Studies in Educational Evaluation 38 (2012) 127–134

Contents lists available at SciVerse ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/stueduc

Science teaching based on cognitive load theory: Engaged students, but cognitive deficiencies

Barbara Meissner *, Franz X. Bogner

University of Bayreuth, Chair of Biology Education, GER-95440 Bayreuth, Germany

ARTICLE INFO

Article history:
Received 14 June 2012
Received in revised form 13 September 2012
Accepted 8 October 2012

Keywords:
Engagement
Guidance
Instructional efficiency
Instructional involvement
Out-of-school
Cognitive load

ABSTRACT

To improve science learning under demanding conditions, we designed an out-of-school lesson in compliance with cognitive load theory (CLT). We extracted student clusters based on individual effectiveness, and compared instructional efficiency, mental effort, and persistence of learning. The present study analyses students' engagement. 50.0% of our sample ($n = 250$, 5th–8th graders) showed satisfying results, 11.2% were not motivated; 38.8% had difficulties to cope with the learning situation. Presumably, most of them had problems in identifying relevant contents, some were precarious about their capabilities. We suppose that those students may have improved performance with extended active support. We may advance future research on guidance in CLT, and its integration in the field of science education.

Rationale and aim

Any student-centred, interactive learning requires extended cognitive resources by definition (Hofstein & Lunetta, 2004; Winberg & Berg, 2007): Students may, for example, need to communicate and coordinate their work, and they may face unusual or even unknown activities. At an out-of-school setting, students are in addition confronted with novel surroundings distracting their attention (Orion & Hofstein, 1994). To compensate for these demanding conditions, we based the design of our interactive out-of-school lesson on the principles of cognitive load theory (CLT). We developed five experiments to examine the effects of freezing point depression, electric conductivity, endothermic solvation processes, density increase, and osmotic activity. We intended the lesson to enrich the educational programme of a commercial salt mine, and our target group comprised students of different ages (from 5th to 8th graders) and of different stratification levels. The educational objective was to provide basic knowledge about salt (NaCl) together with first experiences with laboratory equipment and hands-on learning. By orientation on CLT principles in the lesson design, we aimed to reach all students likewise, independent on their individual preknowledge.

* Corresponding author. Tel.: +49 0 921 552590; fax: +49 0 921 552696.
E-mail addresses: barbara.meissner@uni-bayreuth.de, barbara-meissner@freenet.de (B. Meissner).

0191-491X/$ – see front matter
http://dx.doi.org/10.1016/j.stueduc.2012.10.002

Background and application of cognitive load theory (CLT)

Recently, there have been many discussions about the capabilities and limitations of cognitive load theory (CLT; e.g. de Jong, 2010; Gerjets, Scheiter, & Cierniak, 2009; Moreno, 2010; Schnotz & Kürschner, 2007). However, CLT has proven to be a valuable theory of instructional design (Gerjets et al., 2009; Ozcinar, 2009; Paas, van Gog, & Sweller, 2010), investigated especially in the fields of mathematics and e-learning. The guidelines presented by CLT have been confirmed in various studies (e.g. Kalyuga, 2007; Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Paas & van Merriënboer, 1994; Sweller & Chandler, 1994; van Merriënboer, Kester, & Paas, 2006). In the following, we briefly describe CLT principles and how we applied them in the design of our lesson.

Each task requires cognitive resources of a learner's working memory which in addition depends on individual learner characteristics. Depending on the resulting cognitive load, a learner invests a certain amount of mental effort to perform the task. Mental effort "refers to the amount of capacity or resources that is actually allocated to accommodate the task demands" (Paas & van Merriënboer, 1994, p. 122). Reduced mental effort may result from previous experiences with presumably similar learning situations: If students already have conceptions about how much effort is required for a learning task they may not invest more effort than they suppose to be necessary (Paas, Tuovinen, van Merriënboer, & Darabi, 2005). These "performance-goal-oriented learners" (Paas et al., 2005, p. 28) only try to satisfy the demands specified by the teacher in order to be good at school. Another mechanism that may lead to reduced mental effort is "motivational defaults" (Clark, Howard, & Early, 2006, p. 30), that can occur if learners do not know how to cope with learning material. Learners tend either to concentrate on less demanding issues, or they opt out completely as they feel overchallenged and perceive little control over their success or failure (Clark et al., 2006).

Cognitive load of a task can result from two main causes: (a) intrinsic cognitive load (ICL) is innate to a task and depends on its difficulty and complexity; (b) extraneous cognitive load (ECL), that does not contribute to learning itself, refers to working memory capacity required to deal with the structure of a task and with the associated activities. A third component, (c) germane cognitive load (GCL), depends on ICL. GCL results from intentional learning processes and refers to the mental effort invested to deal with ICL requirements. Guidelines for instructional design of demanding tasks aim at achieving adequate levels of intrinsic, reduction of extraneous, and encouragement of germane cognitive load (Sweller, 2010; Sweller, van Merriënboer, & Paas, 1998).

(a) Intrinsic cognitive load (ICL): to reach an adequate level of ICL implies reaching an adequate level of difficulty and complexity of a task. A subject should be, for example, restructured (van Merriënboer, Schuurman, De Croock, & Paas, 2002) or simplified (Pollock, Chandler, & Sweller, 2002) according to learners' expertise, abilities, and so on. As students of our target group were novices to the subject, we limited descriptions to the effects observed at a phenomenological level.
(b) Extraneous cognitive load (ECL): to reduce cognitive load originating from extraneous factors, split attention (Sweller, Chandler, Tierney, & Cooper, 1990) and redundancy (Chandler & Sweller, 1991; Kalyuga, Ayres, Chandler, & Sweller, 2003) need to be avoided. Hence we provided step-by-step guidelines and prestructured workbook tasks, and consistently excluded redundant information.
(c) Germane cognitive load (GCL): to enhance learning processes (i.e. to foster germane cognitive load), a learner's motivation is crucial (Schnotz & Kürschner, 2007). In CLT research, tasks of high variability (contextual interference) and an appropriate level of guidance are known to foster motivation (van Merriënboer et al., 2006). Our lesson contained a high contextual interference: Students conducted similar experiments (each including salt and water/ice) covering various issues. Prestructured workbook tasks were designed using completion problems (van Gog & Paas, 2008) to provide clear guidance and to help clarify procedures and outcomes. Additionally, we offered help on demand to answer students' questions. We provided supportive information rather than ready-made solutions.

To further foster students' motivation, we employed an autonomy-supportive teaching style introduced by Reeve (2002). The theoretical background of this teaching style lies within intrinsic motivation research (e.g. Reeve, 1996; Ryan & Deci, 2000). There are three pillars that influence students' intrinsic motivation: perceived competence, autonomy, and relatedness. Perceived competence and autonomy increase if students feel adequately challenged by a task, and if they feel they are acting under their own control. The autonomy-supportive teaching style fosters these two components of perceived competence and autonomy: In their study, Reeve and Jang (2006) used a perceived self-determination and competence rating during different lessons in order to develop a list of autonomy-supportive behaviours. They concluded, for example, that teachers need to listen to their students, to foster subject-related conversation with peers, and to give encouraging feedback. An autonomy-supportive teaching style leads to increased intrinsic motivation (Deci, Nezlek, & Sheinman, 1981), more positive state emotions (Patrick, Skinner, & Connell, 1993), and better conceptual learning (Grolnick & Ryan, 1987). We applied an autonomy-supportive teaching style to foster students' motivation, and hence to increase germane cognitive load which in turn should result in increased learning success.

Students' engagement

According to Reeve, Jang, Carrell, Jeon, and Barch (2004), engagement "refers to the behavioural intensity and emotional quality of a person's active involvement during a task" (p. 147). That is, engagement includes both a behavioural and an emotional dimension. Behavioural engagement, supposed to be "most prototypical of engagement" (Skinner, Furrer, Marchand, & Kindermann, 2008, p. 778), includes actions such as, for example, investing effort, being involved, or paying attention. Emotional engagement involves state emotions, such as interest, enjoyment, or anxiety. Engagement results from intrinsically motivating learning environments. Although perceived autonomy seems to strongly predict emotional engagement, perceived competence mainly contributes to behavioural engagement. In addition, the amount of support – as estimated by teachers – influences behavioural engagement more than emotional engagement (Skinner et al., 2008).

We analysed instructional involvement to estimate students' behavioural engagement, and state emotions to estimate students' emotional engagement.

Research questions

We analysed the lesson on an individual level (cf. Goldman, 2009) and characterised students clusters according to cognitive and motivational parameters to obtain a comprehensive profile of the lesson. In a first study (Meissner & Bogner, submitted for publication), we had focused on the following questions: Does CLT provides sufficient guidelines (with respect to factual knowledge gain) for the instructional design of learning settings with predefined high cognitive load? What are the main deficiencies and how could we avoid them? In the present study, we examined our conclusions of the first study with respect to motivation. Our hypotheses were:

1. Students of clusters III and VII are motivated, i.e. show high engagement in the lesson.
2. Students of clusters I and VI show low engagement in the lesson.

Instruments and methods

After implementation of the CLT-based lesson, we characterised students on the basis of cognitive and motivational parameters to identify starting points for improvement. We used achievement data (preknowledge and learning success) to extract student clusters according to the individual effectiveness of the lesson. For each cluster, we calculated cognitive achievement (Scharfenberg, Bogner, & Klautke, 2007), mean mental effort (Paas, 1992), and instrucional efficiency of the lesson (Paas & van Merriënboer, 1993) to obtain a comprehensive cognitive profile. For motivational analyses, we chose instructional involvement (Paas et al., 2005) – a motivational construct in the framework of CLT that is based on performance and mental effort scores – as indicator of students' behavioural engagement. To refine the results, we analysed the state emotions 'interest', 'enjoyment' and 'anxiety' as indicators of emotional engagement. In Table 1, an overview of the instruments and constructs used is shown.

Table 1
Summary of instruments.

Scale/construct	Items/calculation
Knowledge test[a]	⇒Cognitive achievement[b] (CA) Combining knowledge test results... ...before and immediately after lesson → short-term CA ...before and six weeks after lesson → long-term CA = learning success
Mental effort[c]	"How would you rate the mental effort you invested in this workstation, compared to a regular science lesson?" ⇒Instructional efficiency[d] Combining cognitive achievement (long-term) and mental effort ⇒Instructional involvement[e] Combining cognitive achievement (short-term) and mental effort; points to behavioural engagement
State emotions[f]	See Table 3; point to emotional engagement

[a] See Table 2 for quality criteria.
[b] Scharfenberg et al. (2007).
[c] Paas (1992).
[d] Paas and van Merriënboer (1993).
[e] Paas et al. (2005).
[f] Laukenmann et al. (2003).

The lesson

Students participated in an out-of-school lesson in an external seminar room of a local educational centre. Teachers were offered an opportunity for a half-day school excursion. In order to exclude site effects, the programme was not performed at the salt mine itself but we chose neutral surroundings that were unrelated to the topic 'salt'. The student-centred lesson comprised the following experiments that were performed by small groups of students in optional order: (cf. Meissner & Bogner, submitted for publication; Meissner & Bogner, 2011)

"Ice-free" (freezing point depression): using binoculars, students observe that ice is melting more quickly if they put some grains of salt on it. Comparing pictures of differently treated cress, they recognise, that road salt may harm plants.

"My nervs" (electric conductivity): students are provided a simple circuit with a battery as power source, a lamp and hooter as indicator, and two electrodes. They compare electric conductivity of solid NaCl, purified water, and salt water. Additional tasks show the connection to signal transduction in neural cells.

"Salt-cold" (endothermic solvation processes): students use digital thermometers to measure temperatures of ice before and after having added salt. They calculate the difference of temperatures, and recognise a decrease.

"Shipwreck" (density increase): a short text introduces the term of density. Using a magnetic stirrer, students prepare highly concentrated brine out of salt and water. Afterwards, they put an egg into water and the sole, and recognise increased density of the sole.

"Distress at sea" (osmotic activity; demonstrated experiment): in a short film, students see that water on a cucumber slice is getting wet. An animation shows that salt detracts water from cells. Students transfer this explanation to the situation of a shipwrecked person who is drinking only sea water.

Each workstation introduces a special laboratory device (binocular, electric circuit, thermometer, magnetic stirrer, computer). Additionally, methodical and social competencies are enhanced.

Sample and implementation

The sample comprised 113 boys (45.2%) and 137 girls (54.8%). Students' age ranged between 10 and 15 years (M_{age} = 11.6, SD = 1.6). In accordance with local curricula and the target group of the salt mine, we included students of high (10–12 years old) and low (10–15 years old) stratification level. Students were part of the sample of the first study that comprised 276 students (M_{age} = 11.6, SD = 1.6; 46.4% male, 53.6% female). 26 (i.e. 9.4%) incomplete questionnaires were dropped from further analyses.

Knowledge test

We applied an ad hoc multiple-choice test to quantify students' preknowledge and knowledge gain. Students completed the knowledge tests one week before (KT1), immediately after (KT2), and six weeks after the out-of-school lesson (KT3). The knowledge test comprised 13 multiple-choice items. We asked students about the effects observed, the devices used, and about applications of the effects highlighted in workbook tasks. The knowledge test was pilot-tested with 109 students of the 5th grade (high stratification level) who filled in the test one week before and immediately after the lesson (Cronbach's alpha = .72). Knowledge test characteristics and examples of knowledge test items are listed in Table 2.

Preknowledge, cognitive achievement and learning success

The sum-score of correctly answered questions in KT1 quantified the amount of each student's preknowledge. We differentiated learning outcome in short-term cognitive achievement (based on KT2 sum-scores) and long-term cognitive achievement (learning success; based on KT3 sum-scores). For calculation, we applied a weighted difference introduced by Scharfenberg et al. (2007), which takes into account ceiling-effects resulting from the limited number of questions. Hence we calculated cognitive achievement as the difference between the sum-scores of KT2 (KT3, resp.) and KT1, multiplied by the quantifier 'sum-score of KT2 (KT3, resp.)/total number of knowledge test items'. Short-term cognitive achievement, for instance, is defined as follows: $(KT2 - KT1) \times KT2/13$.

As it is changes in long-term memory that characterise learning we employed long-term cognitive achievement as an index of students' learning success.

Cluster development

We decided to perform cluster analyses, as our aim was to find the structures behind the data set, and to get a profile of a CLT based lesson. In the first study, we developed student clusters

Table 2
Core data of the knowledge test as results of a pilot-testing with 109 5th graders, and examples of knowledge test items.

Reliability[a]	Difficulty index[b]	Discrimination index[c]
.72	22–84%	.36
Examples of knowledge test questions		
Which of these conducts electricity the best? Pure salt/pure water/rock salt/saltwater [correct] What are binoculars used for? To: see things amplified [correct]/dissolve substances/measure indoor and outdoor temperature simultaneously/gauge objects exactly		

[a] Cronbach's alpha.
[b] number of correct answers per question.
[c] corrected item-total correlation.

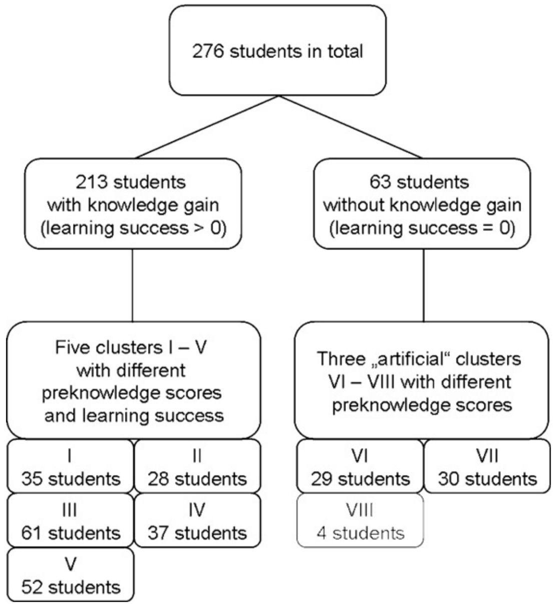

Fig. 1. Overview of the first study cluster formation.

based on students' preknowledge and learning success, as illustrated in Fig. 1.

First, we divided the students into two groups: (a) students with knowledge gain (learning success > 0) and (b) students without knowledge gain (learning success = 0). Second, we conducted hierarchical and cluster-centres analyses with group (a). We calculated contingency c and corrected contingency $c_{corr} = c/c_{max}$ (c_{max} = sqrt($(n-1)/n$); n = number of clusters) to compare solutions of hierarchical and cluster-centres analysis. We found high contingency (c = .84; c_{corr} = .94) between the five-cluster solutions of both methods and, hence, chose the extraction of five clusters as the final solution. We labelled preknowledge and learning success of the clusters according to the whole sample's quartiles 'low', 'medium', or 'high' (cf. Table 4). Third, we manually assigned students without knowledge gain, group (b), to three clusters with low, medium, and high preknowledge. As the latter comprised only four students, we excluded it from further analyses.

To ensure compliance of our sample with the former sample of the first study, we partially repeated cluster analyses: We conducted k-means cluster-centres analysis, assuming a five-cluster solution. We calculated contingency c and corrected contingency c_{corr} to estimate cluster solution comparability of the two samples. Comparison showed high contingency (c = .92, c_{corr} = .99) between the two solutions.

Mental effort and instructional efficiency

During the lesson, students rated the mental effort invested at each workstation on an unidimensional self-rating scale (Paas, 1992; cf. Table 1) ranging from 1 (*very low*) to 7 (*very high*). We used mean scores of the five ratings (Cronbach's alpha = .67) as an estimate of average mental effort (ME) throughout the lesson, and for further calculations of instructional efficiency and instructional involvement. Although self-rating scales have been criticised, their validity has been supported by psychophysical theory (Stevens, 1975; Gopher & Braune, 1984).

To estimate the efficiency of the lesson (cf. van Gog & Paas, 2008) we used the method introduced by Paas and van Merriënboer (1993) whereby instructional efficiency (IE) scores are computed as the difference between z-standardised long-term cognitive achievement (A_z) and mental effort scores (ME_z), divided by square root of 2:

$$\text{Instructional efficiency (IE)} = \frac{A_z - ME_z}{\text{sqrt}(2)}$$

Identical z-scores of A_z and ME_z lead to an IE score of zero that is defined as average IE. Paas and van Merriënboer explained the calculation graphically (cf. Fig. 2): With ME_z as abscissa and A_z as ordinate, a line with slope 1 through the origin displays average IE scores. Above- and below-average scores result from the rectangular distance of a point from this line.

We used long-term cognitive achievement scores, students' learning success, for the calculation of IE as we considered a lesson to be the more efficient the greater the amount learned persistently is. As IE scores were normally distributed (Kolmogorov–Smirnov with Lilliefors correction: p = .07) we used t-tests to compare the results of each cluster with residual sample mean scores. Bonferroni correction resulted in a significance level of p = .007.

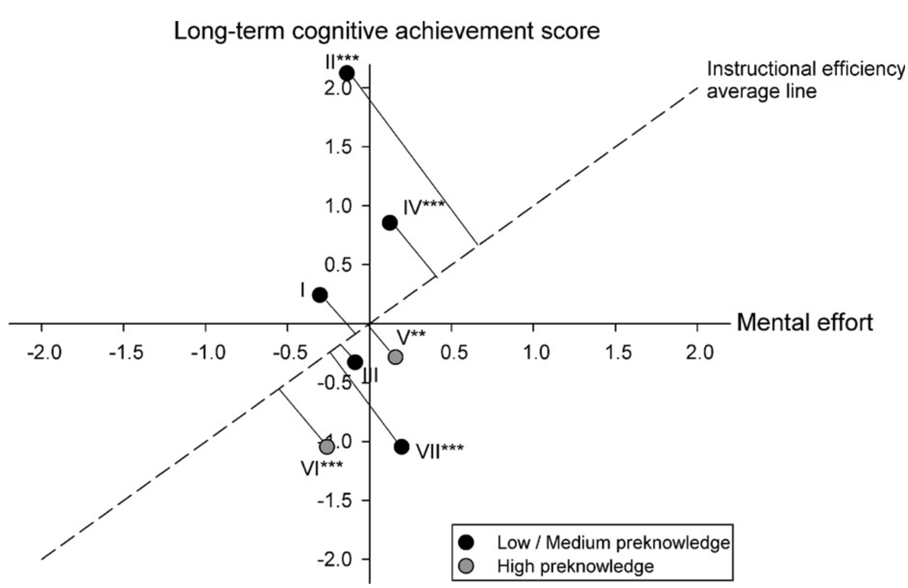

Fig. 2. Instructional efficiency (IE) scores of each cluster, shown as the distance from the mean IE line (dashed). We used t-tests to examine statistical significance. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Students' engagement in the lesson

Engagement comprises the two major components of behavioural and emotional engagement (Skinner et al., 2008). Behavioural engagement means behaviours like investing effort, emotional engagement means state emotions like interest throughout the lesson. We used the construct instructional involvement (Paas et al., 2005) and the state emotions scale (Laukenmann et al., 2003; Randler et al., 2011) to measure behavioural and emotional engagement, respectively.

Behavioural engagement – calculation of instructional involvement

Behaviourally engaged students are supposed to participate in a lesson and to try to do well (Skinner et al., 2008). Mental effort and performance reflect a student's engagement. As a combination of these two factors, Paas et al. (2005) offered the construct of instructional involvement as an alternative to common self-rating scales. To calculate instructional involvement scores, the sum of z-standardised cognitive achievement and mental effort scores is divided by square root of 2:

$$\text{Instructional involvement (InsInv)} = \frac{A_z + ME_z}{\text{sqrt}(2)}$$

Similarly to instructional efficiency, InsInv scores can be displayed in a coordinate system (cf. Fig. 3) with ME_z as abscissa and A_z as ordinate. A line with slope -1 through the origin represents average InsInv scores of 0.

As involvement is a short-term, situation-specific variable, we used short-term cognitive achievement scores for the calculation of InsInv. InsInv scores were normally distributed (Kolmogorov–Smirnov with Lilliefors correction: $p = .20$). Thus, we used t-tests to compare the results of each cluster with residual sample mean scores. Bonferroni correction resulted in a significance level of $p = .007$.

Emotional engagement – state emotions questionnaire

To gain insight in students' emotional engagement (Skinner et al., 2008) during the lesson we analysed the subscales 'interest', 'well-being', and 'anxiety' of the state-emotions questionnaire (Laukenmann et al., 2003; Randler et al., 2011), using a four-point Likert scale (1 = *not at all right* to 4 = *completely right*). The students filled in the questionnaire immediately after the lesson, together with KT2. As well-being contains two items measuring joy and two items measuring satisfaction (cf. Table 3), we suggest the term 'enjoyment' for this subscale.

Table 3
Subscales of the state-emotions questionnaire.

Subscale	Examples of Items
Interest	
Cognitive interest	The lesson was interesting for me
Value	I found that topic important
Enjoyment[a]	
Joy	I enjoyed the lesson
Satisfaction	For me it was a good lesson
Anxiety	
Worry	The lesson frightened me
Emotionality	Several events alarmed me

[a] Originally 'Well-being'.

As data were not normally distributed we applied the nonparametric Mann–Whitney U test for between-group comparisons. As cluster II showed exemplary results we set this cluster as reference for optimally engaged students with optimal cognitive achievement. Bonferroni correction resulted in a significance level of $p = .02$ for each subscale.

Results

Cluster description, learning success, and instructional efficiency (IE)

We obtained the seven clusters described in Table 4. The two clusters with rather low preknowledge (clusters I and II) showed medium or high learning success. We found three clusters with medium preknowledge scores (clusters III, IV, and VII). They comprised students with learning success from none (cluster VII) to high (cluster IV). Clusters V and VI with high preknowledge scores yielded low to medium learning success and no learning success, respectively.

Results of IE calculations are shown in Fig. 2. Clusters II and IV had significantly above-average IE (II: $t = -9.4$, $df = 248$, $p < .001$; IV $t = -3.5$, $df = 248$, $p = <.001$). IE was average for students of

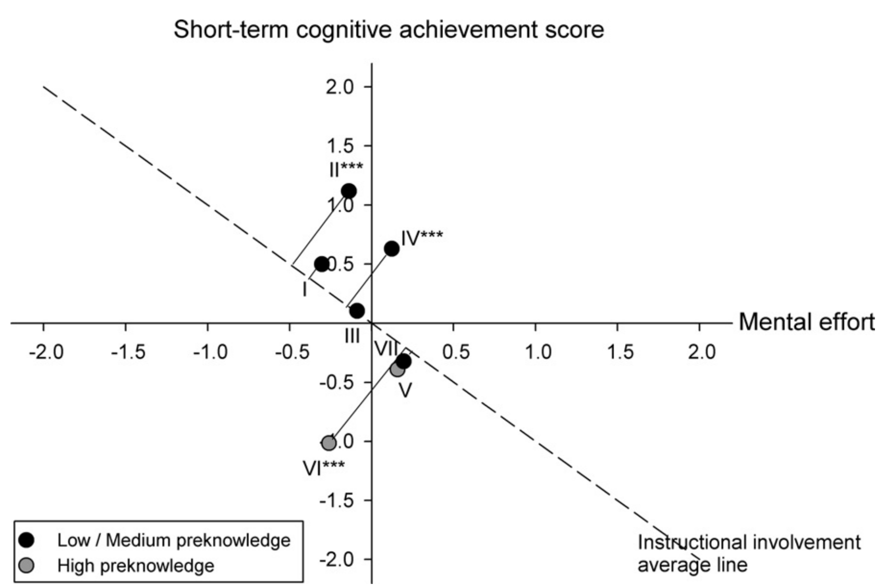

Fig. 3. Instructional involvement (InsInv) scores of each cluster, shown as the distance from the mean InsInv line (dashed). We used t-tests to examine statistical significance. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 4
Cluster description.

	I[a]	II	III[a]	IV	V	VI	VII[a]
n	24	25	44	42	58	28	29
Pre-knowledge	Low	Low/medium	Medium	Medium	High	High	Medium
Learning success	Medium	High	low/medium	Medium/high	Low/medium	None	None

[a] Clusters I, III, and VII are of major interest in the discussion.

clusters I and III, and significantly below-average for students of clusters V, VI, and VII (V: $t = 2.8$, $df = 248$, $p = .005$; VI: $t = 5.2$, $df = 58$, $p < .001$; VII: $t = 5.2$, $df = 248$, $p < .001$).

As can be seen in Figs. 2 and 3, ME of clusters IV, V, and VII was slightly above-average (i.e. positive ME scores; *n.s.*), and ME of clusters I, II, III, and VI was slightly below-average (i.e. negative ME scores; *n.s.*).

Instructional involvement

The majority of the clusters, namely clusters I, III, V, and VII, achieved average instructional involvement scores (Fig. 3). Only instructional involvement of cluster VI was significantly below-average ($t = 7.4$, $df = 44$, $p < .001$), while instructional involvement of both clusters II and IV were significantly above-average (II: $t = -3.6$, $df = 248$, $p < .001$; IV: $t = -4.4$, $df = 73$, $p < .001$).

State emotions

Students of the whole sample showed high interest and enjoyment scores (interest: Q1 – 2.75, Q2 – 3.25, Q3 – 3.75; enjoyment: Q1 – 3.00, Q2 – 3.75, Q3 – 4.00; Q = quartile), and low anxiety scores (Q1 – 1.00, Q2 – 1.25, Q3 – 1.75).

We were mainly interested in the emotional engagement of students of clusters I, III, and VII with similar instructional involvement scores. We defined cluster II as reference as students of this cluster revealed highly desirable results. Nevertheless, there were no significant differences in students' interest between clusters II and clusters I, III, and VII, respectively. Students of cluster VII had significantly lower enjoyment-scores than students of cluster II ($Z = -2.63$, $p = .009$). Clusters I, III, and VII each revealed significantly higher anxiety-scores, compared to cluster II (I: $Z = -2.69$, $p = .007$; III: $Z = -2.90$, $p = .004$; VII: $Z = -3.37$, $p = .001$). Fig. 4 illustrates these results.

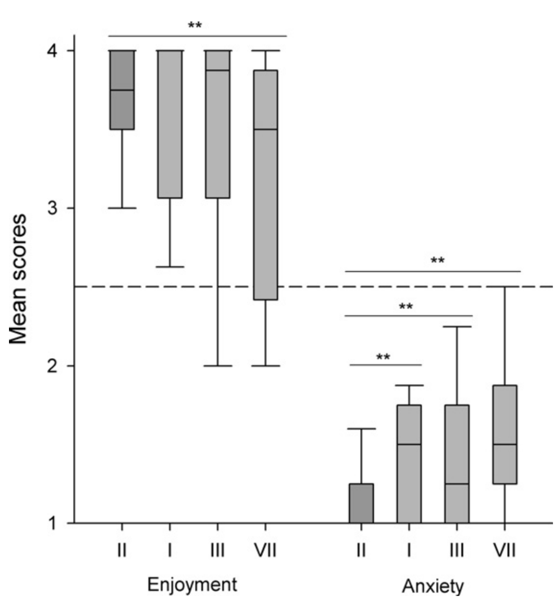

Fig. 4. Enjoyment and anxiety scores of cluster II (as a reference sample), and clusters I, III, and VII; dashed line indicates midpoint of the scale; **.001 ≤ $p < .01$ (Mann–Whitney *U* test).

Discussion

As a first step to value CLT as a guideline for heuristic science education, we characterised students of our sample to gain insight into cognitive and affective outcomes of a CLT-based lesson.

Clusters II, IV, and V (satisfactory learning success)

50.0% of the sample, students of clusters II, IV, and V, with different levels of preknowledge from low to high, showed very satisfactory learning success (cf. Table 5). For the case of students of cluster V with high preknowledge and only low to medium learning success it must be considered that – due to the already high scores in the pretest one week before the lesson – students could not have answered many more questions correctly in the knowledge tests after the lesson. In this case, then, low to medium learning success can be assumed a good result.

Cluster VI (non-satisfactory learning success and low behavioural engagement)

Students of cluster VI, 11.2% of the whole sample, had both below-average IE and below-average instructional involvement. Seemingly, neither was the lesson efficient for these students, nor did they engage in the lesson. The low engagement confirms our first assumption that an expertise-reversal effect (Kalyuga et al., 2003; Schnotz, 2010; Sweller, 2010) may have caused these deficiencies: As their preknowledge was high, students may have been exposed to a high amount of redundant information so that they may not have recognised relevant information. Thus, the lesson may have appeared to be uninteresting for the students, which resulted in their opting out.

Clusters I, III, and VII (non-satisfactory learning success, but high engagement)

Behavioural engagement in clusters I, III, and VII

Students of clusters I and III, 27.2% of the sample, showed about average instructional efficiency (IE; cf. Fig. 2). That is, related to the

Table 5
Overview of results of clusters I, III, and VII, that presumably required more guidance during the lesson.

Cluster	*n*	IE[a]	InsInv[a]	Interest[b]	Enjoyment[b]	Anxiety[b]
I	24	*n.s.*	*n.s.*	*n.s.*	*n.s.*	+**
III	44	*n.s.*	*n.s.*	*n.s.*	*n.s.*	+**
VII	29	−***	*n.s.*	*n.s.*	−**	+**

Note: +: above, −: below reference score.
** $p < .01$.
*** $p < .001$.
[a] Deviations with reference to residual sample mean score.
[b] Deviations with reference to scores of 'prototype-cluster' II.

mental effort invested, cognitive achievement was adequate. In our first study, as learning success of cluster I reached only medium levels, we assumed that students could have performed better if they had invested more mental effort. We considered the two possibilities that students either did not know how to or did not want to invest more mental effort. As instructional involvement, that quantifies behavioural engagement, was about average for both clusters I and III (cf. Fig. 3) we may assume that students could have improved their performance if they had known how to. If students had not been motivated to invest more mental effort, instructional involvement scores would have been lower.

We can assume a similar explanation for students of cluster VII, 11.6% of the students, with medium preknowledge scores: IE was significantly below-average (i.e. related to the mental effort invested, cognitive achievement was too low), but instructional involvement was about average, indicating that students were engaged in the lesson. We can therefore confirm students' difficulties to deal with the learning material as an explanation of the poor performance, as supposed in the first study.

Emotional engagement in clusters I, III, and VII

As instructional involvement scores of students of clusters I, III and VII were similar, we assume the same cause for their deficiencies. However, IE scores of these three clusters were quite different (cf. Table 5): While IE scores of clusters I and III were slightly above-average and roughly average, cluster VII yielded below-average IE scores. Apparently the reasons of these students' difficulties with the learning setting may differ. We analysed the state emotions interest, enjoyment, and anxiety as indicators of emotional engagement to describe clusters I, III, and VII in more detail (cf. Table 5).

Both clusters I and III showed no significant differences to the 'reference-cluster' II in their interest and enjoyment, but did show increased anxiety. We can therefore assume that perceived familiarity (Paas et al., 2005) with required demands did not account for the reduced mental effort, otherwise we would have expected no increased anxiety scores. Hence, students rather might have switched their attention to less demanding issues as they did not know how to deal with the learning material (Clark et al., 2006). Students may hence benefit from more guidance in terms of active support by the teacher: Active communication with students, and asking specific questions about working steps and learning outcomes of the workstations may help to integrate active learning processes – that is, germane cognitive load – in students' performance of the workstations.

Students of cluster VII, compared to their learning success, showed relatively high mental effort, which resulted in below-average IE. Thus, mental effort was not reduced, but inappropriately high. In this case, worry – one aspect of anxiety – may have caused an "interference effect [...] on attention" (Eysenck & Calvo, 1992, p. 410): Worry may have reduced working memory capacity on the one hand, on the other hand students may nevertheless have tried to cope with task demands by investing additional mental effort (cf. Eysenck & Calvo, 1992). The importance of worry is confirmed by the increased anxiety and the decreased enjoyment of students of this cluster (cf. Table 5). Students may have been overchallenged by the lesson and may have needed clear verbal instructions on what to do, and continuous encouraging feedback. Hence, these students may benefit from more guidance in terms of active support by the teacher, as well.

Summary

In summary, we can say that 50.0% (clusters II, IV, and V) of the students showed good results, 11.2% (cluster VII) were not motivated, and 38.8% (clusters I, III, and VII) may have had difficulties in coping with the learning situation. Based on emotional engagement analysis, we identified two reasons why students of clusters I, III, and VII showed deficiencies although they were engaged in the lessons: Students of clusters I and III may not have succeeded in holding to the course set as they mainly concentrated on hands-on activities. Students of cluster VII may have been worried about failures, resulting in decreased working memory capacity for cognitive processes. Both of these explanations let assume that students may have required more guidance in terms of active support by the teacher during the lesson.

Outline

In research on school science laboratories, the aspect of adequate guidance to students during a lesson has been neglected (Lunetta, Hofstein, & Clough, 2007), as is the case in CLT research beyond the field of e-learning. Cognitive tutors (e.g. Koedinger & Aleven, 2007) or software agents (e.g. Moreno, 2004) as a part of virtual learning environments have proved effective, but beyond the field of e-learning, little concrete statements about assembly of adequate guidance have developed to date. Further specification of guidance in the field of science education is needed and may be subject of future research on CLT in science education. Kirschner, Sweller, and Clark (2006) have – among others (e.g. Klahr & Nigam, 2004; Moreno, 2004) – already demonstrated the importance of such adequately developed guidance that leads learners to identify relevant information and to achieve educational objectives. Goldman (2009) pointed to the importance of individual learner characteristics. As a next step, we would specify the students' needs for guidance in the framework of CLT more precisely, specifically for the field of science education. Possibilities may be a stronger emphasis on self-explanations (Renkl, Hilbert, & Schworm, 2009; Sweller, 2010), or students' training on how to discuss in group and how to build arguments (Gillies & Haynes, 2010).

Limitations

One limitation of the study may be that we did only a knowledge test as a pretest. We did not gather any data about, for example, students' attitudes or motivation before the lesson. However, the major limitation is that we had no control group with a variation of the amount of guidance as we did not expect this to be of interest before we had analysed our data. Otherwise, we already could have confirmed our conclusions of the present study.

References

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(3), 293–332.

Clark, R., Howard, K., & Early, S. (2006). Motivational challenges experienced in highly complex learning environments. In J. Elen & R. Clark (Eds.), *Handling complexity in learning environments: Theory and research* (pp. 27–42). Amsterdam: Elsevier.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science, 38*(2), 105–134.

Deci, E., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology, 40*, 1–10.

Eysenck, M., & Calvo, M. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6*(6), 409–434.

Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Psychology Review, 21*, 43–54.

Gillies, R., & Haynes, M. (2010). Increasing explanatory behaviour, problem-solving, and reasoning within classes using cooperative group work. *Instructional Science* http://dx.doi.org/10.1007/s11251-010-9130-9 *(Advance online publication).*

Goldman, S. (2009). Explorations of relationships among learners, tasks, and learning. *Learning and Instruction, 19*, 451–454.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Factors, 26*(5), 519–532.

Grolnick, W., & Ryan, R. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology, 52*, 890–898.

Hofstein, A., & Lunetta, V. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*, 28–54.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Reviews, 19*, 509–539.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*(3), 579–588.

Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiental, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction – Effects of direct instruction and discovery learning. *Psychological Science, 15*, 661–667.

Koedinger, K., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*, 239–264.

Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, P., & von Rhöneck, C. (2003). An investigation of the influence of emotional factors on learning in physics instruction. *International Journal of Science Education, 25*(4), 489–507.

Lunetta, V., Hofstein, A., & Clough, M. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. In S. Abell & N. Ledermann (Eds.), *Handbook of research on science education.* (pp. 393–442). Mahaw, NJ: Lawrence Erlbaum Associates.

Meissner, B., & Bogner, F. (2011). Enriching Students' Education Using Interactive Workstations at a Salt Mine Turned Science Center. *Journal of Chemical Education, 88*(4), 510–515.

Meissner, B., & Bogner, F. Towards Cognitive Load Theory as Guideline for Instructional Design in Science Education. Manuscript submitted for publication.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*, 99–113.

Moreno, R. (2010). Cognitive load theory: More food for thought. *Instructional Science, 38*(2), 135–141.

Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific field trip in a natural environment. *Journal of Research in Science Teaching, 31*(10), 1097–1119.

Ozcinar, Z. (2009). The topic of instructional design in research journals: A citation analysis for the years 1980–2008. *Australasian Journal of Educational Technology, 25*, 559–580.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434.

Paas, F., Tuovinen, J., van Merriënboer, J., & Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development, 53*(3), 25–34.

Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review, 22*, 115–121.

Paas, F., & van Merriënboer, J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737–743.

Paas, F., & van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122–133.

Patrick, B., Skinner, E., & Connell, J. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology, 65*, 781–791.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61–86.

Randler, C., Hummel, E., Gläser-Zikuda, M., Vollmer, P., Bogner, F., & Mayring, C. (2011). Reliability and validation of a short scale to measure situational emotions in science education. *International Journal of Environmental & Science Education, 6*(4), 359–370.

Reeve, J. (1996). *Motivating others – Nurturing inner motivational resources*. Needham Heights, MA: Allyn & Bacon. pp. 19–38, 170–175 and 201–216.

Reeve, J. (2002). Self-determination theory applied to educational settings. In E. Deci & R. Ryan (Eds.), *Handbook of self-determination research* (pp. 183–204). Rochester, NY: University of Rochester Press.

Reeve, J., & Jang, H. (2006). What teachers say and do to support students autonomy during a learning activity. *Journal of Educational Psychology, 98*(1), 209–218.

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion, 28*(2), 147–169.

Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review, 21*, 67–78.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54–67.

Scharfenberg, F.-J., Bogner, F., & Klautke, S. (2007). Learning in a gene technology laboratory with educational focus – results of a teaching unit with authentic experiments. *Biochemistry and Molecular Biology Education, 35*(1), 28–39.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review, 19*, 469–508.

Schnotz, W. (2010). Reanalyzing the expertise reversal effect. *Instructional Science, 38*(3), 315–323.

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781.

Stevens, G. (Ed.). (1975). *Psychophysics – Introduction to its perceptual, neural, and social prospects*. New York: John Wiley & Sons.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*, 123–138.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185–233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119*, 176–192.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*(1), 16–26.

van Merriënboer, J., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20*, 343–352.

van Merriënboer, J., Schuurman, J., De Croock, M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*, 11–37.

Winberg, T., & Berg, C. (2007). Students' cognitive focus during a chemistry laboratory exercise: Effects of a computer-simulated prelab. *Journal of Research in Science Teaching, 44*(8), 1108–1133.