International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Influence of using challenging tasks in biology classrooms on students' cognitive knowledge structure: an empirical video study

Jigna Nawani, Julia Rixius & Birgit J. Neuhaus

To cite this article: Jigna Nawani, Julia Rixius & Birgit J. Neuhaus (2016) Influence of using challenging tasks in biology classrooms on students' cognitive knowledge structure: an empirical video study, International Journal of Science Education, 38:12, 1882-1903, DOI: 10.1080/09500693.2016.1213456

To link to this article: http://dx.doi.org/10.1080/09500693.2016.1213456

Published online: 26 Aug 2016.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 12, 1882–1903
http://dx.doi.org/10.1080/09500693.2016.1213456

Routledge
Taylor & Francis Group

Influence of using challenging tasks in biology classrooms on students' cognitive knowledge structure: an empirical video study

Jigna Nawani, Julia Rixius and Birgit J. Neuhaus

Biology Education, Department I, Faculty of Biology, Ludwig-Maximilians-Universität München, Munich, Germany

ABSTRACT

Empirical analysis of secondary biology classrooms revealed that, on average, 68% of teaching time in Germany revolved around processing tasks. Quality of instruction can thus be assessed by analyzing the quality of tasks used in classroom discourse. This quasi-experimental study analyzed how teachers used tasks in 38 videotaped biology lessons pertaining to the topic 'blood and circulatory system'. Two fundamental characteristics used to analyze tasks include: (1) required cognitive level of processing (e.g. low level information processing: repetiition, summary, define, classify and high level information processing: interpret-analyze data, formulate hypothesis, etc.) and (2) complexity of task content (e.g. if tasks require use of factual, linking or concept level content). Additionally, *students' cognitive knowledge structure* about the topic 'blood and circulatory system' was measured using student-drawn concept maps ($N = 970$ students). Finally, linear multilevel models were created with *high-level cognitive processing tasks* and *higher content complexity tasks* as class-level predictors and *students' prior knowledge*, *students' interest in biology*, and *students' interest in biology activities* as control covariates. Results showed a positive influence of *high-level cognitive processing tasks* ($\beta = 0.07$; $p < .01$) on *students' cognitive knowledge structure*. However, there was no observed effect of *higher content complexity tasks* on *students' cognitive knowledge structure*. Presented findings encourage the use of *high-level cognitive processing tasks* in biology instruction.

ARTICLE HISTORY
Received 17 April 2015
Accepted 12 July 2016

KEYWORDS
Challenging tasks; video studies; biology teaching; concept maps

International studies like the 'Programme for International Student Assessment – PISA' and 'The Third International Mathematics and Science Study – TIMSS' found country-specific differences in the way students performed in science and mathematics tests. Thus, the TIMSS – 1995, 1999 video studies were conceptualized to explore unique features of teaching science and mathematics in various countries (like Japan, United States, etc. (Jacobs et al., 2006; Roth et al., 2006). However, TIMSS did not investigate the influence of teaching features on students' learning outcomes. Furthermore, Germany participated in the mathematics part of this cross-country study (TIMSS 1995) and thus

CONTACT Jigna Nawani jigna.nawani@biologie.uni-muenchen.de

German science – especially biology classrooms are rarely investigated (Wadouh, Liu, Sandmann, & Neuhaus, 2014). In that regard, the presented study investigated the way teachers' use of *challenging tasks* (e.g. *high-level cognitive processing tasks* requiring high level or deeper information processing; *higher content complexity tasks* requiring the use of linking or concept level content) in German secondary biology classrooms influenced students' learning outcomes.

The TIMSS – 1995, 1999 video studies identified *cognitive activation* as one important characteristic of effective mathematics and science teaching. This study found that higher achieving countries like Japan and Australia used cognitively challenging tasks and complex content like interconnections between facts, concepts and core ideas to engage students in learning domain-related content (Hiebert et al., 2003, Klieme and Bos 2000; Roth et al., 2006). *Cognitively activating instruction* can be defined as a teaching practice that encourages deeper processing of new content presented during the classroom discourse (Lipowsky et al., 2009). In science and mathematics classrooms, cognitive activation is usually studied at three different instructional levels: (1) teaching of complex domain content, (2) use of challenging tasks and (3) practicing thoughtful discourse (see Figure 1). However, teaching effectiveness studies differ in the way they define and operationalize this construct (Kunter et al., 2013). Lipowsky et al. (2009) found a positive

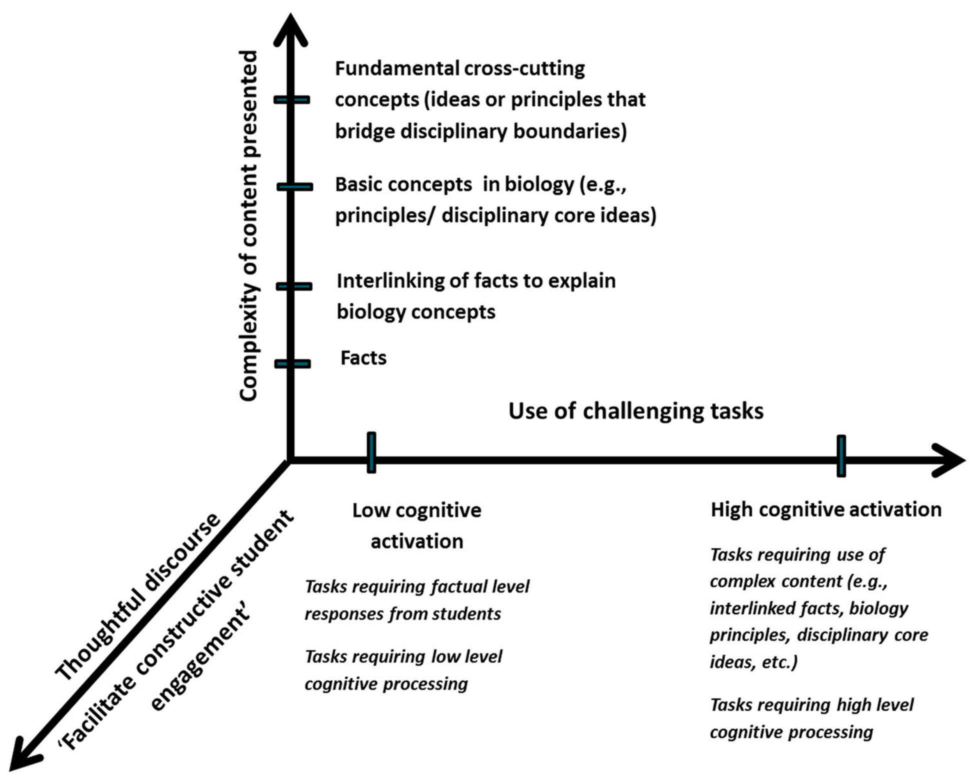

Figure 1. Cognitively activating instruction. Based on the oretical descriptions presented by Förtsch et al. (2016a); Lipowsky et al. (2009); Wadouh et al. (2014).

correlation between cognitively activating instruction and students mathematics achievement. One recent study found a positive influence of cognitively activating instruction on students' situational interest and student achievement in biology classrooms (Förtsch, Werner, Dorfner, von Kotzebue, & Neuhaus, 2016a). However, these empirical studies usually measure the cognitive activation potential of complete lessons (Förtsch, et al., 2016a; Lipowsky, et al., 2009). Thus, it is unclear whether enhancement in student outcomes was due to (1) teaching of complex domain content, (2) use of challenging tasks or (3) practicing thoughtful discourse (see Figure 1). Jatzwauk, Rumann, & Sandmann (2008) analyzed the use of tasks in German biology classrooms. This study found that *two-third of class time* in German biology classrooms was utilized for processing tasks. Thus, frequency with which teachers used challenging tasks (e.g. *high-level cognitive processing tasks* and *higher content complexity tasks*) during classroom discourse can be used to measure the cognitive activation potential of lessons (Blömeke, Risse, Müller, Eichler, & Schulz, 2006; Klieme & Bos, 2000; Lipowsky et al., 2009). Furthermore, teaching effectiveness studies so far have investigated the influence of teaching features on student outcomes like 'performance in knowledge tests' or 'situational interest'. Thus, *students' cognitive knowledge structure* (i.e. interconnectedness of students' knowledge about a topic or domain), an important competence reflecting domain expertise, is rarely investigated (Ruiz-Primo & Shavelson, 2005; Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005). To that end, this empirical investigation examined the influence of using challenging tasks on students' cognitive knowledge structure, measured using student-drawn concept maps. To begin with, we first present the literature review guiding this study.

Cognitively activating instruction

Pauli, Drollinger-Vetter, and Hugener (2008) defined cognitive activation as active, constructive and discursive engagement with domain-related content. Cognitively activating instruction can be described as 'use of learning activities or tasks that engage students in developing conceptual level content' (Kunter et al., 2013). Described below are three features that together depict the cognitive activation potential of science or mathematics lessons (Förtsch et al., 2016a; Lipowsky et al., 2009).

(1) Teaching of complex domain content that includes interconnected facts, biology concepts, principles and disciplinary core ideas (Hiebert et al., 2003; Jacobs et al., 2003, 2006; Neumann, Fischer & Summefleth, 2008; Schönborn & Bögeholz, 2009; Wadouh et al., 2014).
(2) Use of challenging tasks that involve higher order cognitive processing and high content complexity (Bloom, 1972; Craik & Lockhart, 1972; Ergönenc, Neumann, & Fischer, 2014; Hiebert et al., 2003; Jacobs et al., 2003).
(3) Practicing thoughtful discourse that constructively engages students in new knowledge generation process (Chi, 2009; Hugener et al., 2009; Mayer 2009) (see Figure 1).

In order to meaningfully inform the instructional practice, it is important to understand how the above-mentioned instructional levels influence students' domain knowledge. In that regard, Neumann et al. (2008) and Wadouh et al. (2014) showed a positive influence of teaching complex domain content on *students' cognitive knowledge* *structure* in physics and biology, respectively. However, teaching effectiveness in science and mathematics classrooms studies advocate the 'use of tasks' (i.e. teacher-initiated questions or activities) as an important instructional tool for developing deeper conceptual (Ergönenc et al., 2014; Förtsch et al., 2016a; Jacobs et al., 2003, 2006; Jatzwauk et al., 2008; Lipowsky et al., 2009). In this regard, the presented study analyzed the effectiveness of using *challenging tasks* in biology classrooms on *students' cognitive knowledge structure*.

Challenging tasks

Tasks are basic treatment units that could be used to orchestrate transfer of new domain knowledge (Blumenfeld & Meece, 1988; Doyle, 1979, 1983). 'Tasks as classrooms learning opportunities' form an interactive interface between students' already acquired knowledge and new content being taught in lessons (Jatzwauk, 2007; Jatzwauk et al., 2008). Teachers use tasks to redirect students' attention on specific aspects of content. In that regard, tasks could encourage students to cognitively process the new information and share it with the class for further discussion. Two elements that can help differentiate tasks used during classroom discourse are: 'required cognitive level of processing' and 'complexity of task content' (Blumenfeld & Meece, 1988). The presented study used these fundamental task characteristics (i.e. 'required cognitive level of processing' and 'complexity of task content') to identify *challenging tasks* in 38 videotaped biology lessons.

Two types of *challenging tasks* analyzed in this study include:

(1) *High-level cognitive processing tasks*: Several empirical studies have shown that tasks requiring deeper analysis of content enhance students' conceptual understanding and overall performance (Brown, 1994; Klieme & Bos, 2000; Lipowsky et al., 2009; Stein & Lane, 1996). Such tasks include deeper information-processing situations such as designing an experiment, formulating a hypothesis, presenting reasons or explanation for a given problem, interpreting or analyzing data, reflecting or evaluating a given scenario (Anderson, Krathwohl, & Bloom, 2001; Krathwohl, 2002). On the other hand, tasks requiring repetition, enlisting, classifying or comparing do not engage students in deeper processing of the new information presented during lessons. Hence, this study endeavored to examine teacher-initiated tasks for their level of cognitive processing (e.g. high level: analysis, reasoning, interpretation, etc. low level: repetition, classifying, comparing, etc.) (Anderson et al., 2001; Bloom, 1972; Blooms Taxonomy, n.d.; Craik & Lockhart, 1972; Ergönenc et al., 2014; Krathwohl, 2002) (see Table 1).

(2) *Higher content complexity tasks:* Teaching effectiveness studies have found that mathematics and science lessons usually focus on presenting and reinforcing facts related to the topics being taught (Jacobs et al., 2006; Neumann et al., 2008; Wadouh et al., 2014). These studies have advocated that mathematics and science lessons must include content as well as tasks that enable students to see: 1) how facts can be interconnected to describe concepts (e.g. facts like 'cytoplasm of red blood cells is rich in hemoglobin' or 'red blood cells help carry gases' can be interlinked to explore how oxygen is transported from lungs to body cells), 2) how basic principles can be used to explain biology concepts (e.g. antigen–antibody interactions during blood transfusion can be explained using the key-lock principle, gas exchange across

Table 1. Challenging tasks in biology lessons.

Elements that help differentiate tasks: 'Complexity of task content' and 'Level of cognitive processing demanded' (Anderson et al., 2001; Bloom, 1972; Blumenfeld & Meece, 1988; Hiebert et al., 2003; Wadouh et al., 2014)

Complexity of task content	Fact-level content (Lower content complexity tasks)	Linking or conceptual-level content (Higher content complexity tasks)
Cognitive processing level of tasks	Low-level cognitive processing situations: Repetition, Summary, Define, List, Classify, Arrange, Compare, Contrast (Low-level cognitive processing tasks)	High-level cognitive processing situations: Explaining – giving reasons, Designing experiment, Formulating hypothesis, Interpret or analyze data, Reflect –rethink, Evaluate (High-level cognitive processing tasks)

thin-walled air sacs or alveoli in lungs' can be explained using the principle of diffusion gradient), 3) how core domain ideas can be used to acquire a deeper understanding about important biology ideas and phenomenon (e.g. red blood cells have flexible, disc-like shape to increase the surface area for gas exchange and enhance their flexibility to fit through narrow blood vessels - this makes more sense, when looked through the lens of core domain idea of 'form follows function') (Hiebert et al., 2003; Jacobs et al., 2003, 2006; Neumann et al., 2008; Wadouh et al., 2014).

In that regard, quality of instruction can be described by investigating the content complexity of tasks used in these lessons. In other words, if the task presented by the teacher requires the use of simple factual level content or complex linking and conceptual level content. Analysis of TIMSS teaching videos showed an extensive use of higher content complexity tasks (40%) in higher achieving Japanese mathematics classrooms (Hiebert et al., 2003). Similarly, analysis of TIMSS science videos revealed that most activities and teacher utterances in higher achieving Japanese and Australian science classrooms focused on conceptual linking of domain content (Jacobs et al., 2006). In that regard, the presented study investigated the content complexity of tasks used by German biology teachers. Teacher-initiated tasks were analyzed for use of factual (e.g. name the components of blood) and linking or conceptual-level content (e.g. Why can't we transfuse blood from a donor with a different blood group?) (Neumann et al., 2008; Schönborn & Bögeholz, 2009; Wadouh et al., 2014) (see Table 1).

Students' cognitive knowledge structure in biology

Learning involves assimilating new knowledge and connecting it with prior knowledge to form an integrated knowledge structure about any topic (Ausubel 1963, 1968). Ausubel (1963) described two ways of acquiring knowledge: rote and meaningful learning. Rote learning focuses on assimilating knowledge of isolated facts, whereas meaningful learning involves assimilation of new information and linking it with prior knowledge to develop a more complex knowledge structure (Wadouh et al., 2014). Research in this regard has shown that experts possess complex cognitive knowledge structures, while novices have simpler knowledge structures, which consist of isolated facts or propositions (Duschl, Schweingruber, & Shouse, 2007; Glaser, 1991). Several empirical studies found that students who acquired interconnected and integrated knowledge could remember content more successfully than students who acquired knowledge in the form of isolated facts (Osborne & Wittrock, 1985; Wadouh et al., 2014). Thus, meaningful learning involves continuous refining of knowledge structures (about domain related topics) to develop expertise in the domain (Mayer, 1998; Resnick, 1989).

Knowledge structure in domains like biology consists of interconnections or links between biology concepts and appreciation of underlying principles or disciplinary core ideas (Wadouh et al., 2014). However, school assessments rarely focus on evaluating *students' cognitive knowledge structure* about the topics being taught in classrooms (Ruiz-Primo & Shavelson, 2005; Yin et al., 2005). To that end, our study examined the relation between teacher-initiated tasks used during biology classroom discourse and *students' cognitive knowledge structure*, evaluated using the concept mapping exercise.

Concept maps

Concept mapping is a valuable tool in assessing *students' cognitive knowledge structure* about a certain topic (Zele & Lenaerts, 2004). Concept maps reflect conceptual terms and interconnectedness of terms related to a topic. Concept mapping exercises involve both linear and hierarchical structures of knowledge (Kinchin & Hey, 2000; Kinchin, 2011). Several scoring systems have been suggested for assessing the linear and hierarchical structures (or connections) in concept maps (Ruiz-Primo & Shavelson, 2005; Zele & Lenaerts, 2004; Yin et al., 2005). The quantitative scoring systems count the number of valid structures or propositions (Ruiz-Primo & Shavelson, 2005). A valid proposition is a structure that includes two conceptual terms connected by a labeled arrow. The qualitative scoring systems rely upon expert evaluation to analyze the content and quality of maps (Kinchin & Hey, 2000; Kinchin, 2011). Quantitative methods are hence objective and more reliable (Zele & Lenaerts, 2004). Wadouh et al. (2014) used the quantitative scoring method to evaluate concept maps for variables like: (1) number of relations (propositions) drawn, (2) number of cross-relations drawn, (3) number of separate networks or concept maps drawn (4) number of correct relations drawn and (5) number of relations with deeper explanations for connections drawn. Here, the term cross-relation can be defined as a relation between the concept (or term) of the topic 'blood and circulatory system' and concepts (or terms) of other topics such as immune system and respiration. We used the above-mentioned variables related to concept maps, while investigating the influence of teachers' *use of challenging tasks* on *students' cognitive knowledge structure.*

Students' prior knowledge

According to constructivists, learning is an active process of acquiring new knowledge in a way that it is linked with pre-existing knowledge (Gerstenmaier & Mandl, 1995). Acquisition of new knowledge leads to extension or correction of learners' existing knowledge structures (Wadouh, 2008). Several studies have found that availability of relevant knowledge is a crucial parameter for acquiring new knowledge (Alexander, Kulikowich, & Schulze, 1994; Garner & Gillingha, 1991). *Students' prior knowledge* is thus an important parameter in determining their success in acquiring new knowledge and developing complex knowledge structures. In that context, we used *students' prior knowledge* (related to the topic) as a control covariate for this research investigation.

Students' motivation and interest to learn biology

Motivation can be described as individual preferences or reasons that lead to a certain behavior (Gredler, Broussard & Garisson, 2004; Guay et al., 2010). Self-determination theory described different types of achievement motivations based on the reasons that lead to a behavior or action (Deci & Ryan, 1985). Based on this theory, intrinsic motivation could be described as individual engagement in an activty because they feel rewarded by completing the task. In the classroom context, such a behavior reflects autonomy where student involvement is sustained due to their inherent interest in the content, discussion or activities presented by the learning environment (Krapp, 2002; Wadouh et al., 2014). On the other hand, extrinsic motivation could be described as individual engagement in tasks to receive an external positive outcome or avoid a negative outcome. Extrinsic learning motivation could thus be described as student engagement in learning activities to achieve good grades, teacher approval, or just to avoid negative teacher response. Schiefele and Schreyer (1994) found a positive relationship between intrinsic learning motivation and student achievement.

However, researchers argued that achievement motivation does not account for content specificity of learning motivation and thus it is important to explore how interest with regard to a specific context, theme or activity can influence achievement (Schiefele, Krapp, Prenzel, Heiland, & Kasten, 1983). Individual interest in that regard could be defined as a person's preference or affinity for certain themes, objects or activities. This person–object interaction is also referred to as 'object engagement' (Krapp, 2002). In classroom contexts, this engagement is deliberately aimed at enhancing the student understanding of various topics. Researchers suggest that such intentional learning environments could gradually enhance student disposition to learn about a given topic or domain. Empirical research in that regard has also found that thematic interest is an important predictor of performance (Schiefele, Krapp, & Winteler, 1992; Prenzel, 1988).

Thus, the presented study used motivation (intrinsic and extrinsic) and interest (interest in the subject, interest in subject-related activities) as control variables while investigating the hypotheses defined for the study.

Video-based observation of classroom instruction

Video-based direct observation is increasingly being used to analyze deeper features of classrooms instruction and correlating them with student learning outcomes (Rakoczy et al., 2007). The TIMSS – 1999 study compared mathematics teaching in seven countries that include Australia, Czech Republic, Hong Kong, Japan, Netherlands, Switzerland and United States. This study found that more than 50% of tasks in high-achieving Japan emphasized on making connections between mathematical facts, concepts and procedure. Moreover, 40% of tasks in Japanese mathematics classrooms demanded high-level procedural complexity (Hiebert et al., 2003). A similar study about German physics classrooms found that 80% of teacher-initiated tasks demanded lower order cognitive processing that is, reproducing factual knowledge (Seidel et al., 2007). One recent study analyzed high-complexity and high-cognitive-processing tasks in videotaped Grade 6 biology lessons. This study found a positive influence of high-cognitive-processing tasks on students' conceptual knowledge (Förtsch, Werner, von Kotzebue, & Neuhaus, 2016b).

Here, conceptual knowledge was defined as tasks or questions that require students to explain interconnections between facts or explain the biology phenomenon using disciplinary core ideas. Another study analyzed tasks in Grade 9 secondary biology lessons. This study reaffirmed that German biology lessons were usually orchestrated using low cognitive-level tasks (Jatzwauk et al., 2008). This study found a positive that time spent on task related activities was a predictor of student learning, specifically when students showed very little topic-related prior knowledge. Thus, the presented study used video-based observation method as a tool to analyze teachers' use of *challenging tasks* in German biology lessons.

Hypotheses

To summarize, several empirical studies have investigated the influence of cognitively activating instruction on students' learning outcomes like situational interest and knowledge test. However, the presented study investigated the influence of using *challenging tasks (high-level cognitive processing tasks & higher content complexity tasks)* in classrooms on *students' cognitive knowledge structure*, when controlled for students' prior knowledge related to the topic, motivation and interest-related variables. Therefore, we investigated following hypotheses in the study presented here:

> H1: There is a positive influence of using *high-level cognitive processing tasks* on the *students' topic-related cognitive knowledge structure*, measured using student-drawn concept maps (Brown, 1994; Klieme & Bos, 2000; Lipowsky et al., 2009; Stein & Lane, 1996).

> H2: There is a positive influence of using *higher content complexity tasks* on *students' cognitive knowledge structure*, measured using a concept mapping exercise (Jacobs et al., 2003, 2006).

Method

The presented study is part of a larger teaching effectiveness project funded by Federal Ministry of Education & Research (BMBF). A quasi-experimental pre-post design was used to collect classroom teaching videos and student tests-questionnaire data. All data were collected from Gymnasium secondary schools of the state North Rhine-Westphalia, Germany, in a prior study (Wadouh, 2008; Jatzwauk, 2007; DFG FG 511).

Research design

In the first phase of data collection, students from 47 participating grade 9 classrooms were given: (1) a pre-test to evaluate prior knowledge about the topic and (2) an interest and motivation in biology questionnaire. In the second phase, we videotaped one biology lesson per teacher on the topic 'blood and circulatory system'. In the final phase of data collection, all students completed: (1) a post-unit knowledge test and (2) a concept mapping exercise (see Figure 2). Some of the previous studies that used this dataset to describe biology teaching processes include: Jatzwauk (2007), Wüsten (2010), Wadouh et al. (2014), and so on.

As school and teacher participation in the study was voluntary, we collected videos and other data from biology teachers who gave their consent in the beginning of this study.

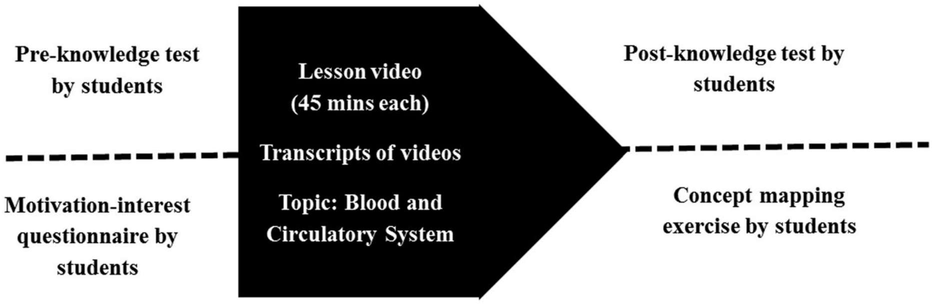

Figure 2. Pre–post design for collecting data used in this empirical investigation.

Furthermore, this study examined the correlation between 'instructional components' and students' knowledge structure about a given topic. Hence, we standardized the content by videotaping lessons pertaining to one topic 'blood and circulatory system'. Videotaping lessons pertaining to a common topic helped to administer topic-related pre–post tests and post-unit concept mapping exercises that helped to evaluate students' knowledge about 'blood and circulatory system' (Hugener et al., 2009; Praetorius, Pauli, Reusser, Rakoczy & Klieme, 2014).

Participants

Forty-seven biology lesson videos (approx. 45 min each) on the topic 'blood and circulatory system' were collected from Grade 9 classrooms ($N = 1214$ students) of the state North Rhine-Westphalia. Teachers who participated in this study were on average 46 years old (min = 28, max = 60, SD = 10; $N = 47$) with 18 years of teaching experience (min = 1, max = 31, SD = 11; $N = 47$). However, our study examined 38 out of 47 biology lesson videos. Here, 7 out of 47 classrooms were dropped because students from these classrooms could not participate in the concept mapping exercise. Two more classrooms were dropped because these lessons had very few utterances related to the content. Average class size of participating 38 classrooms was approximately 26 students (min = 20; max = 31; SD = 2.4).

Instruments

Concept maps

Students constructed concept maps based on 15 terms related to the topic blood and circulatory system. Followinig terms were provided to construct the concept maps: Heart, blood groups, cellular respiration, circulation, blood, muscles, nutrients, blood donation, blood cells, pathogens, oxygen, arteries, blood pressure, exercise and energy (Wadouh, 2008). Quantitative scoring system based on a frequency was used to evaluate concept maps (Friege & Lind, 2000; Stoddart, Abrams, Gasper, & Canaday, 2000). Student-drawn concept maps were scored for the following variables:

(1) number of concept maps drawn (i.e. whether the concept maps consisted of disconnected networks);

(2) number of total relations with legible labels (i.e. the number of relations and number of cross-relations drawn);
(3) number of relations drawn with technically correct explanations;
(4) number of relations drawn with deeper explanations for the relations drawn.

We would like to mention here that 'number of total relations with legible labels' (i.e. variable 2 mentioned above) included two sub-variables: (1) total number of relations drawn between terms pertaining to the topic 'blood and circulatory system' (e.g. blood and heart; blood and blood pressure) and (2) total number of cross-relations drawn between terms related to the topic 'blood and circulatory system' and other topics like 'immune biology' (e.g. blood and pathogens; blood cells and cellular respiration).

Analyzed by two raters, student-drawn concept maps showed satisfactory values of Cohens' kappa coefficients (κ) (Landis & Koch, 1977; Wirtz & Caspar, 2002). Kappa values for inter-rater agreement were: number of total relations: $\kappa = 0.93$, number of correct relations: $\kappa = 0.61$ and depth of description used to explain the relation between two terms: $\kappa = 0.73$ (Wadouh, 2008). These data were also used for another study to investigate the influence of *teaching interconnected complex domain content* on *students' cognitive knowledge structure* (Wadouh, 2008; Wadouh et al., 2014).

However, we used *principal component analysis with varimax rotation* to extract factors from the four concept maps variables mentioned above. As depicted in Table 2, the principal component analysis resulted in a single component. No sub-scales could be extracted. Hence, the *z*-standardized values of four concept map variables were added together to form one aggregate variable: *students' cognitive knowledge structure*. Four variables that were *z*-standardized and added to form this variable included: (1) number of concept maps drawn, (2) number of total relations with legible labels, (3) number of relations with technically correct explanations and (4) number of relations with deeper explanations for the relations drawn.

Category system for video coding

A three-step coding scheme was theoretically devised to analyze teachers' use of *challenging tasks* in lesson videos. All 38 videos were event coded with the software Videograph (Rimmele, 2002). Thus, each teacher-initiated task (an event) was coded in three steps described below:

(1) At first, the coders observed the teacher-initiated tasks and coded them for 'new teacher-initiated tasks' (e.g. why are red blood cells important?) and 'connecting teacher tasks' – used as links, connectors to continue the discussion (e.g. let me ask some one else, you: why are they (i.e. red blood cells) important?) (see Table 3).

Table 2. Principal component matrix (Varimax-rotated) of 4 variables analyzed in student concept maps.

Student concept map variables	Loading
Number of correct relations drawn	.91
Number of total relations drawn	.89
Number of relations with deeper explanation for connections drawn	.68
Number of concept maps drawn	.65

Note: The matrix shows the loadings of the 4 variables on one factor. Only loadings >.4 are shown.

Table 3. New and connecting teacher initiated tasks (function of tasks during teacher-student interactions).

Category	Description and indicators	Example
New teacher initiated tasks	(1) Task that presents new content and facilitates the process of new content development. (2) Task refers to new information, text or artifacts/ material presented in classrooms. (3) First task of teaching conversation (4) Teacher formulates new task without expecting any response to a previous task.	T: Explain how cells are supplied with oxygen. S: The oxygen inhaled via lungs reaches cells. T: What happens to oxygen in cells? (Although teacher uses the same word oxygen in new task but asks about the molecular processing of oxygen in cells). T: What is a blood type? S: A, B, AB and O. T: And how is knowledge of blood type important for blood transfusion?
Connecting teacher task to generate further responses	(1) Teacher passes the same task to new student. (2) Teacher reformulates a task (further clarification). However, this new task has same content as the previous task.	T: What is a blood type? S: A, B, AB T: Anyone else, What are blood groups?
Connecting teacher task after a student answer	(1) Teacher formulates a task after student answer for further clarification, justification, error-correction. (2) Teacher asks for clarification of terms used by student while answering previous task. (3) Task relates to whole or part of student answer.	T: What is a blood type? S: A, B, AB and zero. T: What is the meaning of A, B, AB and O? (A, B, AB and O should be further explained). T: What happens when you mix the blood of Hanna with the blood of Tom? S: It clumps. T: Why/ How?
None of above	Tasks with no content or tasks which could not be connected to the content being discussed during the lesson.	T: Where is your assignment? Give me? T: Did you all bring your books/ student card.

Source: Adapted from Rixius (2016, in preparation).

Cohens' kappa coefficient ($\kappa = 0.72$) indicated a substantial inter-rater agreement between observers coding new and connecting teacher tasks.

(2) In the second step, each 'new teacher-initiated task' was coded for level of cognitive processing (i.e. Low level: repetition, summary, list, describe, etc.; High level: explain, justify, formulate hypothesis, interpret, etc.) (see Table 4). Cohens' kappa coefficient for observers coding high- and low-level cognitive processing tasks was satisfactory ($\kappa = 0.68$).

(3) In the third step, 'new teacher-initiated tasks' were coded for their level of content complexity (i.e. fact, linking or conceptual-level content) (see Table 5). Cohens' kappa for coding complexity of task content was again satisfactory ($\kappa = 0.72$).

Observer coding from this three-step process was used to report total number of (1) *high-level cognitive processing tasks* used in each class and (2) *higher content complexity tasks* used in each class.

Students' prior knowledge

All students from participating biology classrooms completed the 31-item factual knowledge test before and after the teaching unit 'blood and circulatory system'. This instrument measured students' factual knowledge about 'blood and circulatory system' (Wadouh, 2008). This test consisted of multiple-choice items ($N = 25$), match the terms ($N = 1$),

Table 4. Cognitive processing level of teacher initiated tasks.

Category	Description	Example
Low level cognitive processing (Cognitive objectives levels – Knowledge, Comprehension) (Anderson et al., 2001; Bloom, 1972; Blooms Taxonomy, n.d.; Krathwohl, 2002)		
Summary	Tasks that ask students to concisely summarize content in their own words.	T: What were some key learning points in todays' discourse?
Define, List, Specify terms	Tasks asking for definitions, naming of specific technical biological terms, verification of given definition, examples, analogies, etc.	T: Give names of different types of blood cells. T: Give percent rates of occurrence of different types of blood cells.
Describe	Description about how something works, appears, etc. Description of actual circumstances, structures, contexts or procedures with or without pictures, graphs or diagrams.	T: Describe the structure of erythrocytes. T: What relation is shown in this diagram?
Classify Arrange	Characteristics, elements, members should be classified into categories.	T: Arrange the images of immune response to individual texts.
Compare Contrast	Tasks asking to state differences or similarities between elements, members, features, contexts.	T: How are platypus and mammals similar and different from each other?
High level cognitive processing (Cognitive objectives levels – Application, Analysis, Synthesis & Evaluation) (Anderson et al., 2001; Bloom, 1972; Blooms Taxonomy, n.d.; Krathwohl, 2002)		
Explain, give reasons, Justify	Tasks that ask for logical explanation, justification of phenomenon using biology concepts, principles or disciplinary core ideas.	T: Explain why is water the most appropriate habitat for fish/ Explain why fish live in water? T: How do you know that Mr. Roth's blood type is A+
Design an experiment/ Formulate hypothesis	Tasks that ask students to design an experiment and formulate hypothesis to prove a scientific phenomenon or observed process	T: What factors do you think influence the photosynthetic activity of plant? Formulate hypothesis to investigate various factors
Interpret and Analyze	Tasks that ask students to draw substantive conclusions after evaluating multiple evidences, clues	T: Observe the results from clumping reactions of various blood antigen-antibodies and explain what Mr. Roth's blood type is?
Reflect, rethink	Tasks that ask students to recheck the answer given by another student to confirm or refute its accuracy	T: Consider again whether the platypus actually descended from birds?
Evaluation	Asking opinion/ judgment/ justification	T: Should we donate our organs after death? T: Should blood donation be a common practice for all healthy human beings?

Source: Adapted from Rixius (2016, in preparation).

draw and label diagram ($N = 1$) and filling the gaps ($N = 4$). Student pre-unit performance in this test was used as a covariate in the study presented here.

Students' motivation and interest to learn biology

Questionnaire developed by Wild, Hofer and Pekrun (2001) and adapted for the subject biology was completed by students from all participating biology classrooms in the beginning of the teaching unit 'blood and circulatory system'. It consisted of four scales: Interest in subject biology ($N = 3$ items, $\alpha = 0.89$), Interest in subject-related activities ($N = 3$ items, $\alpha = 0.56$), Intrinsic Motivation ($N = 7$ items, $\alpha = 0.83$) and Extrinsic Motivation ($N = 9$ items, $\alpha = 0.54$). Students' rated their agreement on a 4-point Likert scale ranging from 0 (not true) to +3 (true). Four sub-scales of this instrument showed good reliability (Cortina, 1993; Wadouh, 2008). Z-standardized values of individual student scores on all four sub-scales of this instrument were used to calculate the four motivation and

Table 5. Content complexity of teacher initiated tasks.

Category	Example
Task content - Fact level	
When a task asks for fact or more facts. The task could ask for a definition, features, specific properties, technical term used in domain.	T: Tell me one task of erythrocytes! S: red blood cells carry oxygen. (Task). T: You know the blood groups A, B, AB, and .? S: O. (Here, one specific term is requested as a response.) T: What do you mean by clumping? S: agglutination. (Here, students are asked to answer by mentioning what do they understand by the term clumping.)
Task content - Linking level	
When a task requires that students to explain the interconnections between facts or present an explanation for the biology process or phenomenon using a set of interconnected facts.	T: When does agglutination of blood happen? S: In case of injury (temporal condition of agglutination) (own example).
Additionally, linking level tasks could demand explanation about how facts influence each other or a third factor, dependence of two factors on each other, conditions required for occurrence of biological phenomenon, causal relations for biological processes, functional explanation of biological terms, etc.	T: What happens during blood clumping? S: Red blood cells are combined. (Process, not concept). T: What happens if a person with A type blood gets B type blood? S: The anti B antibody from A blood type person will bind with antigen B from B type blood received during transfusion (interaction between antigens and antibodies must be explained to answer the question)
Task content - Concept level	
When tasks require explanation of causal relation using a biology concept or disciplinary core idea. When a task requires explanation based on underlying biological concept or disciplinary core idea.	T: Explain the process of oxygen transport at organ level and cellular level. S: The oxygen passes from lungs via pulmonary vein and from there to heart and then into the aorta. Aorta helps transport oxygen to various organs in the body. If you look at the cellular level, oxygen is bonded via hemoglobin and either stored or passed over to various organs in the body. (You will be prompted to explain a biology idea or phenomenon using basic concepts. In this case, students are asked to explain the process of oxygen transport - both at organ and cellular level).

Source: Adapted from Rixius (2016, In preparation).

interest-related variables: (1) *students' extrinsic motivation*, (2) *students' intrinsic motivation*, (3) *students' interest in subject biology* and (4) *students' interest in biology activities*.

The above-described variables were used as control variates to examine the influence of instructional features (here: *challenging tasks*) on students' cognitive knowledge structure. These data were also used in another study to examine the influence of *teaching interconnected complex domain content* on *students' cognitive knowledge structure* (Wadouh et al., 2014).

Data analysis using linear multilevel modeling

This study investigated the influence of class-level predictors *high-level cognitive processing tasks* & *higher content complexity tasks* on *students' cognitive knowledge structure*. Additionally, *students' prior knowledge*, *students' extrinsic motivation*, *students' intrinsic motivation*, *students' interest in subject biology* and *students' interest in biology activities* were used as covariates, while examining the influence of teachers' *use of challenging tasks* on *students' cognitive knowledge structure*. As explained earlier, this study collected hierarchically nested data, linear multilevel modeling in SPSS was used to test the hypotheses formulated for this study (Field, 2009; Heck, Thomas, & Tabata, 2013).

Results

Results of this study are divided into two parts. The first part presents descriptive statistics pertaining to the independent variables, dependent variables and control covariates investigated for this study. The second part describes findings from linear multilevel modeling in SPSS where *high-level cognitive processing tasks & higher content complexity tasks* were used as class-level predictors to study to investigate their influence on *students' cognitive knowledge structure.*

Descriptive statistics

Videotaped lessons

All 38 videotaped lessons were first coded for the frequency of new teacher-initiated tasks. We found 1704 instances where teacher-initiated new tasks during classroom discourse (min = 16; max = 88; SD = 18.42). Later, each *new teacher-initiated task* was coded for their level of cognitive processing and complexity of task content. 366 *high-level cognitive processing tasks* (min = 0; max = 32, SD = 6.63) were found in 38 investigated biology lessons. Higher level cognitive processing tasks involved deeper information-processing situations like justifying, formulating hypotheses, interpreting, reflecting and evaluating (see Table 4). Furthermore, 614 *higher content complexity tasks* (min = 0; max = 37, SD = 9.50) were found in biology lessons. Higher content complexity tasks involved linking and conceptual-level content (see Table 5).

Students' prior knowledge test

The 31-item testing instrument measuring students knowledge related to the topic 'blood and circulatory system' exhibited satisfactory internal consistency ($\alpha = 0.72$). Mean task difficulty was 0.64 (min = 0.18; max = 0.89) and selectivity ranged from 0.04 to 0.40 (Wadouh, 2008). Student performance in this test, before the teaching unit 'blood and circulatory system', was used as a control variable, while investigating the influence of *challenging tasks* on students' performance in the concept mapping exercise.

Findings from linear multilevel modeling in SPSS

As explained earlier, data collected for this study included both class-level and individual student-level variables. We also calculated the intra-class correlation (ICC), which indicates how students from various classes differed in their performance in the concept mapping exercise. When *students' cognitive knowledge structure* (i.e. aggregate student performance in concept mapping) was used to generate the 'Restricted Maximum Null Model', ICC value calculated was about 0.070. This means that 7.0% variance in *students' cognitive knowledge structure* was located at the class level. However, as data were hierarchically nested, it warranted the use of multilevel modeling to examine the correlations proposed in hypotheses.

Models showing influence of high-level cognitive processing tasks on students' cognitive knowledge structure

To explore the influence of class-level predictor *high-level cognitive processing tasks* on outcome variable *students' cognitive knowledge structure*, we generated various 'Maximum Likelihood Random Intercept Models'. Initial model (see Model 1: Table 6) was created with *high-level cognitive processing tasks* as a class-level predictor of *students' cognitive knowledge structure*. Later on, covariates like *students' prior knowledge*, *students' extrinsic motivation, students' intrinsic motivation, students' interest in the subject biology*, *students' interest in subject-related activities* were gradually introduced as grand mean-centered predictors. 'Maximum Likelihood (ML)' and 'Bayesian Information Criterion (BIC)' estimates for these models were compared to choose the best model that predicted *students' cognitive knowledge structur*e. Thus, the final model was chosen where ML and BIC estimates showed a significant decline (Field, 2009, p. 753) (see Model 2 & 3: Table 6). BIC value of this model (see Model 3: Table 6) was 4421. 42 and this estimate (BIC) did not decline further, when additional student-level covariates were added. The final model (see Model 3: Table 6) depicted that *students' prior knowledge* related to blood and circulatory had significant, but very low impact on *students' cognitive knowledge structure* ($\beta = 0.01$). Besides that, *students' interest in biology activities* improved the ML and BIC estimates (see Model 2,3: Table 6) but did not correlate with *students' cognitive knowledge structure*, while *high-level cognitive processing tasks* showed a moderate impact on *students' cognitive knowledge structure* ($\beta = 0.07$). Here, β represents the partial regression coefficient or unstandardized regression estimates, presented as 'estimates of fixed effects' in SPSS (see Model 1, 2 & 3: Table 6) (Bring, 1994; Heck et al., 2013). In the end, it is important to note that several models were created in SPSS with covariates related to students' intrinsic motivation, extrinsic motivation and interest in biology. However, we did not report

Table 6. Maximum likelihood random intercept models for 'High-level cognitive processing tasks'.

Dependent variable – Students Knowledge Structure (SKS)						
	Model 1 (Model 1SKS) Estimate		Model 2 (Model 2SKS) Estimate		Model 3### (Model 3SKS) Estimate	
Predictors	ß	SE	ß	SE	ß	SE
Intercepts						
Class level						
High-level cognitive processing tasks	0.07**	0.03	0.07*	0.02	0.07**	0.02
Individual level						
Pre-knowledge			0.01**	0.001	0.01**	0.001
Interest in subject activities					−0.08	0.20
BIC	3666.15		3435.08		3432.89	
(ML)	(3639.83)		(3402.44)		(3393.74)	

Note: Maximum likelihood (ML) & Schwartźs Bayesian Criterion/ Bayesian Information Criterion (BIC) were used for model selection. All Individual-level variables were grand mean centered before creating likelihood models. Dependent variable – students' knowledge structure was the composite variable, calculated by adding the z-standardized values for (1) number of concept maps drawn, (2) number of total relations drawn with legible labels, (3) number of relations with technically correct explanations; (4) depth of description of relations drawn between two terms.
ß = SPSS regression weights (Estimate of fixed effects); SE = standard error for estimates of fixed effects.
***$p \leq .005$.
**$p \leq .01$.
*$p \leq .05$.
+$p \leq .10$.
###Best-fit model (based on BIC and Maximum Likelihood comparison).

Table 7. Maximum likelihood random intercept models for 'Higher content complexity tasks'.

Dependent variable – Students' Knowledge Structure (SKS)		
	Model 1	
	(Model 1SKS)	
Predictors	ß	SE
Intercepts		
Class level		
Higher content complexity tasks	.01	.02
Individual level		
Pre-knowledge		
Interest in subject		
Interest in subject activities FVV		
BIC (ML)	3673.59 (3647.27)	

Note: Maximum likelihood (ML) & Schwartźs Bayesian Criterion/ Bayesian Information Criterion (BIC) were used for model selection. All Individual-level variables were grand mean centered before creating likelihood models. *D*ependent variable – students' knowledge structure was calculated as a sum of *z*-standardized values for (1) number of concept maps drawn, (2) number of total relations drawn with legible labels, (3) number of relations with technically correct explanations; (4) depth of description of relations drawn between two terms.

ß = SPSS regression weights (Estimate of fixed effects); SE = standard error for estimates of fixed effects.

###Best-fit model (Based on BIC and Maximum Likelihood comparison.).

these models as the covariates neither imporved the ML and BIC values nor significantly correlated with *students' cognitive knowledge structure*.

Multilevel models showing influence of higher content complexity tasks on students' cognitive knowledge structure

In order to explore the influence of class-level predictor *higher complexity tasks* on student outcome variable – *students' cognitive knowledge structure*, we generated various 'Maximum Likelihood Random Intercept Models'. The initial model (see Model 1: Table 7) was created with *higher content complexity tasks* as a class-level predictor of *students' cognitive knowledge structure*. Later on, student-level covariates were gradually added. However, these maximum likelihood random intercept models did not show any significant influence of class-level predictor *higher complexity tasks* on *students' cognitive knowledge structure*.

Discussion

This section will first endeavor to relate study aims with key findings described in the previous section. Later, we will discuss methodological and generalizability concerns pertaining to this study. Last, this section will briefly describe the key implications of the results obtained and present perspectives about the way ahead for future studies.

First, this study successfully used the three-step coding manual to objectively and reliably identify *high-level cognitive processing tasks* (Anderson et al., 2001; Bloom, 1972; Blooms Taxonomy, n.d.; Krathwohl, 2002) and *higher content complexity tasks* (Hiebert et al., 2003; Jacobs et al., 2003, 2006; Wadouh et al., 2014) used in biology lesson videos.

However, we assumed that 'complexity of task content' and 'cognitive processing level of tasks' are two defining characteristics of *challenging tasks* (Blumenfeld & Meece, 1988).

Furthermore, this study confirmed the first hypothesis, which proposed that *high-level cognitive processing tasks* will positively predict *students' cognitive knowledge structure* in biology (Brown, 1994; Klieme & Bos, 2000; Lipowsky et al., 2009; Stein & Lane, 1996). These results are in line with findings from a similar investigation using Grade 6 biology lessons that found a positive influence of *high-level cognitive processing tasks* in biology classrooms on students' factual knowledge and structural knowledge (Förtsch et al., 2016b).

In this regard, we used student-level covariates related to prior knowledge, motivation and interest, while constructing linear multilevel models. Comparing ML and BIC values for these models showed a significant but low impact of *students' prior knowledge* on their performance in concept mapping. Students' prior knowledge related to the topic usually consists of factual information and pre-concepts, while concept mapping requires students to describe links or conceptual relation between any two terms. This could be one reason for the minimal impact of prior knowledge on performance in concept mapping. Moreover, *students' interest in biology activities* improved the ML and BIC values and hence was retained in the final model (Field, 2009, p. 753). However, *students' interest in biology activities* along with other interest and motivation variables did not correlate with *students' cognitive knowledge structure*. One reason for such findings could be the way teaching lessons were implemented. As shown in the descriptive section, teachers rarely used challenging tasks during classroom discourse and thus teacher–student interactions could not activate students to benefit from their individual interest and motivation attitudes to acquire in-depth understanding of the topic being taught.

Furthermore, we could not confirm the second hypothesis which states that *higher content complexity tasks* will positively influence *students' cognitive knowledge structure* in biology classrooms. Our investigation used a three-level coding manual for coding content complexity: (1) Tasks at the fact level; (2) Tasks at the linking level; (3) Tasks at the conceptual level for identifying *higher content complexity tasks* in biology lessons (see Table 5). Here, descriptive statistics shows that very few *conceptual-level tasks* were used by teachers during the classroom discourse. This could be one reason why our statistical analysis could not find a correlation between *higher content complexity tasks* and *students' cognitive knowledge structure*. Future studies in this regard could include interventions where teachers are trained/ encouraged to use *tasks involving conceptual-level content* to meaningfully examine the correlation between use of *higher content complexity tasks* and *students' cognitive knowledge structure*.

Limitations

Data pertaining to this study were collected from the German state of North Rhine-Westphalia. As participation in this study was not compulsory, we collected data from schools and biology teachers who gave their consent in the beginning of the study. Such a strategy of data collection could present concerns regarding the generalizability of results obtained. Nevertheless, it must be noted that empirical studies that use external observer ratings for analyzing data usually collect one or few lessons per teacher (Praetorius, Pauli, Reusser, Rakoczy, & Klieme, 2014). Researchers in this regard have argued that instructional competence does not change in a short time and hence daily teaching practice will show sufficient stability, especially in the absence of planned interventions or training.

Furthermore, due to the resource and time constraints, the presented study videotaped one lesson per teacher ($N = 47$ teachers) about the topic 'blood and circulatory system'. As mentioned earlier, collecting data related to one common topic (1) helped to standardize the content (2) facilitated comparison of instructional practices and (3) helped to collect pre–post assessment data related to the topic 'blood and circulatory system' to examine the correlations (Hugener et al., 2009). However, the limited sample size and use of lessons pertaining to one topic could again raise concerns about the generalizability of results presented. Future studies in this regard could videotape multiple lessons related to two or more topics to triangulate data and enhance the validity and generalizability of results obtained (Bush, 2012; Mathison, 1988).

To conclude, findings presented here contribute to the existing attempts towards understanding effective science instruction. These results could provide significant ideas for teacher trainers as well as in-service and future teachers to refine their practice and facilitate student understanding about a given topic. These results and ensuing discussions would be significantly informative for designing video-based in-service teacher training programs for enhancing teaching effectiveness in science, especially biology.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This paper is part of the video study 'Teacher Communication in Biology Classrooms' funded by the Federal Ministry of Education and Research (01JG1074).

Notes on contributors

Ms. Jigna Nawani is a doctoral candidate at the Biology Education Group, Department I, Faculty of Biology, Ludwig-Maximilians-Universität München. She received her Masters in Environment Sciences from Gujarat University. Her current research interests include investigating the effectiveness of teaching practices used in biology classrooms.

Ms. Julia Rixius was a doctoral student at the at the Biology Education Group, Department I, Faculty of Biology, Ludwig-Maximilians-Universität München. Currently, she is working as a biology teacher.

Prof. Dr. Birgit J. Neuhaus is chair of the Biology Education Group, Department I, Faculty of Biology, Ludwig-Maximilians-Universität München. Her research interests include investigating biology teachers' professional knowledge as well as the effectiveness of teaching practices used in German biology classrooms.

References

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*(2), 313–337.

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

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York, NY: Grune.
Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York, NY: Holt, Rinehart & Winston.
Blömeke, S., Risse, J., Müller, C., Eichler, D., & Schulz, W. (2006). Analyse der Qualität von Aufgaben aus didaktischer und fachlicher Sicht. Ein allgemeines Modell und seine exemplarische Umsetzung im Unterrichtsfach Mathematik [Analysis of the quality of tasks from a didactic and technical point of view. A general model and its exemplary implementation in teaching mathematics]. *Unterrichtswissenschaft*, *34*(4), 330–357.
Bloom, B. S. (Ed.). (1972). *Taxonomy of education objectives book 1-cognitive domain*. New York: David McKay.
Blooms Taxonomy. (n.d.). Retrieved March 28, 2015, from http://www.nyiteez.org/EDIT652/bloom.htm
Blumenfeld, P. C., & Meece, J. L. (1988). Task factors, teacher behavior, and students' involvement and use of learning strategies in science. *The Elementary School Journal*, *88*(3), 235–250.
Bring, J. (1994). How to standardize regression coefficients. *The American Statistician*, *48*(3), 209–213.
Brown, H. D. (1994). *Teaching by principles* (pp. 152–157). Englewood Cliffs, NJ: Prentice Hall.
Bush, T. (2012). Authenticity in research: Reliability, validity and triangulation. In A. R. Briggs, M. Morrison, & M. Coleman (Eds.), *Research methods in educational leadership and management* (3rd ed.). London: Sage.
Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.
Cortina, J. M. (1993). What is coefficient alpha? Examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104.
Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.
Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134.
Doyle, W. (1979). Classroom tasks and students' abilities. In P. L. Peterson & H. J. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 183–209). Berkeley, CA: McCutcheon.
Doyle, W. (1983). Academic work. *Review of educational research*, 53(2), 159–199.
Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.
Ergönenc, J., Neumann, K., & Fischer, H. E. (2014). The impact of pedagogical content knowledge on cognitive activation and student learning. In H. E. Fischer, P. Labudde, K. Neumann, & J. Viiri (Hrsg.), *Quality of instruction in physics – Results from a trinational video study* (pp. 145–160). Münster: Waxmann.
Field, A. P. (2009). Discovering statistics using SPSS: and sex and drugs and rock 'n' roll (3rd ed.). London: Sage publications.
Förtsch, C., Werner, S., Dorfner, T., von Kotzebue, L., & Neuhaus, B. J. (2016a). Effects of cognitive activation in biology lessons on students' situational interest and achievement. *Research in Science Education*, 1–20. doi:10.1007/s11165-016-9517-y
Förtsch, C., Werner, S., von Kotzebue, L., & Neuhaus, B. J. (2016b). Effects of high-complexity and high-cognitive-level instructional tasks in biology lessons on students' Acquirement of Factual and Structural Knowledge. Manuscript submitted for publication.
Friege, G., & Lind, G. (2000). Begriffsnetze und expertise. In H. Fischler & J. Peuckert (Eds.), *Concept mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie* [Concept mapping in research projects on physics and chemistry didactics] (Vol. 1, pp. 147–178). Berlin: Logos Verlag.
Garner, R., & Gillingham, M. G. (1991). Topic knowledge, cognitive interest, and text recall: A microanalysis. *Journal of Experimental Education*, *59*, 310–319.

Gerstenmaier, J., & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive [Knowledge acquisition from a constructivist perspective]. *Zeitschrift für Pädagogik*, *41*(6), 867–888.

Glaser, R. (1991). Expertise and assessment. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition*. Englewood Cliffs, NJ: Prentice-Hall.

Gredler, M. E., Broussard, S. C., & Garrison, M. E. B. (2004). The relationship between classroom motivation and academic achievement in elementary school aged children. *Family and Consumer Sciences Research Journal*, *33*(2), 106–120.

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H. W., Larose, S., & Boivin, M. (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *British Journal of Educational Psychology*, *80*(4), 711–735.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2013). *Multilevel and longitudinal modeling with IBM SPSS*. New York: Routledge.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., … Stigler, J. (2003). Teaching mathematics in seven countries: Results from the TIMSS 1999 video study. *Education Statistics Quarterly*, *5*(1), 7–15.

Hugener, I., Pauli, C., Reusser, K., Lipowsky, F., Rakoczy, K., & Klieme, E. (2009). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning and Instruction*, *19*(1), 66–78.

Jacobs, J., Garnier, H., Gallimore, R., Hollingsworth, H., Givvin, K. B., Rust, K., … Stigler, J. (2003). Third international mathematics and science study 1999 video study technical report: Volume 1 – Mathematics. Technical Report. NCES 2003-012. *National Center for Education Statistics*.

Jacobs, J. K., Hiebert, J., Givvin, K. B., Hollingsworth, H., Garnier, H., & Wearne, D. (2006). Does eighth-grade mathematics teaching in the United States align with the NCTM Standards? Results from the TIMSS 1995 and 1999 video studies. *Journal for Research in Mathematics Education*, 37 (1), 5–32.

Jatzwauk, P. (2007). *Aufgaben im Biologieunterricht: eine Analyse der Merkmale und des didaktisch-methodischen Einsatzes von Aufgaben im Biologieunterricht* [Tasks in biology lessons: An analysis of characteristics and the use of didactic methodological tasks in biology lessons]. Berlin: Logos-Verlag.

Jatzwauk, P., Rumann, S., & Sandmann, A. (2008). Der Einfluss des Aufgabeneinsatzes im Biologieunterricht auf die Lernleistung der Schüler–Ergebnisse einer Videostudie [The effect of usage of tasks in biology education on learning performance – A video study]. *Zeitschrift für Didaktik der Naturwissenschaften*, *14*, 263–283.

Kinchin, I. M. (2011). Visualising knowledge structures in biology: Discipline, curriculumand student understanding. *Journal of Biological Education*, *45*(4), 183–189.

Kinchin, I. M., & Hay, D. B. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research*, *42*(1), 43–57.

Klieme, E., & Bos, W. (2000). Mathematikleistung und mathematischer Unterricht in Deutschland und Japan [Mathematics performance and mathematics teaching in Germany and Japan]. *Zeitschrift für Erziehungswissenschaft*, *3*(3), 359–379.

Krapp, A. (2002). An educational-psychological theory of interest and its relation to SDT. In E. L. Deci, & R. M. Ryan (Eds.), *The handbook of self-determination research* (pp. 405–427). Rochester, NY: Rochester University.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (Vol. 8). New York: Springer Science & Business Media.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, *19*(6), 527–537.

Mathison, S. (1988). Why triangulate? *Educational Researcher*, *17*(2), 13–17.

Mayer, R. E. (1998). Cognitive theory for education: What teachers need to know. In N. Lambert & B. L. McCombs (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 351–377). Washington, DC: American Psychological Association.

Mayer, R. E. (2009). Constructivism as a theory of learning versus constructivism as a prescription for instruction. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: success or failure?* (pp. 184–200). New York: Routledge.

Neumann, K., Fischer, H. E., & Sumfleth, E. (2008). Vertikale Vernetzung und kumulatives Lernen im Chemie-und Physikunterricht [Vertical linking and cumulative learning in chemistry and physics teaching]. In E.-M. Lankes (Ed.), *Pädagogische Professionalität als Gegenstand empirischer Forschung* (pp. 141–151). Munster: Waxmann.

Osborne, R. J., & Wittrock, M. C. (1985). The generative learning model and its implications for science education. *Studies in Science Education*, *12*, 59–87.

Pauli, C., Drollinger-Vetter, B., Hugener, I., & Lipowsky, F. (2008). Kognitive Aktivierung im Mathematikunterricht [Cognitive activation in mathematics teaching]. *Zeitschrift für Pädagogische Psychologie*, *22*(2), 127–133.

Praetorius, A. K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, *31*, 2–12.

Prenzel, M. (1988). *Die Wirkungsweise von Interesse. Ein Erklärungsversuch aus pädagogischer Sicht* [Functioning of interest. An explanation from pedagogical perspective]. Opladen: Westdeutscher Verlag.

Rakoczy, K., Klieme, E., Drollinger-Vetter, B., Lipowsky, F., Pauli, C., & Reusser, K. (2007). Structure as quality feature in mathematics instruction of the learning environment vs a structured presention of learning content. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (pp. 101–120). Münster: Waxmann.

Resnick, L. B. (1989). *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum.

Rimmele, R. (2002). *Videograph. Multimedia-Player zur Kodierung von Videos* [Video-graph. Multimedia-Player for coding videos]. Kiel: IPN.

Rixius, J. (2016). *Das Klassengespräch im Biologieunterricht* [Teacher-student talk in biology lessons] (Dissertation Manuscript in preparation). Munich: Ludwig-Maximilians-Universität München.

Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., … & Stigler, J. (2006). *Teaching science in five countries: Results from the TIMSS 1999 video study* (Statistical Analysis Report. NCES 2006-011. National Center for Education Statistics).

Ruiz-Primo, M. A., & Shavelson, R. J. (2005). Problems and issues in the use of concept map in science assessment. *Journal of Research in Science Teaching*, *33*(6), 569–600.

Schiefele, H., Krapp, A., Prenzel, M., Heiland, A., & Kasten, H. (1983, July). *Principles of an educational theory of interest*. 7th Meeting of the International Society for the Study of Behavioral Development in Munich.

Schiefele, U., & Schreyer, I. (1994). Intrinsische Lernmotivation und Lernen. Ein Überblick zu Ergebnissen der Forschung [Intrinsic motivation and learning. An overview of results from research]. *Zeitschrift für Pädagogische Psychologie*, *8*(1), 1–13.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A metaanalysis of research. In: K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale, NJ: Erlbaum.

Schönborn, K., & Bögeholz, S. (2009). Knowledge transfer in biology and translation across external representations: Experts' views and challenges for learning. *International Journal of Science and Mathematics Education*, *7*(5), 931–955.

Seidel, T., Prenzel, M., Rimmele, R., Herweg, C., Kobarg, M., Schwindt, K., & Dalehefte, M. (2007). Science teaching and learning in German physics classrooms. Findings from the IPN video study.

In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (pp. 79–99). Münster: Waxmann.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, *2*(1), 50–80.

Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2000). Concept maps as assessment in science inquiry learning-a report of methodology. *International Journal of Science Education*, *22*(12), 1221–1246.

Wadouh, J. (2008). *Vernetzung und kumulatives Lernen im Biologieunterricht der Gymnasialklasse 9 [Interconnectedness and cumulative learning in biology teaching of grade 9]*. Dissertation University of Duisburg-Essen.

Wadouh, J., Liu, N., Sandmann, A., & Neuhaus, B. J. (2014). The effect of knowledge linking levels in biology lessons upon students' knowledge structure. *International Journal of Science and Mathematics Education*, *12*(1), 25–47.

Wild, E., Hofer, M., & Pekrun, R. (2001). Psychologie des Lerners. In A. Krapp & B. Weidemann (Eds.), *Pädagogische Psychologie* [Pedagogical Psychology] (pp. 218–241). Weinheim: Beltz Psychologie Verlag Union.

Wirtz, M. A., & Caspar, F. (2002). *Beurteilerübereinstimmung und beurteilerreliabilität* [Inter-rater reliability and Inter-rater agreement]. Göttingen: Hogrefe.

Wüsten, S. (2010). *Allgemeine und fachspezifische Merkmale der Unterrichtsqualität im Fach Biologie: Eine Video-und Interventionsstudie* [General and domain-specific charactercs of instructional quality in biology: A video and intervention-study]. Berlin: Logos Verlag Berlin GmbH.

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

Zele, E. V., & Lenaerts, J. (2004). Improving the usefulness of concept maps as a research tool for science education. *International Journal of Science Education*, *26*(9), 1043–1064.