scenario that required integrating the knowledge from all four source texts to derive implications for the application of the knowledge.

The participants of the control group (CG: no prompts) received just organizational information in the instructional phase of the learning process and no aid regarding knowledge integration for the reading and writing task (Figure 1).

The participants of experimental group 2 (EG2: relevance prompts) received *relevance prompts* in verbal and written form embedded in their learning processes in the instructional phase as well as the reading and writing phase. Experimental group 1 (EG1: relevance prompts and instructional prompts) received both *relevance prompts* and *instructional prompts* (Figure 1). Both types of prompts were embedded in the students' learning process at three time points: the instructional phase, the reading phase, and the writing phase.

Following the learning process, we again assessed students' perceived relevance of knowledge integration (3 minutes) as well as their post-experimental intrinsic motivation (Intrinsic Motivation Inventory, IMI; Deci & Ryan, n.d.; German version, 12 minutes). All survey instruments used are introduced below in the section *Survey Instruments*.

Instructional prompts in experimental group 1

In the instructional phase, the participants of experimental group 1 received a 10-minute PowerPoint-based introduction to knowledge integration to stimulate their pre-flexion prior to reading. The introduction used the example of lesson planning to outline how knowledge integration works, how it works, and why it is important for future teachers. This served as a pre-actional cognitive prompt.

To support the reading phase, focus questions were developed in our research team specifically to the text material at hand. The participants received the following focus questions related to knowledge integration to apply to the texts to be read (translated from German):

- What are similarities and differences in the understanding of competence between the texts?
- What level does the knowledge from the different texts refer to?

- o Does the knowledge refer to what competencies are?
 - o Does the knowledge relate to abstract or specific objectives? (What is to be achieved?)
 - o Does the knowledge relate to application? Does it relate to the process of how to accomplish something?
 - o Does the knowledge relate to why? (Why does something work this way and not another way?)
- How do the competency perspectives interact with each other? Do the knowledge contents and their levels complement each other or do contradictions arise? (Bridges between texts?)
 - Can different things be derived from the different perspectives for application?
 - What can be derived for the application from the integrated impression of all 4 texts?

Furthermore, the participants were asked to model the interrelation between the texts, for instance by drawing a mind map. The focus questions served as pre-actional cognitive prompts for the reading task, while the modeling served as an actional cognitive prompt for the reading phase and a pre-actional cognitive prompt for the writing phase.

Before writing another pre-actional prompt was given: The participants in this group were explicitly asked to use their elaborations from the reading phase while writing and to explicitly connect the knowledge instead of just summarizing it.

Relevance prompts in experimental group 1 and 2

During instruction, the relevance prompt was provided to the participants via an oral explanation of the importance of knowledge integration for their future teaching proficiency and the reasons for it, supported by an anthology on knowledge integration research held up during the presentation of a PowerPoint slide. This served as pre-actional prompt.

In both the reading and the writing phase, the participants were again reminded in writing of the relevance of knowledge integration.

For both experimental groups, the task sheet remained with the participants so that they could access the prompts even while performing the task.

The control group received neither relevance prompts nor instructional prompts but just the reading and writing task.

Material

Reading material

The reading material was four selected texts on the concept of competence in teaching. The texts come from different disciplines and are all academic in source and nature (educational science/humanities, educational psychology, didactics, and policy-making). The texts were selected and discussed beforehand by an interdisciplinary team both for their relevance within each field and for their potential for not being too easy to integrate. However, since the four texts are from different disciplines, each of them takes a different professional perspective on the topic.

Text 1 takes the perspective of educational science or pedagogy. It is taken from a textbook on the introduction to educational science and two pages long (Textbook: Thompson, 2020, pp. 131–133). This text is about how the concept of competence is defined from the perspective of competence research.

Text 2 takes the point of view of educational psychology. It is two and a half pages long and about what distinguishes the concept of competence from established categories such as ability, skill, or intelligence. (Article: Wilhelm & Nicolaus, 2013, pp. 23–26)

Using the example of learning to read, text 3 deals with the distinction between different levels of competence (Textbook: Philipp, 2012, pp. 11–15). It is a text from the didactics and one and a half pages long.

Text 4 is a curricular description from the Standing Conference of the Ministers of Education and Cultural Affairs (Kultusministerkonferenz, 2009, pp. 1–5). It takes an educational policy perspective on the subject matter and is two and a half pages long. This text is about the competency level model for the educational standards in the competency area speaking and listening for secondary school.

Writing task

The writing task was part of a fictional scenario requiring participants to integrate knowledge from the four source texts and to draw integrated conclusions for its application to teaching in order to help a friend in need. The scenario with task (translated from German) read as follows:

During your school internship, a future colleague has guided you through many a challenging situation thanks to her professional experience and appreciative nature. This teacher, Monique Gerber, recently turned to you and somewhat bashfully told you that she herself is currently facing a rather challenging situation. She has a school evaluation coming up next week. In itself, this is not a problem for Monique. However, she has learned in advance that dealing with competence and its scientific foundation in teaching is a central theme of the evaluation. She says that the academic discussion has been going on for far too long and that she would like you to give her an informative summary of the topic of competence:

- What should she look for when teaching?
- How should she justify things?
- How does she relate what she does well in class to existing scientific knowledge?

A little flattered, and knowing of Monique's distress, you set out to help her.

Task:

Write a text yourself on the basis of the four short texts on the topic of competence. Your text for Monique should explain step by step the current scientific understanding of competence and show her how it can be used to help her plan and design lessons.

(The text should be written in complete sentences. It must be at least 400 words.)

This complex task requires both knowledge-telling and knowledge-transforming (Scardamalia & Bereiter, 1987, 1991). An integrated mental model of the text sources (Bråten & Strømsø, 2009; Wiley & Voss, 1999) is needed to derive implications for lesson planning.

Survey instruments

The StEB Inventory (Hähnlein, 2018) is designed to assess pre-service teachers' epistemological beliefs. The instrument development is based on the theoretical conceptualization of the epistemological belief system by Schommer (1994), the core dimensions by Hofer and Pintrich (1997) and Conley et al. (2004), as well as the *Integrative Model for Personal Epistemology (IM)* (Bendixen & Rule, 2004; Rule & Bendixen, 2010) to explain the mechanism of change, and the *Theory of Integrated Domains in Epistemology (TIDE)* (Muis et al., 2006) to explain the context dependency and discipline specificity of epistemological beliefs. The StEB questionnaire consists of four subscales: *beliefs about the simplicity of knowledge*, the *absoluteness of knowledge*, the *multimodality of knowledge*, and the *development of knowledge*. The questionnaire consists of 26 items. Agreement with the statements is indicated on a 5-point Likert scale (from *does not apply at all* to *completely applies*).

The MiLP Inventory (Hähnlein & Pirnay-Dummer, 2019) assesses students' metacognitive activities in the form of learning judgments. The instrument development is based on the theoretical model of Nelson and Narens (1990, 1994). It distinguishes between metacognitive monitoring and control as well as the three phases of learning: knowledge acquisition, retention, and retrieval. The questionnaire consists of 33 items and six subscales. Four subscales concern the metacognitive activities in the knowledge acquisition phase. Two each concern metacognitive monitoring and metacognitive regulation. One subscale concerns metacognitive observation in the retention phase and one that of knowledge retrieval. The response scale has a 5-level Likert format (from *does not apply* to *does apply*). The six subscales are as follows:

- Anco: Assesses a learners' ability to regulate his/her learning in the phase of knowledge acquisition by means of adequate learning strategies. (10 items)
- Abmo: Assesses a learners' ability to monitor his/her retrieval of knowledge in a way that her/she is able to successfully remember the learning content. (8 items)
- Anmo: Assesses a learners' ability to monitor his/her knowledge acquisition by means of assessing the difficulty of the learning content. (5 items)
- Akco: Assesses a learners' ability to regulate his/her knowledge acquisition in a way that he/she is able to successfully differentiate between important and unimportant learning content. (3 items)
- Bemo: Assesses a learners' ability to monitor his/her retention of knowledge in a way that he/she is able to remember the learning content. (4 items)
- Akmo: Assesses a learners' ability to monitor his/her knowledge acquisition in a way that he/she is able to figure out if the knowledge acquisition was successful. (3 items)

The SILTE Short Scales (Lehmann et al., 2020) are used to measure the self-reported knowledge integration of pre-service teacher students in teacher education across domains. With its two dimensions, it measures integrative learning with 7 items and separative learning with 5 items. The two scales have a five-point response format (ranging from *does not apply at all* to *fully applies*). According to Lehmann et al. (2020, p. 156), the theoretical foundation of the SILTE questionnaire is the model of knowledge building (e.g. Chan et al., 1997; Scardamalia & Bereiter, 1994, 1999), which can be assigned to the constructivist approaches to strategic learning. In addition, the

questionnaire is based on the concepts of cognitive fragmentation and knowledge integration in teacher education and learning to teach (e.g. Ball, 2000; Darling-Hammond, 2006; Lehmann, 2020b).

Study-specific self-efficacy is assessed using the WIRKSTUD scale (Schwarzer & Jerusalem, 2003). It is one-dimensional and has 7 items with a four-point rating scale (*does not apply at all, hardly applies, applies, applies completely*). The conception of the scale is based on Bandura's (1978) social-cognitive learning theory and the concept of positive situation-action expectations contained therein.

The Intrinsic Motivation Inventory (IMI) is a multidimensional measurement that comes in different versions. It is intended to assess "participants' subjective experience related to a target activity in laboratory experiments" (Deci & Ryan, n.d., p. 1). The Post-Experimental Intrinsic Motivation Inventory (Deci & Ryan, n.d.) originally consists of 45 items and seven scales that can be selected according to the requirements of the experimental setting. In our study, the following six scales were used: *Interest/enjoyment* (7 items), *perceived competence* (6 items), *effort/importance* (5 items), *pressure/tension* (5 items), *perceived choice* (7 items), and *value/usefulness* (7 items). The scale *relatedness* (8 items) was not used in this study. A five-point response format (ranging from *does not apply at all* to *fully applies*) were used for all IMI measures.

The Text Material Questionnaire consists of three of the subscales of the IMI questionnaire (Deci & Ryan, n.d.) adapted to text material. It assesses students' interest and pleasure (5 items), felt pressure (2 items), and perceived competence (2 items) in dealing with the text.

The *value/usefulness* (7 items) subscale of the IMI (Deci & Ryan, n.d.) which is adaptable to different content, was used to measure pre-service teachers' *perceived relevance of knowledge integration* in the pretest and posttest.

Table 1 shows the internal consistencies of the survey instruments used. Reliabilities are reported for the current study as well as for the previous development and validation studies. All in all, the reliabilities can be considered acceptable. For individual scales, however, there are very low internal consistencies with values below $\alpha = .70$.

Text rating measure

To score the quality of the participants' texts, we used a rating scheme with three criteria: degree of transfer, validity of conclusions, and degree of integration. Each criterion was rated on a scale of 0 to 3 points. The criteria were initially developed by an expert group of five persons regarding content validity. For this study, a two-person group re-evaluated the criteria, but only minor changes were made and only with respect to the specific content. The text criteria were not revealed to the participants. The rating criteria (translated from German) were as follows:

Degree of transfer (transfers are only given if they can also be derived from the texts read).

- 0: There is no transfer to the action level.
- 1: A transfer to the action level is only made by naming goals. For this purpose, less specified should statements are used (for example, as in "the lessons should be designed in a friendly way" or "the lessons should have a good relationship level").
- 2: Ideas are formulated sporadically (or in an unconnected list form) on how aspects from the transfer can be implemented.

Table 1 Surveys used for the assessment of control variables

Scales (items)	Cronbachs's α (development studies)	Cronbachs's α (current study)
Students' epistemological beliefs (StEB)		
Simplicity of knowledge (6 items)	.70 < α < .82	.69
Absoluteness of knowledge (8 items)		.78
Multimodality of knowledge (7 items)		.68
Development of knowledge (5 items)		.74
Metacognition in the learning process (MiLP)		
Anco (10 items)	.71 < α < .91	.89
Abmo (8 items)		.85
Anmo (5 items)		.81
Akco (3 items)		.74
Bemo (4 items)		.71
Akmo (3 items)		.75
Study-specific self-efficacy (Wirkstud)		
One-dimensional (7 items)	.71 < α < .89	.77
Separative and integrative learning in teacher education (SILTE)		
Separative learning (5 items)	.75	.77
Integrative learning (7 items)	.77	.66
Relevance of knowledge integration		
Value/usefulness (7 items)		Pre .89 Post .88
Text material questionnaire I		
Interest/ enjoyment (5 items)		.71
Pressure (2 items)		.78 ^a
Perceived competence (2 items)		.67 ^a
Post-experimental intrinsic motivation inventory		
Interest/enjoyment (7 items)	.71 < α < .91	.77
Perceived competence (6 items)		.85
Effort/importance (5 items)		.71 ^b
Pressure/tension (5 items)		.89
Perceived choice (7 items)		.89
Value/usefulness (7 items)		.91

^aUsing Spearman-Brown-Coefficient as an estimation of internal consistency for pressure: .63, and perceived competence: .48

^bItem effort2 "I didn't try very hard to do well at this activity". was dropped. Without drop of item: $\alpha = .55$ 95%-CI [.35;.65]

- 3: There are concrete, interrelated derivations from the texts with regard to a realistic lesson design.

Validity of the conclusions

- 0: The conclusions cannot be derived with certainty from the sources (e.g., purely intuitive assumptions).

- 1: Unconnected (e.g., purely abductive) assumptions are present, but at most as a list-like series of unrelated individual statements.
- 2: The conclusions are largely clear from the sources.
- 3: The conclusions emerge unambiguously and deductively from the sources and are logically related to application.

Degree of integration

- 0: Assumptions are treated separately per text.
- 1: The assumptions are treated separately, but, e.g., any contradictions and compatibilities discovered are contrasted, mentioned, and/or discussed.
- 2: The different models are treated together, with reference to each other. They are not worked through sequentially.
- 3: The different models in the texts are processed in an integrated and coherent way, explicitly integrating the areas of knowledge into each other.

Computer linguistic methods

Computer linguistic methods were used in this research project for computational modeling of the semantic knowledge structures contained in both source texts and student texts (Pirnay-Dummer, 2006, 2010, 2014, 2015a, 2015b; Pirnay-Dummer et al., 2010).

Mental model-based (Seel, 1991, 2003) knowledge elicitation techniques have relied on recreating propositional networks from human knowledge (Ifenthaler, 2010; Jonassen, 2000, 2006; Jonassen & Cho, 2008; Pirnay-Dummer, 2015a, 2015b).

The computational linguistic heuristic technology T-MITOCAR (Text-Model Inspection Trace of Concepts and Relations; Pirnay-Dummer, 2006, 2007; Pirnay-Dummer et al., 2010) was developed as a means of automatically analyzing, modeling, visualizing, and comparing the semantic knowledge structure of texts. The approach behind T-MITOCAR is closely based on the psychology of knowledge, knowing, and epistemology (Spector, 2010; Strasser, 2010). Its associative core functions are founded strictly on mental model theory (Gentner & Stevens, 1983; Johnson-Laird, 1983; Pirnay-Dummer et al., 2012; Pirnay-Dummer & Seel, 2018; Seel, 2012; Seel et al., 2013) and on how, when, and why parts of knowledge are reproduced in the semantics of natural language (Evans & Green, 2006; Helbig, 2006; Partee, 2004; Taylor, 2007). The algorithms work through the propositional relations of a text to identify central relations between concepts and build a network (a graph), while always heuristically reconstructing the parts as closely as possible to how human knowledge is constructed, conveyed, and reconstructed through text.

T-MITOCAR can also automatically compare different knowledge structures from text quantitatively using measures based on graph theory (Tittmann, 2010), four structural measures (Table 2), and three semantic measures (Table 3). The measures lead to similarity measures s between zero and one. Tables 2 and 3 only provide an overview for their interpretations, which is necessary to understand them as criteria within this study.

In this research project, we used T-MITOCAR technology to compare the student texts with a reference model of the source texts on the basis of the seven similarity measures. Although we classified the measures into structural and semantic indices, as Tables 2 and 3 show, the similarity indices measure different features and can therefore not be treated like a subsuming scale (e.g., on a test): The measures are not items for the same but for different properties of knowledge graphs. The structural measures indicate *different* properties of

Table 2 Structural measures used for the quantitative comparison of texts (see Pirnay-Dummer, 2020, p. 141)

Measure	Definition
Surface matching (SUR)	The number of links contained in two graphs is compared. This is a simple measure comparing just how many overall propositions are in each graph
Graphical matching (GRA)	The structural ranges of two graphs are compared. This is achieved through comparison of the longest and the shortest path between each two indirectly connected concepts in the graph. A long path could resemble a chain of effects, as often found in causal models, while a short path could be found in reasoning networks, for instance resembling a case frame or simple hierarchy
Gamma matching (GAMMA)	This measure of structural density involves deriving the quotient of the number of concepts per links within each graph and then comparing them to each other for similarity. Usually, individual weak network structures are either represented by a graph that connects everything or by a graph that only connects pairs of terms but otherwise has no hierarchy or network of any kind. A medium density per graph is expected for usable knowledge structures
Structural matching (STRU)	The complete structure of two graphs is algorithmically mapped on all possible network-building blocks, allowing the components of knowledge structure to be traced. Possible reasoning paths are fully reconstructed and compared between two graphs. This similarity measure is particularly important for hypotheses about features of structure (e.g., when one wishes to investigate whether experts would have differently structured knowledge than non-experts)

Table 3 Semantic measures used for the quantitative comparison of texts (see also Pirnay-Dummer, 2020, p. 142)

Measure	Definition
Concept matching (CONC)	This measure compares how many words match between two graphs. This does not say very much in itself, because there can always be a number of reasons for words just being in a network. It can, however, be highly informative together with other measures, especially when the structure and the use of words match
Propositional matching (PROP)	This measure compares fully identical propositions (concept-link-concept) between two graphs. In contrast to concept matching, the words must be linked in the same way. The measure is dependent on concept matching (if no words match, then no propositions can match)
Balanced semantic matching (BSM)	This measure makes it easier to read the interpretation of the dependence between concept matching and propositional matching. For instance, if many concepts but only very few propositions match, this may stand for word spamming without proper knowledge. On the other hand, if there is a partial concept matching and thus a very low propositional matching, then BSM can still be high, resembling a case where a learner might only partly have access to the whole expected content/expertise

structure, whereas structure itself is not a property. The same holds true for the semantic measures.

Analysis and results

The statistical software R (R-Core-Team, 2022) was used for analysis. Of the 119 students enrolled in the four courses at the beginning, five dropped out during term and terminated their participation in this study. Two people were absent on the day of the experiment for health reasons. Three other participants participated in the experiment but did not submit their self-written texts for unknown reasons. Since this was interpreted as a withdrawal of consent, these three datasets were not evaluated. The data of 109 participants were available for further analysis.

First, we checked for differences in the control variables between the experimental conditions using MANOVA and ANOVA. Results from MANOVA showed no significant differences in students' separative and integrative learning in teacher education between the groups (Wilks' $\lambda = 0.97$, $F[4,220] = 0.93$, $p = .45$). Also, no significant differences between the three groups were found in the students' metacognitive abilities (MANOVA, Wilks' $\lambda = 0.88$, $F[12,222] = 1.27$, $p = .24$), epistemological beliefs (MANOVA, Wilks' $\lambda = 0.95$, $F[8,226] = 0.81$, $p = .60$), and study-specific self-efficacy (ANOVA, $F[2,110] = 2.29$, $p = .11$). The participants of the three groups did not perceive the reading material differently (MANOVA, Wilks' $\lambda = 0.95$, $F[6,212] = 0.99$, $p = .43$). For post-experimental intrinsic motivation, no significant differences between the three groups were found (MANOVA, Wilks' $\lambda = 0.79$, $F[12,176] = 1.79$, $p = .053$).

Using ANOVA, we analyzed the group differences in the time (minutes) participants spent studying the text material, to check if the instructional prompts (focus questions) were used. Time spent studying the text material differed significantly between the three groups (ANOVA, $F[2,89] = 238.2$, $p < .001$; $\eta_p^2 = .84$). Participants with the combined prompts spent an average of 63 minutes ($SD = 7.40$) studying the text material, while participants with only relevance prompts (EG2: $M = 29$, $SD = 7.25$, $p < .001$) or no prompts (CG: $M = 27$, $SD = 5.67$, $p < .001$) spent significantly less time.

As a further check on prompt usage, participants of EG1 were asked to hand in their models showing the interrelations between the texts. Three subjects who had not created a model, i.e. had not completed this learning phase, were excluded from further analysis.

Quality of knowledge integration (DV1)

We analyzed the quality of the pre-service teachers' texts (DV1). Two trained raters assessed the student texts ($N = 109$, EG1: $n1 = 37$, EG2: $n2 = 37$, CG: $n3 = 35$) for quality on a scale of 0 to 3 points on the basis of the following criteria: degree of transfer, validity of conclusions, and degree of integration.

Table 4 shows the mean ratings for the criteria per group for both raters. The average points achieved for integration, transfer and conclusions are lower for rater 1. Reviewer one seems to evaluate the student texts more strictly than reviewer two.

Interrater reliability (Kendalls- τ , Type b) was high for transfer ($r_\tau = .52$, $z = 6.08$, $p < .001$) but low for conclusion ($r_\tau = .21$, $z = 2.23$, $p = .02$) and integration ($r_\tau = .13$, $z = 1.60$, $p = .11$), as well as for the overall rating ($r_\tau = .24$, $z = 3.30$, $p < .001$). The ratings were treated separately in the following analysis due to low interrater reliability. All raters

Table 4 Mean and standard deviation for text quality ratings per rater per group

Rater 1				Rater 2			
Criterion	Group	<i>M</i>	<i>SD</i>	Criterion	Group	<i>M</i>	<i>SD</i>
Integration	EG1	0.81	0.88	Integration	EG1	1.43	1.04
	EG2	1.11	0.91		EG2	1.57	1.14
	CG	0.80	0.93		CG	1.94	1.11
Transfer	EG1	1.00	0.88	Transfer	EG1	1.16	0.93
	EG2	0.86	0.67		EG2	1.03	0.96
	CG	0.74	0.66		CG	1.17	0.98
Conclusion	EG1	1.81	0.57	Conclusion	EG1	1.43	0.73
	EG2	1.78	0.48		EG2	1.76	0.80
	CG	1.74	0.56		CG	1.89	0.58

used the whole range of the criteria (0–3) for each item. MANOVA with three rating scores per rater as dependent measures and the experimental conditions as independent measures indicated no significant differences between type of prompting (Wilks' $\lambda = 0.85$, $F[12,202] = 1.43$, $p = .16$).

Computer linguistic analyses (DV2)

Comparison of participants' texts with the reference model of source material

Using T-MITOCAR technology, we combined the four source texts to generate a reference model of their semantic knowledge structures across the respective domains (see Figure 3). Figure 2 shows only a section of the resulting reference model for illustrative purposes because the entire model would be too large.

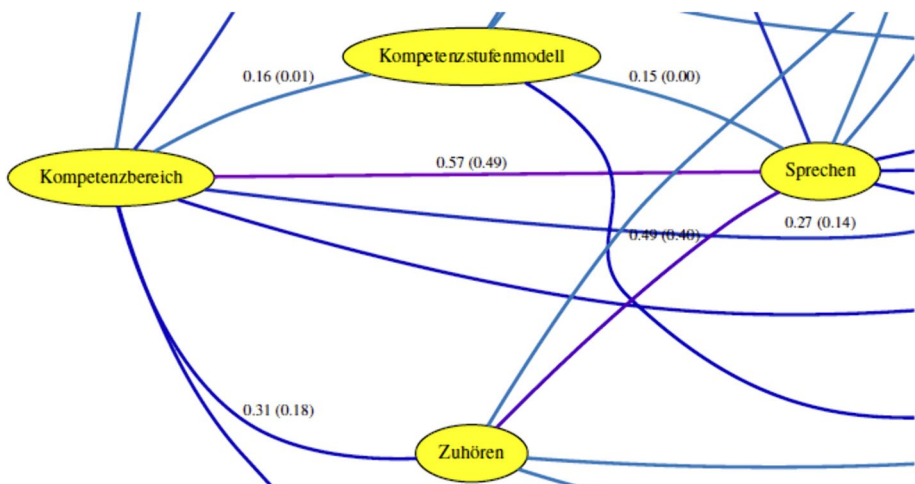


Fig. 2 Section of the reference model of the source material

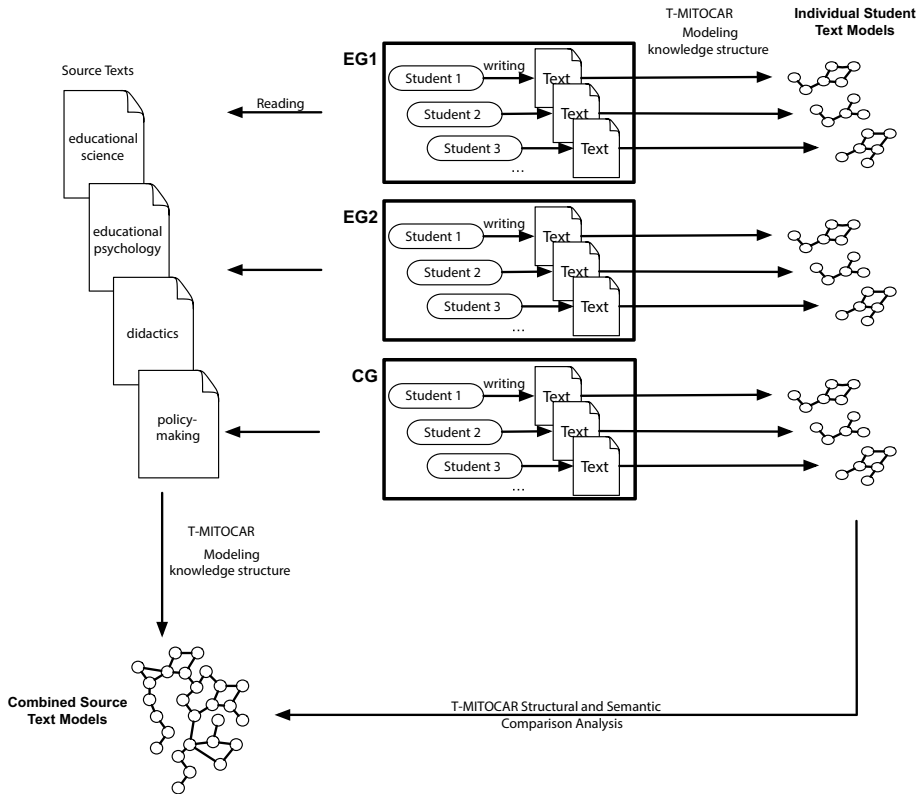


Fig. 3 Computer-linguistic analysis of the student texts

The source reference model consists of concepts that are bound by links. The links are associations strengths as determined by T-MITOCAR. At the links (Figure 2) are measures of association as weights between 0 and 1. One stands for the strongest association within the text and zero would stand for no association. Only the strongest ones are included in such a graph. Within the parenthesis is a linear transformation of the same value, so that the weakest links that still made it to the graph show a zero and the strongest show a one (Pirnay-Dummer, 2015b). The meaning of a concept is constituted by the context structure in which it is located in the network. The meaning of such a T-MITOCAR generated semantic knowledge structures (in this case, the reference model of the source texts) lies the way concepts are linked with each other (e.g. cyclic, hierarchical, sequential) and the connections they make to specific concepts but not others.

To explore the effect of prompting on the pre-service teachers' knowledge structures, we compared the student texts ($N = 109$, EG1: $n_1 = 37$, EG2: $n_2 = 37$, CG: $n_3 = 35$) with the reference model of the source texts using T-MITOCAR similarity measures. The participants in CG and EG2 tended to write longer texts than those in EG1. However, differences in the average word count between the groups were not significant (ANOVA, $F[2,106] = 1.67$, $p = .19$; $M_{EG1} = 407.11$, $SD_{EG1} = 122.59$; $M_{EG2} = 429.73$, $SD_{EG2} = 86.21$; $M_{CG} = 448.8$, $SD_{CG} = 74.56$; $F(2,106) = 1.75$).

The logic of the computer-linguistic analysis in this study is shown in Figure 3.

Table 5 Mean and standard deviation for computer-linguistic comparison measures per group between participants' texts and reference text model

Measure	Submeasure	Experimental condition		
		EG1	EG2	CG
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Structural measure	SUR	.13 (.06)	.14 (.05)	.15 (.05)
	GRA	.45 (.16)	.46 (.17)	.52 (.15)
	GAMMA	.39 (.15)	.36 (.07)	.40 (.09)
	STRU	.31 (.15)	.34 (.14)	.43 (.10)
Semantic measure	CONC	.25 (.08)	.26 (.06)	.28 (.06)
	PROP	.06 (.03)	.07 (.03)	.08 (.03)
	BSM	.25 (.08)	.26 (.08)	.30 (.09)

Means and standard deviations for the computer-linguistic comparison measures per group between participants' texts and reference text model are shown in Table 5.

MANOVA (Type III), with the four structural (SUR, GRA, STRU, GAMMA) and three semantic similarity scores (CONC, PROP, BSM) as dependent measures and the experimental conditions as independent measures, indicated significant differences between type of prompting (Wilks' $\lambda = 0.79$, $F[14,200] = 1.83$, $p = .04$).

We conducted follow-up univariate ANOVA (Type III) for each of the dependent measures. The results indicated significant differences between the experimental conditions for structural matching (STRU: $F[2,106] = 7.44$, $p < .001$, $\eta_p^2 = .12$; Figure 3), propositional matching (PROP: $F[2,106] = 4.78$, $p = .01$, $\eta_p^2 = .08$; Figure 4), and balanced semantic matching (BSM: $F[2,106] = 3.49$, $p = .03$, $\eta_p^2 = .06$). The differences in surface (SUR: $F[2,106] = 1.69$, $p = .19$), graphical (GRA: $F[2,106] = 2.10$, $p = .13$), concept matching (CONC: $F[2,106] = 1.80$, $p = .17$) and gamma were not significant (GAMMA: $F[2,106] = 1.28$, $p = .28$).

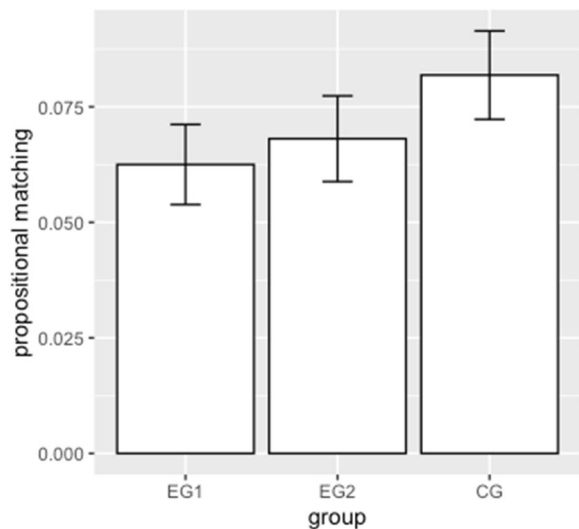
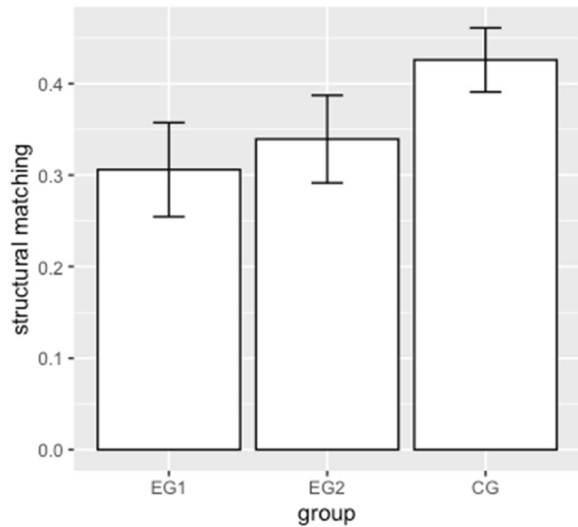
Fig. 4 Propositional similarity between participants' texts and reference source model per group

Fig. 5 Structural similarity between participants' texts and reference source model per group



Tukey HSD post-hoc for structural matching revealed that the control group (STRU: $M = 0.43$, $SD = 0.10$) achieved significantly higher similarities with the reference model than experimental group 1 (STRU: $M = 0.31$, $SD = 0.15$, $p < .001$) and experimental group 2 (STRU: $M = 0.34$, $SD = 0.14$, $p = .02$; see Figure 4). The control group (PROP: $M = 0.08$, $SD = 0.03$) achieved significantly higher propositional similarity to the reference model than experimental group 1 (PROP: $M = 0.06$, $SD = 0.03$, $p = .009$; see Figure 5). The contrasts for balanced semantic matching were not significant (BSM: $p > .05$).

Perceived relevance (DV3)

We used mixed ANOVA (Type III, with Greenhouse-Geisser correction for violation of sphericity) to analyze the differences in the students' perceived relevance of knowledge integration over time (R-Package: afex; Function: aov_4). There were no significant main effects of the within-subject factor (time, $F[1,105] = 0.77$, $p = .38$) or the between-subject factor (group, $F[2,105] = 0.20$, $p = .82$) on the differences between the groups over time. However, as Figure 6 illustrates, a significant interaction effect of time (within) and group (between) on perceived relevance was found ($F[2,105] = 5.68$, $p = .005$, $\eta^2_G = .027$).

Tukey-HSD-adjusted mixed post-hoc analysis revealed a significant ($p = .02$) increase in perceived relevance of knowledge integration for only EG1, the group receiving combined prompts ($M_{Diff} = -0.31$, 95% - CI[-.58, -.033]).

Discussion

The main objective of this study was to examine the effects of instructional prompts and relevance prompts embedded in pre-service teachers' learning processes on the quality of their knowledge integration in multiple document comprehension across domains.

We found that students receiving no prompts (CG) achieved a closer structural match to the overall reference model than the students aided by prompts (EG1, EG2). Also, their propositions were more similar to the reference model than those of the students with

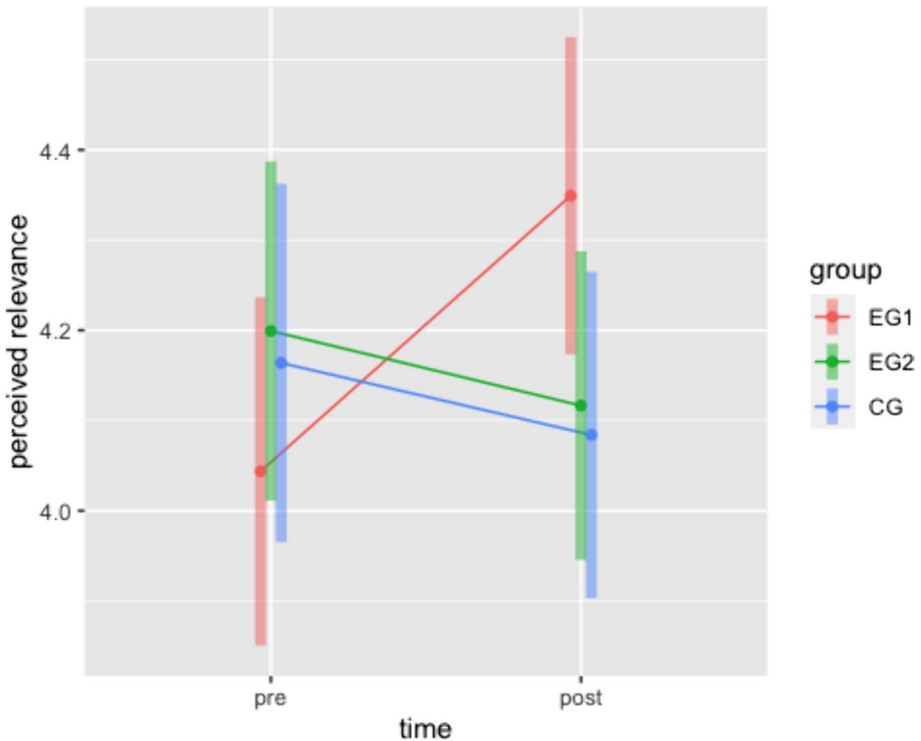


Fig. 6 Differences in perceived relevance of knowledge integration between groups over time

combined prompts (EG1). When interpreting the results, it is important to separate the knowledge level from the integration level of a text. The way the reference model was created to which the student texts were compared explains why higher similarity does not indicate higher knowledge integration, but rather more knowledge-telling (Scardamalia & Bereiter, 1987, 1991). The reference model was created by combining all four source texts into one document. From this, the overall model of the textual knowledge structure was then created using T-MITOCAR. Thus, the resulting model of the source texts is not an integrated knowledge model but rather a combined one. Students who have solved the writing task by preparing a short summary for each of the reference texts in turn achieved a higher structural and propositional similarity than students who actually made the effort to produce an integrated text of their own from the four reference texts. Thus, our results provide for structural and propositional matching provide a substantial empirical indication that students who do not receive prompts about knowledge integration are more prone to knowledge-telling than students who receive prompts.

However, no conclusions can be drawn from this as to whether students supported by the prompts actually integrated their knowledge better. We found no direct evidence that pre-service teachers were supported in their knowledge integration by pre-actional cognitive prompts in the form of task-supplemental focus questions in combination with repeated content-independent relevance instruction. This is contrary to previous studies who succeeded in prompting knowledge integration (e.g. Lehmann, Rott et al., 2019a, 2019b; Zeeb, Biwer, et al., 2019; Zeeb et al., 2020).

Overall, very few of the pre-service teachers in our study succeeded in writing integrated texts that contained transfer. This again is in line with previous empirical findings that student teachers struggle with knowledge integration (Graichen et al., 2019; Harr et al., 2014; Harr et al., 2015; Janssen & Lazonder, 2016). Pre-service teachers have little or no formal experience in knowledge integration, because their study domains are taught separately (Ball, 2000; Darling-Hammond, 2006; Hudson & Zgaga, 2017). The complex task we used in our experiment required students not only to integrate knowledge from four texts across domains but also to draw integrated conclusions for transfer. Low integration despite of repeated, combined prompting can be interpreted as an indicator that transfer-oriented knowledge integration should not be treated as a function of multi-document integration, even when it is as easy to control as it is within studies for knowledge integration. Instead of a text-inherent process of a task, which is limited to a particular domain and particularly trained expectations towards kinds of integration, transfer extending to the true expected task (what students believe to be their knowledge application later in “real life”) seems to modify the kind of integration as well as its outcome. Students in this study seem to be inspired to actually leave their academic knowledge and rely more on their word knowledge. When the students of this study are induced by the task and its way of introducing integration and thus transfer to leave the specific domain, they no longer perceive the same specific content to be as accessible or relevant. This is only just post-hoc at this point and not at all finite evidence, but it poses a recognizable danger to academic transfer and may even help to explain a lack of dissemination. In future analyses, we will also try to map commonplace knowledge to the solutions—this could be difficult to do, however, because we would first need to carefully create a comparable knowledge base as an additional outside criterion.

Low integration and transfer might suggest a floor effect that may have made it additionally difficult for raters, as there was little transfer and integrated knowledge to be found. This suggests that knowledge integration training and practice in the use of knowledge integration support is needed in handling complex tasks especially when the integration is directed at academic transfer.

We did not find significant group differences for text quality (transfer, conclusion, integration criteria). However, the low interrater reliability is a limitation of this study. It is suggested that new raters to be trained in more detail to obtain more reliable assessments of the texts. At the same time, we found group differences for time participants spent reading and working with the given text material. This shows that the focus questions given as instructional prompts to EG1 were indeed used and stimulated a significantly longer engagement with the texts compared to the relevance prompts alone or no prompts.

Contradicting our third hypothesis (see Objectives) and previous findings (Zeeb, Biwer, et al., 2019; Zeeb et al., 2020), relevance prompts alone were not enough to enhance the relevance perception, even though they were repeated and emphasized knowledge integration independently from source material content or the task (DV3). Only in combination with the complex instructional prompts did the perceived relevance increase over time. However, long-term effects were not part of this study, which limits the validity of this point.

Conclusion

The results of this study are surprising in some important ways: Knowledge integration seems to be even more complex than already known, particularly when it uses interdisciplinary domains aimed at transfer. Everyday knowledge may get in the way of academically

sound transfer in a much deeper sense than previously assumed. This should be considered as a prerequisite of transfer. Just leaving it to the practical imagination clearly does not suffice. Scholars and practitioners alike need to know about this gap before they can effectively train for integrated transfer.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The data that support the findings of this study are not openly available but are available from the corresponding author upon request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51(3), 241–247. <https://doi.org/10.1177/0022487100051003013>
- Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, 1(4), 139–161. [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)
- Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23(2), 139–145. <https://doi.org/10.1024/1010-0652.23.2.139>
- Barzilai, S., & Strømsø, H. I. (2018). Individual differences in multiple document comprehension. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 99–115). Routledge. <https://doi.org/10.4324/9781315627496-6>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, 39(1), 69–80. https://doi.org/10.1207/s15326985Sep3901_7
- Blömeke, S. (2009). Lehrerausbildung. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtasik, & W. Sacher (Eds.), *Handbuch schule. Theorie—organisation—entwicklung* (pp. 483–490). Klinkhardt.
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., & Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, 40, 450–465. <https://doi.org/10.3758/s13421-011-0160-6>
- Bråten, I., & Braasch, J. L. G. (2018). The role of conflict in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 184–201). Routledge. <https://doi.org/10.4324/9781315627496-11>
- Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, 47(1), 1–31. <https://doi.org/10.1080/01638530902959646>
- Bråten, I., & Strømsø, H. I. (2012). Knowledge acquisition: Constructing meaning from multiple information sources. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1677–1680). Springer. https://doi.org/10.1007/978-1-4419-1428-6_807
- Britt, M. A., & Rouet, J.-F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276–314). Cambridge University Press. <https://doi.org/10.1017/CBO9781139048224.017>
- Britt, M. A., Rouet, J.-F., & Durik, A. M. (2018). Representations and processes in multiple source use. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 17–32). Routledge. <https://doi.org/10.4324/9781315627496-2>
- Britt, M. A., & Sommer, J. (2004). Facilitating textual integration with marco-structure focusing tasks. *Reading Psychology*, 25(4), 313–339. <https://doi.org/10.1080/02702710490522658>

- Bromme, R. (2014). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens* (Vol. 7). Waxmann Verlag.
- Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T., & Neubrand, M. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? *Zeitschrift für Erziehungswissenschaft*, 9(4), 521–544. <https://doi.org/10.1007/s11618-006-0166-1>
- Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, 100(1), 209–222. <https://doi.org/10.1037/0022-0663.100.1.209>
- Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, 15(1), 1–40. https://doi.org/10.1207/s1532690xci1501_1
- Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Epistemological Development and Its Impact on Cognition in Academic Domains*, 29(2), 186–204. <https://doi.org/10.1016/j.cedpsych.2004.01.004>
- Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education*, 57(3), 300–314. <https://doi.org/10.1177/0022487105285962>
- Deci, E. L., & Ryan, R. M. (n.d.). Intrinsic motivation inventory (IMI). <http://selfdeterminationtheory.org/intrinsic-motivation-inventory/>
- Dörner, D. (1976). *Problemlösen als Informationsverarbeitung* (1st ed.). Kohlhammer.
- Evans, V., & Green, M. (2006). Cognitive linguistics. *An Introduction*. Edinburgh: Edinburgh University Press. <https://doi.org/10.4324/9781315864327>
- Gentner, D., & Stevens, A. L. (1983). *Mental models*. Erlbaum.
- Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology*, 35(3), 157–173. <https://doi.org/10.1016/j.cedpsych.2009.11.002>
- Graichen, M., Wegner, E., & Nückles, M. (2019). Wie können Lehramtsstudierende beim Lernen durch Schreiben von Lernprotokollen unterstützt werden, dass die Kohärenz und Anwendbarkeit des erworbenen Professionswissens verbessert wird? *Unterrichtswissenschaft*, 47(1), 7–28. <https://doi.org/10.1007/s42010-019-00042-x>
- Hähnlein, I. (2018). *Erfassung epistemologischer Überzeugungen von Lehramtsstudierenden. Entwicklung und Validierung des StEB Inventar*. (Dissertation). Universität Passau. <https://opus4.kobv.de/opus4-uni-passau/frontdoor/index/index/docId/588>.
- Hähnlein, I., & Pirnay-Dummer, P. (2019). *Assessing Metacognition in the Learning Process. Construction of the Metacognition in the Learning Process Inventory MILP*. Paper presented at the European Association for Research on Learning and Instruction (EARLI) Conference, Aachen, Germany.
- Harr, N., Eichler, A., & Renkl, A. (2014). Integrating pedagogical content knowledge and pedagogical/psychological knowledge in mathematics. *Frontiers in Psychology*, 5, 1–10. <https://doi.org/10.3389/fpsyg.2014.00924>
- Harr, N., Eichler, A., & Renkl, A. (2015). Integrated learning: Ways of fostering the applicability of teachers' pedagogical and psychological knowledge. *Frontiers in Psychology*, 6, 1–16. <https://doi.org/10.3389/fpsyg.2015.00738>
- Helbig, H. (2006). Knowledge representation and the semantics of natural language. *Springer*. <https://doi.org/10.1007/3-540-29966-1>
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on students' achievement. *American educational research journal*, 42(2), 371–406. <https://doi.org/10.3102/00028312042002371>
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.3102/00346543067001088>
- Hudson, B., & Zgaga, P. (2017). History, context and overview: Implications for teacher education policy, practice and future research. In H. Brian (Ed.), *Overcoming Fragmentation in Teacher Education Policy and Practice* (pp. 1–10). Cambridge University Press.
- Ifenthaler, D. (2010). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research and Development*, 58(1), 81–97. <https://doi.org/10.1007/s11423-008-9087-4>
- Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Journal of Educational Technology & Society*, 15(1), 38–52.
- Ifenthaler, D., & Lehmann, T. (2012). Preactional self-regulation as a tool for successful problem solving and learning. *Technology, Instruction, Cognition and Learning*, 9(1–2), 97–110.

- Janssen, N., & Lazonder, A. W. (2016). Supporting pre-service teachers in designing technology-infused lesson plans. *Journal of Computer Assisted Learning*, 32(5), 456–467. <https://doi.org/10.1111/jcal.12146>
- Johnson-Laird, P. N. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness*. Cambridge Univ Press. <https://doi.org/10.2307/414498>
- Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking*. Cham: Prentice hall.
- Jonassen, D. H. (2006). On the role of concepts in learning and instructional design. *Educational Technology Research and Development*, 54(2), 177–196. <https://doi.org/10.1007/s11423-006-8253-9>
- Jonassen, D. H. (2012). Problem typology. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning*. Springer. https://doi.org/10.1007/978-1-4419-1428-6_209
- Jonassen, D. H., & Cho, Y. H. (2008). Externalizing mental models with mindtools. In D. Ifenthaler, P. Pirnay-Dummer, & J. M. Spector (Eds.), *Understanding models for learning and instruction Essays in honor of Norbert M Seel* (pp. 145–159). Springer. https://doi.org/10.1007/978-0-387-76898-4_7
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological review*, 95(2), 163–182. <https://doi.org/10.1037/0033-295X.95.2.163>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Koehler, M. J., & Mishra, P. (2008). Introducing TPCK. In AACTE Committee on Innovation and Technology (Eds.), *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 2–29). Routledge.
- König, J. (2010). Lehrerprofessionalität. Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In J. König & B. Hofmann (Eds.), *Professionalität von Lehrkräften. Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können* (pp. 40–106). Dgls.
- Krauskopf, K., Zahn, C., & Hesse, F. W. (2020). Concetualizing (pre-service) teachers' professional knowledge for complex domains. In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 31–57). Brill Sense. https://doi.org/10.1163/9789004429499_003
- Kulturministerkonferenz. (2009). *Kompetenzstufenmodell zu den Bildungsstandards im Kompetenzbereich: Sprechen und Zuhören—hier Zuhören—für den Mittleren Schulabschluss*. Beschluss der Kultusministerkonferenz.
- Lehmann, T. (2020a). What is knowledge integration of multiple domains and how does it relate to teachers' professional competence? In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 9–29). Brill Sense. https://doi.org/10.1163/9789004429499_002
- Lehmann, T. (2020b). *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education*. Brill Academic Publishers. <https://doi.org/10.1163/9789004429499>
- Lehmann, T. (2022). Student teachers' knowledge integration across conceptual borders: The role of study approaches, learning strategies, beliefs, and motivation. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-021-00577-7>
- Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on reflection in self-regulated online learning. *Computers in Human Behavior*, 32, 313–323. <https://doi.org/10.1016/j.chb.2013.07.051>
- Lehmann, T., Klieme, K., & Schmidt-Borcherding, F. (2020). Separative and integrative learning in teacher education: Validity and reliability of the “SILTE” short scales. In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 155–177). Brill Academic Publishers.
- Lehmann, T., Pirnay-Dummer, P., & Schmidt-Borcherding, F. (2019a). Fostering integrated mental models of different professional knowledge domains: Instructional approaches and model-based analyses. *Educational Technology Research and Development*, 68(3), 905–927. <https://doi.org/10.1007/s11423-019-09704-0>
- Lehmann, T., Rott, B., & Schmidt-Borcherding, F. (2019b). Promoting pre-service teachers' integration of professional knowledge: Effects of writing tasks and prompts on learning from multiple documents. *Instructional Science*, 47(1), 99–126. <https://doi.org/10.1007/s11251-018-9472-2>
- Linn, M. C. (2000). Designing the knowledge integration environment. *International Journal of Science Education*, 22(8), 781–796. <https://doi.org/10.1080/095006900412275>
- Mishra, P. (2019). Considering contextual knowledge: The TPACK diagram gets an upgrade. *Journal of Digital Learning in Teacher Education*, 35(2), 76–78. <https://doi.org/10.1080/21532974.2019.1588611>

- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18(1), 3–54.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. *The Psychology of Learning and Motivation*, 26, 125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition knowing about knowing* (pp. 1–25). Cambridge University Press.
- Norton, P., van Rooij, S. W., Jerome, M. K., Clark, K., Behrmann, M., & Bannan-Ritland, B. (2009). Linking theory and practice through design: An instructional technology program. In M. Orey, V. J. McClendon, & R. M. Branch (Eds.), *Educational media and technology yearbook* (pp. 47–59). Springer. https://doi.org/10.1007/978-0-387-09675-9_4
- Partee, B. H. (2004). Compositionality in formal semantics: Selected papers of Barbara Partee. *Blackwell Pub.* <https://doi.org/10.1002/9780470751305>
- Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 99–122). Lawrence Erlbaum Associates.
- Pirnay-Dummer, P. (2006). *Expertise und Modellbildung-MITOCAR*. (Dissertation). Albert-Ludwigs-Universität Freiburg (Breisgau).
- Pirnay-Dummer, P. (2007). *Model inspection trace of concepts and relations: A heuristic approach to language-oriented model assessment*. Paper presented at the annual meeting of AERA, Chicago IL.
- Pirnay-Dummer, P. (2010). Complete structure comparison. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 235–258). Springer. https://doi.org/10.1007/978-1-4419-5662-0_13
- Pirnay-Dummer, P. (2014). Gainfully guided misconception: How automatically generated knowledge maps can help companies within and across their projects. In D. Ifenthaler & R. Hanewald (Eds.), *Digital knowledge maps in higher education: Technology-enhanced support for teachers and learners* (pp. 253–274). Springer. https://doi.org/10.1007/978-1-4614-3178-7_14
- Pirnay-Dummer, P. (2015a). Knowledge elicitation. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (pp. 438–442). SAGE. <https://doi.org/10.4135/9781483346397.n183>
- Pirnay-Dummer, P. (2015b). Linguistic analysis tools. In C. A. MacArthur, M. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 427–442). Guilford Publications.
- Pirnay-Dummer, P. (2020). Knowledge and structure to teach: A model-based computer-linguistic approach to track, visualize, compare and cluster knowledge and knowledge integration in pre-service teachers. In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 133–154). Brill Sense. https://doi.org/10.1163/9789004429499_007
- Pirnay-Dummer, P., Ifenthaler, D., & Seel, N. M. (2012). Designing model-based learning environments to support mental models for learning. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 55–90). Routledge. <https://doi.org/10.4324/9780203813799>
- Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development*, 58(1), 3–18. <https://doi.org/10.1007/s11423-009-9119-8>
- Pirnay-Dummer, P., & Seel, N. M. (2018). The sciences of learning. In L. Lin & J. M. Spector (Eds.), *The Sciences of learning and instructional design: Constructive articulation between communities* (pp. 8–35). Routledge. <https://doi.org/10.4324/9781315684444-2>
- R-Core-Team (2022). R: A language and environment for statistical computing (Version 4.2.2): R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), *Instructional design theories and models: An overview of their current status* (pp. 335–382). Erlbaum. <https://doi.org/10.4324/9780203824283>
- 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
- Rouet, J.-F. (2006). The skills of document use: From text comprehension to web-based learning. *Lawrence Erlbaum Associates*. <https://doi.org/10.4324/9780203820094>

- Rouet, J.-F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). IAP Information Age Publishing.
- Rouet, J.-F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200–215. <https://doi.org/10.1080/00461520.2017.1329015>
- Rule, D. C., & Bendixen, L. D. (2010). The integrative model of personal epistemology development: Theoretical underpinnings and implications for education. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom* (pp. 94–123). Cambridge University Press. <https://doi.org/10.1017/CBO9780511691904.004>
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics* (Vol. 2, pp. 142–175). Cham: Cambridge University Press. <https://doi.org/10.1017/S0142716400008596>
- Scardamalia, M., & Bereiter, C. (1991). Literate expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*. Cham: Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences*, 3(3), 265–283. https://doi.org/10.1207/s15327809jls0303_3
- Scardamalia, M., & Bereiter, C. (1999). Schools as knowledge-building organizations. In D. P. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations: Social, biological, and educational dynamics* (pp. 274–289). Guilford Press.
- Schiefele, U., & Pekrun, R. (1996). Psychologische Modelle des selbstgesteuerten und fremdgesteuerten Lernens. In F. E. Weinert (Ed.), *Enzyklopädie der Psychologie—Psychologie des Lernens und der Instruktion Pädagogische Psychologie* (Vol. 2, pp. 249–287). Hogrefe.
- Schneider, M. (2012). Knowledge integration. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1684–1686). Springer. https://doi.org/10.1007/978-1-4419-1428-6_807
- Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (Eds.), *Beliefs about text and instruction with text* (pp. 25–40). Erlbaum.
- Schwarzer, R., & Jerusalem, M. (2003). SWE: Skala zur Allgemeinen Selbstwirksamkeitserwartung [Verfahrensdokumentation, Autorenbeschreibung und Fragebogen]. In L.-I. f. P. (ZPID) (Ed.), *Open Test Archive*. <https://doi.org/10.23668/psycharchives.4515>
- Seel, N. M. (1991). *Weltwissen und mentale Modelle*. Hogrefe Verlag f.
- Seel, N. M. (2003). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning*, 1(1), 59–85.
- Seel, N. M. (2012). Learning and thinking. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (Vol. 4, pp. 1797–1799). Springer.
- Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2013). Mental models and their role in learning by insight and creative problem solving. In J. M. Spector, B. B. Lockee, S. E. Smaldino, & M. Her-ring (Eds.), *Learning, problem solving, and mind tools: Essays in honor of David H. Jonassen* (pp. 10–34). Springer.
- Seel, N. M., Lehmann, T., Blumschein, P., & Podolskiy, O. A. (2017). Instructional design for learning: Theoretical foundations. *Sense*. <https://doi.org/10.1007/978-94-6300-941-6>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Taylor, J. R. (2007). Cognitive linguistics and autonomous linguistics. In D. Geeraerts & H. Cuyckens (Eds.), *Cognitive Linguistics* (pp. 566–588). Oxford University Press.
- Thompson, C. (2020). *Allgemeine Erziehungswissenschaft*. Eine Einführung (Vol. 3). <http://www.klinkhardt.de/ewr/97831726165.html>.
- Tittmann, P. (2010). Graphs and networks. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 177–188). Springer. https://doi.org/10.1007/978-1-4419-5662-0_10
- Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Winne, P. H. (2016). Identity and epistemic emotions during knowledge revision: A potential account for the backfire effect. *Discourse Processes*, 53(5–6), 339–370. <https://doi.org/10.1080/0163853X.2015.1136507>
- van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. *Academic Press*. <https://doi.org/10.2307/415483>
- 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>

- Wäschle, K., Lehmann, T., Brauch, N., & Nückles, M. (2015). Prompted journal writing supports pre-service history teachers in drawing on multiple knowledge domains for designing learning tasks. *Peabody Journal of Education*, 90(4), 546–559. <https://doi.org/10.1080/0161956X.2015.1068084>
- Whitehead, A. N. (1929). *The aims of education*. Macmillan.
- Wiley, J., & Voss, J. F. (1996). The effects of 'playing historian' on learning in history. *Applied Cognitive Psychology*, 10(7), 63–72.
- Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory from text. *Journal of Educational Psychology*, 91(2), 301–311. <https://doi.org/10.1037/0022-0663.91.2.301>
- Wilhelm, O., & Nickolaus, R. (2013). Was grenzt das Kompetenzkonzept von etablierten Kategorien wie Fähigkeit, Fertigkeit oder Intelligenz ab? *Zeitschrift für Erziehungswissenschaft*, 16, 23–26. <https://doi.org/10.1007/s11618-013-0380-6>
- Zeeb, H., Biwer, F., Brunner, G., Leuders, T., & Renkl, A. (2019). Make it relevant! How prior instructions foster the integration of teacher knowledge. *Instructional Science*, 47(6), 711–739. <https://doi.org/10.1007/s11251-019-09497-y>
- Zeeb, H., Spitzmesser, E., Röddiger, A., Leuders, T., & Renkl, A. (2020). Using relevance instructions to support the integration of teacher knowledge. In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 201–229). Brill, Sense. https://doi.org/10.1163/9789004429499_010
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

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

Inka Sara Hähnlein is a research assistant (post-doc) at the Department of Educational Psychology at the Institute of Pedagogy of the Martin Luther University Halle-Wittenberg. She holds a doctorate in educational science. Her current research interests in the field of empirical teaching and learning research, teacher education, and pre-service teacher training are cross-domain knowledge integration from texts, multiple document comprehension, computer-linguistic methods for the psychological identification of knowledge and metacognition in the learning process.

Pablo Pirnay-Dummer is Full Professor of Educational Psychology at the Department of Pedagogy of the Martin Luther University Halle-Wittenberg. His current research interests are the processes of language in learning, computer-linguistic methods for the psychological identification of knowledge in text (knowledge assessment, including software development), modeling and comparison of knowledge domains, expertise and complex problem solving in processes of learning and instruction (including task synthesis and simulation) with a focus on teacher education and pre-service teacher training.