TU Delft

Delft University of Technology

Explicit Teaching and Scaffolding to Enhance Concept Learning by Design Challenges

van Breukelen, Dave; Smeets, M; de Vries, Marc

Publication date
2015

Document Version
Final published version

Published in
Journal of Research in STEM Education

Citation (APA)
van Breukelen, D., Smeets, M., & de Vries, M. (2015). Explicit Teaching and Scaffolding to Enhance Concept Learning by Design Challenges. *Journal of Research in STEM Education*, *1*(2), 87-105.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Journal of Research in STEM Education
ISSN:2149-8504 (online)
Vol 1, No 2, December 2015, PP 87-105

RESEARCH REPORT

Explicit Teaching and Scaffolding to Enhance Concept Learning by Design Challenges

Dave van Breukelen [a 1], Maurice Smeets[a], Marc de Vries[b]

[a]*Fontys University of Applied Sciences for Teacher Education Sittard, Sittard, The Netherlands*

[b]*Delft University of Technology, Science Education and Communication, Delft, The Netherlands*

Abstract: *This paper presents a mixed methods study, carried out among 21 first-year student teachers, that investigated learning outcomes of a modified Learning by Design (LBD) task. The study is part of a series of studies that aims to improve learning, teaching and teacher training. Design-based science challenges are reasonably successful project-based approaches for breaking down the boundaries between traditional STEM subjects. Previous learning outcomes of the extensively studied LBD approach demonstrated a strong positive effect on students' skills. However, compared to traditional classroom settings, LBD provided little profit on (scientific) concept learning. For this, according to two preliminary studies, a lack of explicit teaching and scaffolding strategies, both strongly teacher-dependent, bears a share of responsibility. The results of the third study discussed in this paper indicate that emphasizing these strategies strengthens concept learning without reducing positive effects on skill performances.*

Keywords: Learning by Design, science, technology, skills, concepts, learning outcomes

The world around us is constantly changing and getting more complex. Partly because science and technology have grown progressively denser in our personal lives, where most of the world's issues ask for an interdisciplinary approach to meet humans' needs. We might expect that school systems respond accordingly by delivering juveniles ready to face these issues. However, many curricula are traditionally dominated by separate disciplines (Scott, 2008) where international studies, e.g. ROSE (Sjöberg & Schreiner, 2010), demonstrate a decreasing interest in and understanding of science and technology. Aikenhead (2006) states that unidisciplinary science curricula result in sterile, dehumanized science content that has little appeal to students and is often perceived by them to be irrelevant. Several studies indicate that a holistic understanding of science and technology, through interdisciplinary teaching, may improve students' motivation and understanding (Lustig et al., 2009; Osborne & Dillon, 2008; Rennie, Venville, & Wallace, 2012). If we want students to learn how to apply knowledge and skills in daily life, their education experiences must involve them in learning and applying knowledge and skills of related disciplines in recognizable contexts (Bybee, 2013). Therefore, many national governments aim for interdisciplinary science, technology, engineering, and mathematics (STEM) education (National Science and Technology Council, 2013; Office of the Chief Scientist, 2013; Parliamentary Office of Science & Technology, 2013). In this context technology should be seen as purposeful and goal-directed activities where knowledge (e.g. conceptual, procedural) and skills (e.g. design, experimentation, craft) are used to solve practical problems and meet needs (International Technology Education Association, 2007).

A quite successful approach to respond to the strong interplay of STEM subjects, called Learning by Design (LBD; Kolodner, 2002b), has been studied extensively from 1999 until 2003. Those studies showed high

[1] Corresponding author. Fontys University of Applied Sciences for Teacher Education Sittard, Sittard, The Nethertlands,
E-mail: d.vanbreukelen@fontys.nl
van Breukelen, D., Smeets, M., & de Vries, M. (2015). Explicit teaching and scaffolding to enhance concept learning by design challenges. *Journal of Research in STEM Education*, *1*(2), 87-105.

© i-STEM 2015, j-stem.net

student-involvement and, compared to non-LBD classes, significantly better collaboration, metacognitive and science skills. However, concept learning lagged behind (Kolodner, 2002; Kolodner, Camp, et al., 2003; Kolodner, Gray, & Fasse, 2003) what ties in with the call to educate students with a more conceptual understanding of design technology (International Technology Education Association, 2007).

Prior to this study, Study 1 (Van Breukelen, De Vries, & Schure, 2016) and Study 2 (Van Breukelen, Van Meel, & De Vries, 2016) took place that studied LBD's conceptual learning process in detail from a student's and teacher's perspective. By unravelling the learning process it became clear to what extent students used and learned scientific concepts during design activities and how much room for improvement was left. The preliminary studies showed that students were able to manage a conceptual learning gain comparable to average learning gains of many (traditional) physics-related courses (Hake, 1998). In that way, the findings of Kolodner, Gray, et al. (2003) are confirmed. Although students learned science at an apparently acceptable level more progress should be possible because LBD, compared to traditional educational settings, provides a sound theoretical basis for a higher level of concept learning, which will be discussed later on. The preliminary studies revealed two (interrelated) causes that prevented concept learning from reaching a (much) higher level. First, the complexity and extendedness of design challenges obscured scientific content (What to learn?) and forced students to become product- and process-focused (What to do and deliver?). This resulted in the use and learning of loose facts with too little coherence. Second, explication of underlying science had too little attention during task construction and teacher intervention, resulting in addressing an incomplete and disguised framework of conceptual knowledge. This study aims to address these problems by suggesting improvements, based on literature on learning sciences, concerning explicit teaching and scaffolding strategies, which will be discussed in detail later on. The central research questions are: To what extent will application of explicit teaching and scaffolding strategies result in learning a more comprehensive, coherent and explicit framework of scientific knowledge, and how this affects the conceptual learning gains?; Does skill performance, despite the interventions, still increase and how it is (cor)related with concept learning? Finally, a comparison takes place with previous studies to discover why and how the improvements affect (concept) learning.

LBD and desing composition of the challenge used for this study

LBD contains, as shown in Figure 1, two essential components for learning skills, practices and content: (re)design and investigation. Within these components are a variety of reflective hands-on and heads-on activities concerning design technology, science practices, public presentation, collaboration and teacher-guided class discussions. Students (operating in design groups) are faced with a design challenge where they first have to explore and establish things they need to know/learn for succeeding. By information seeking and experimentation they find answers to questions raised in order to apply them in the design

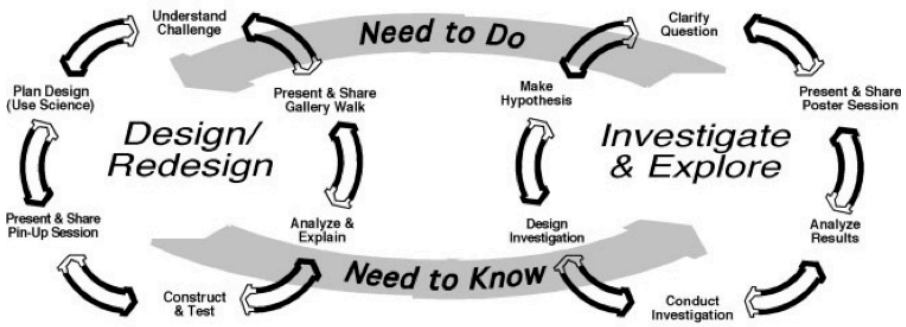

Figure 1. *Learning by Design's Cycles. Reprinted from Kolodner (2002b, p. 339).*

© i-STEM 2015, j-stem.net

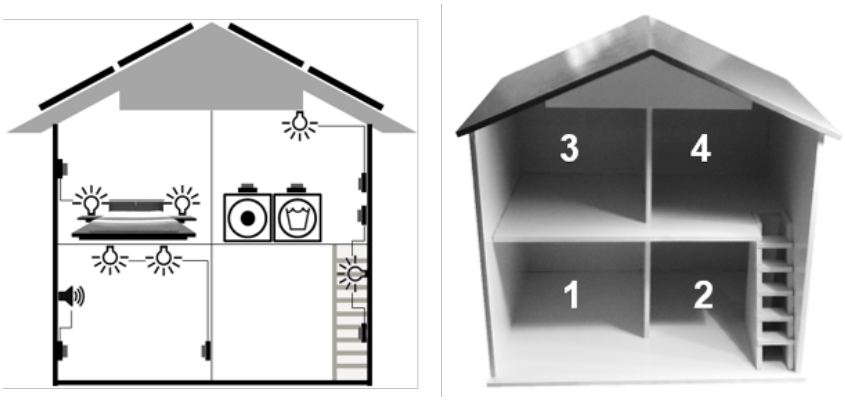

Figure 2. *Model house and layout*

Investigation of this application may lead to additional questions and reinvestigation. To incentivize the understanding of design-related principles and concepts, teacher-guided activities take place (poster and pin-up sessions, whole-class discussions and gallery walks). During these activities experiences and insights are shared among groups, feedback is being given and science is being made explicit. In general, LBD provides a constructivist learning environment where students experience the necessity to learn (Kolodner, Hmelo, & Narayanan, 1996) driven by the fact that students' pre-task conceptions are not sufficient for succeeding: design challenges deliberately address cognitive conflicts. A more scientific framework of knowledge is necessary to cope conflicts and reach conceptual change (Abdul Gafoor & Akhilesh, 2013; Cobern, 1994). Based on literature, e.g. Brandsford, Brown, Donovan, and Pellegrino (2003), LBD contains several elements that are beneficial to conceptual change: collaboration, reflection, contextual learning, applying what is learned, learning from failures and iteration, and connecting skills, practices and concepts.

For this study, an existing LBD challenge was modified for better concept learning. The challenge originated from Study 2 that also concerned first-year student teachers (science). Design groups (3 students per group, randomly chosen) were challenged to design a solar power system for a model house, shown in Figure 2, by taking into account the design specifications in Table 1 that headed for using underlying science, decision-making and creative thinking. Regarding these specifications and the scientific objectives in Table 2 the most fundamental design principles concerned proper wiring (combining series and parallel parts) and regulating current, voltage and resistance for maximum efficiency. Table 3 shows how the LBD elements in Figure 1 were applied resulting in an activity, guided through an instructive presentation and a student's and teacher's guide, that took six periods of 90 minutes.

Table 1

Design specifications and components

Design specifications			
A. [Room layout] Room 1: 2 lamps operated by 1 switch and a doorbell (SP) operated by 1 switch. Room 2: 1 lamp operated by a set of 2-way switches (staircase wiring). Room 3: 2 lamps operated by 1 switch. Room 4: 1 lamp operated by 1 switch, 1 washing machine (M2) with adjustable speed operated by 1 switch, 1 dryer (M1) operated by 1 switch.			
B. [Solar powering] The entire lightning has to be connected to a separate (combination of) solar cell(s). The same applies to the doorbell and washer-dryer combination.			
C. [Efficiency] The energy efficiency of the entire wiring has to be as high as possible and, in addition, the use of materials as less as possible. In any case, it is not allowed to use more components than available.			
Component	Quantity	Component	Quantity
Motor 1.5V DC (M1)	1	Solar cell 4V/35mA	4
Motor 3.0V DC (M2)	1	Solar cell 5V/81mA	4
Mini-speaker 800mW (SP)	1	Solar cell 0.5V/400mA	4
Set of LEDs, resistors, wires and switches	1	Solar cell 0.5V/800mA	4

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016).

Table 2

Scientific objectives and interrelatedness with the challenge

DC electric circuits objectives	Interrelatedness
1. Physical aspects of electric circuits: resistance is a property of an object and hinders current flow (Ohm's Law); equivalent resistance in series increases and in parallel decreases as more elements are added; the necessity of a closed circuit to enable current flow; interpret pictures, diagrams, symbols of a variety of circuits.	• Resistors are necessary to reduce current flow and a variable resistor is necessary to adjust the washing machine's speed. Furthermore, students have to interpret and design a variety of circuit parts in order to meet the requested wiring.
2. Energy and power: apply the concepts of energy (dissipation, conversion and conservation) and power (work done per unit time) to a variety of circuits.	• Students have to establish the amount of energy supply and consumption by the designed circuit in order to reach maximum efficiency.
3. Current: understand and apply conservation of current (Kirchhoff's point rule) to a variety of circuits; explaining the behavior of an ideal current source.	• Combining series and parallel parts (solar cells and components) to meet design specification forces students to investigate and calculate current flow and potential differences. Furthermore, students have to investigate the behavior of (combined) solar cells to get informed about differences compared to (well-known) voltages sources.
4. Potential difference, voltage: the amount of current is influenced by potential difference; apply the concept of Kirchhoff's loop rule (åV = 0 around a closed loop); explaining the behavior of an ideal voltage source.	

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016).

© i-STEM 2015, j-stem.net

Table 3

Stages and activities

Stages [time]	Activities[a]	Final products[b]
Introducing the Challenge and Context [15-20 min]	Introduction of context, design challenge, activities, organization, learning sources, time schedules, materials, objectives, etc.	
Understanding the Challenge, Messing About, White boarding [50-60 min]	Exploration of the challenge, context and objectives (G) Writing down ideas, (research) questions and hypotheses (G): what to do and learn? White boarding: sharing results; feedback session (C)	Design diary stage 2 Flip chart for white boarding (G)
Investigate & Explore, Poster Session [120-180 min]	Formulate and distribute (scientific) research questions (C) Discussion "fair test rules of thumb" (C) Design and conduct experiments, collect data, conclude (G) Presentation of results: poster session; feedback session (C) Discussion about results and fair testing: redoing/adjustments (C/G)	Design diary stage 3 Final research questions (C) Fair test rules of thumb (C) Laboratory notebook (G) Experiment poster (G)
Establishing Design Rules of Thumb [20-30 min]	Determination of design rules using experiment results (C) Focus on the science content: science vocabulary and concepts (C)	Design diary stage 4 Design rules of thumb (C)
Design Planning, Pin-Up Session [80-90 min]	Devise, share and discuss design solutions: divergent thinking (G) Poster: provisional design solution (G) Pin-up session (posters): feedback session (C) Adjusting and redoing until satisfied: final design solution (C/G)	Design diary stage 5 Design posters (G) Design sketch (G)
Construct & Test, Analyze & Explain, Gallery Walk [120-180 min]	Prototyping: realization of the design solution (G) Testing the design: realization of design specifications (G) Gallery walk: determine shortcomings; feedback/reflection (C) Adjustments of the design rules and design solutions (C/G)	Design diary stage 6 Prototype design (G)
Iterative Redesign [50-60 min]	Iteration of previous steps depending on decisions made (C/G) Improving the design (G) Final discussion about design solutions and scientific concepts (C)	Design diary stage 7 Final design solution (G) Final reflection (individual)

Note. Reprinted from Van Breukelen, Van Meel, et al. (2016). C = class activity or product; G = design group activity or product.

a Available resources: electronic learning environment (ELE), smartphones, laptops, tablets, Microsoft Office software, internet access, materials and tools for design realization, materials for conducting experiments

b Design diary (ELE-archived): reflections, feedback, process descriptions and pictures/movies. Bulleted lists are stage-specific.

Preliminary studies and task adjustments

The challenge discussed in the previous section was adjusted towards better concept learning for use in this study. As discussed the adjustments listed below mainly concerned explicit teaching and scaffolding strategies.

Backward design. The pre- and post-exam outcomes of the preliminary studies revealed that high gain question were strongly task-related and crucial for succeeding. Thus, detailed task analysis is important to unravel task-exposed and -underexposed concepts and to predict learning outcomes. Additional less directive concepts, complementing the knowledge domain, should be addressed otherwise (teacher-driven) through additional teaching interventions. This approach corresponds to the idea of backward design (Wiggins & McTighe, 2006) that states that education designers must begin to think about assessment and objectives before

deciding what to do and how to teach. Regarding the solar challenge, initially designed for Study 2, there were four topics of underexposed science: (1) conceptual nature of resistance, (2) nature of electrical energy and energy dissipation, (3) behavior of current in components, (4) effect of voltage changes on circuit operation. To explore 1 through 3 students used simulation software and the fourth topic was addressed by additional experimentation. All topics, addressed as interludes, were complemented by class discussions and didactic analogies for clarification. Topic 1 and 3 were addressed after experimentation (stage 3) because experimentation contained resistance-current measurements. Topic 2 was addressed after design testing (stage 6) because then efficiency calculations took place. Topic 4 was addressed after the final stage by replacing the solar cells in the final design with a traditional voltage source.

Guided discussion. For teacher guidance during class discussion the technique of guided discussion (Carpenter, Fennema, & Franke, 1996) was used to highlight and explicate underlying science. When students worked in groups they were challenged to think and make sense of what they were doing. Then, by observing students' thinking and doing it became clear what individual students understood about science. Based on this, the teacher made notes about which students should present their insights during class-discussion. This might concern insights that are incorrect but useful to initiate a discussion of common misconceptions. Eventually, more sophisticated insights are used as input to head for proper reasoning and understanding. Both inputs and class discussion provide the teacher with information about students' knowledge and (existing) cognitive gaps, whereupon better understanding can be obtained.

Informed design. Informed design aims for enhancing students' prior knowledge through preparatory activities named knowledge and skill builders (KSBs) (Burghardt & Hacker, 2004). In this way, students are better prepared to approach design challenges from a more knowledgeable base and to tackle design problems by conceptual closure. Based on Study 2 an preparatory activity was created for this study around the behavior of solar cells. Students involved in Study 2 incorrectly assumed, without testing, that collar cells behave like (ideal) voltage sources. This assumption resulted in insignificant and time-consuming experimentation and finally trial-and-error behavior during design planning. To prevent this from happening students had to do, during stage 2, some information seeking accompanied by a class discussion regarding characteristics of (combined) solar cells.

Explicit instruction and scaffolding. According to Archer and Hughes (2011) explicit instruction is characterized by a series of supports or scaffolds, where students are guided through the learning process with clear statements about the purpose of and rationale for learning activities. It embraces 16 instructional elements that aim for a systematic method of teaching with emphasis on proceeding in small steps, checking for student understanding, and achieving active and successful participation by all students. LBD takes account of most of the elements and the adjustments mentioned before also fit into explicit instruction. However, teacher guidance should also fit the educational setting. Design challenges face teachers with a new kind of classroom control (Wendell, 2008) where teachers must relinquish directive control (Burghardt & Hacker, 2004). Thus, teachers need to develop pedagogical strategies for guiding complex design-based science tasks (Bamberger & Cahill, 2013). Study 2 that investigated these strategies resulted in a framework of important teaching guidelines that were directive for teacher handling during this study.

Adjustment of the design diary. Students involving Study 1 and 2 hackled the amount of administration (design diary) mainly because little administration was necessary to learn or move on. For example, there was a lot of requested reflection but in too few occasions this affected advancement directly. As a result reflection became disturbing and abortive. Therefore, administration was reduced and many written proceedings were replaced by process pictures accompanied by short subscriptions.

Methodology

21 first-year student teachers (science) took part in this design-based mixed methods study where they faced the improved solar challenge. All participants had prior experiences on characteristic LBD elements and

© i-STEM 2015, j-stem.net

sufficient prior knowledge regarding the science domain. The challenge was guided by the principal investigator (teacher trainer) because of the relatively small number of participants. According to Crouch and McKenzie (2006) this offers the possibility to establish a sustainable relationship with participants and to provide added depth to the study, all resulting in an increased validity.

Quantitative data was collected to study students' progress in concept learning and video recordings were used to generate quantitative data about skill performances. Qualitative data was used to discover how task improvements affected concept learning by comparing students' comments to previous studies.

Data collection

To study a change in conceptual understanding the pre-post-exam developed for Study 2 was used. This multiple choice test is based on the validated Determining and Interpreting Resistance Electric Circuit Test (DIRECT) specially designed for use with high school and university students (Engelhardt & Beichner, 2004). The exam contains 46 items where each objective in Table 2 is served by multiple questions. Study 2 showed, by using a control group, that there was no task learning effect from just completing the test.

Study 1 showed that students mainly learned incomplete concepts and had difficulty in making proper knowledge connections and therefore did not achieve deeper conceptual understanding. This conclusion was partially based on multiple choice tests. According to Stoddart, Abrams, Gasper, and Canaday (2010) close-ended tests like this often fail to measure conceptual understanding because students easily can make guesses and therefore knowledge structures remain invisible. Using concept maps is suggested as a more meaningful way of assessing conceptual understanding. Therefore, beside multiple choice testing, students were asked to create a concept map before and after the challenge. For this, a proposition-based concept map test was developed, based on Yin, Vanides, Ruiz-Primo, Ayala, and Shavelson (2005), where students had to create 16 fundamental propositions (a connection between two concepts by using linking words or phrases) within a set of 10 predefined task-related concepts.

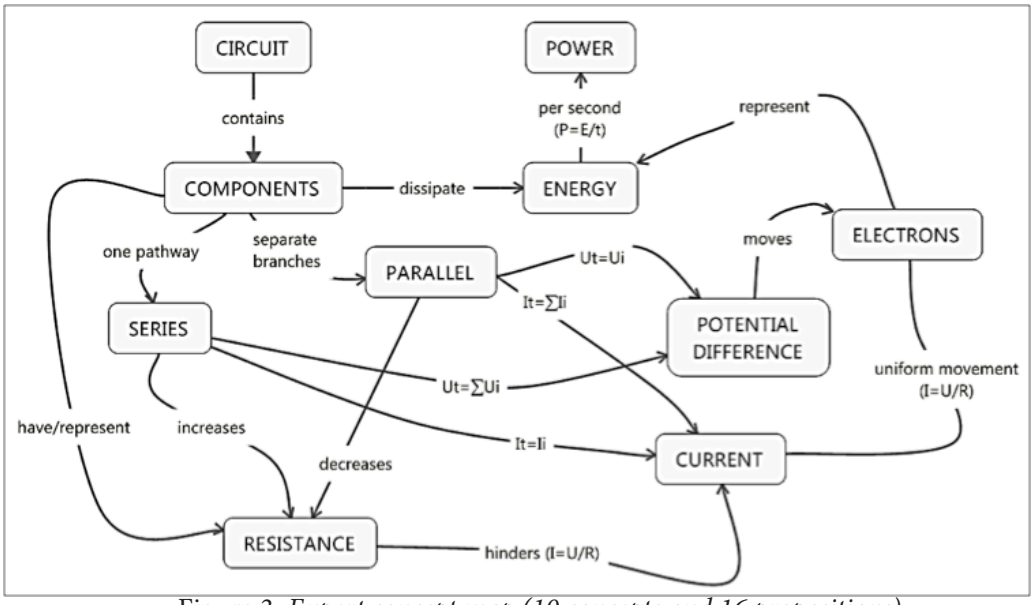

Figure 3. *Expert concept map (10 concepts and 16 propositions)*

The selection of concepts and the amount of propositions was based on a peer reviewed expert map (Figure 3) designed by two experts. According to literature, proposition-based concept map tests, based on an expert map, appear to be superior to other mapping strategies in assessing conceptual understanding (Cañas et al., 2003; Rye & Rubba, 2002). It is important to note that Yin et al. (2005) established small task learning effects in some cases due to the development of mapping skills. Those effects will be minimal for this study because students are familiar with mapping techniques.

To study an increase in students' skill performances we chose and slightly adapted the approach used in previous LBD research by Holbrook, Gray, Fasse, Camp, and Kolodner (2001) in order to make comparison possible. Students were videotaped when working, partially in groups, on similar performance tasks before and after the challenge. Tasks were taken from the Performance Assessments Links in Science Website database (SRI International Center for Technology in Learning, 1999) and were suitable for use with senior high school students (age 16-18); comparable to first-year student teachers in this study. During the pre-task students had to determine the power dissipated in a combination of two resistors connected in series to a battery. The post-task concerned the determination of how well different wires radiate heat when voltage is applied across each wire. Both tasks included three parts: (a) students designed an experiment or procedure for fair testing, (b) students ran an specified experiment and collected data, and (c) students analyzed the data to draw conclusions and make recommendations. The videotapes were analyzed, also according to Holbrook et al. (2001), on seven science-related dimensions (Table 4): negotiations during collaboration, distribution of efforts and tasks, attempted use of prior knowledge, adequacy of prior knowledge, scientific reasoning, experimentation skills and self-checks. Because the dimensions contain a mix of individual and collaboration skills each activity (a-c) started with an individual preparation, followed by a sharing session and ended with task completion by teamwork.

Afterwards an open-ended questionnaire was used to investigate students' views on which activities stimulated or impeded concept learning. Questions were based on the STARR method that provides a framework for reflection on learning outcomes (Verhagen, 2011). By interpreting students' answers, also in the light of preliminary studies, it is possible to establish whether the improvements are appreciated or room for improvement is left. Open-ended questions were used to prevent students' views from being swayed by possible answers. To confirm that the questionnaires' data reduction and interpretations were fair (respondent validation) nine students, the number that made themselves available, were interviewed simultaneously. During this session also some remarkable differences and correlations regarding learning outcomes were discussed for deeper understanding.

Analysis

The multiple choice tests were scored per student by the proportion of correct answers among 46 items. Proportions were used to calculate the gain (g): ratio of actual average gain (post – pre) to the maximum possible average gain (1 – pre) (Hake, 1998). A paired samples t test and a Wilcoxon signed-rank test were performed to investigate differences between pre- and post-scores. This combination was used because literature indicates that for relative small sample sizes using both tests increases the possibility to detect type I and II errors (Meek, Ozgur, & Dunning, 2007). Establishing Cronbach's alpha revealed the internal consistency of the exam.

The concept maps were scored per student. For this, all propositions (16 per concept map) were rated by two experts individually. Based on Yin et al. (2005) and Rye and Rubba (2002) the following scores were awarded: 3 points for a scientifically correct expert proposition (analogous to the expert map in Figure 4), 2 points for other correct propositions, 1 point for a weak or partially correct proposition and no points for incorrect propositions. Based on the experts' allocated scores the linear weighted Cohen's Kappa was calculated, which was sufficient. Then, the experts' average scores were assigned as final scores. Finally, the proportion scores (based on a 48 maximum score) were used, similar to multiple choice test analysis, to calculate gains and to investigate pre-post-score differences.

Analysis of the videotaped performance assessments took place by using a scoring rubric (See Appendix) where each performance dimension was served by a 5-point rating scale (1-5), with 5 being the highest level/score. Although the rubric's scale and dimensions are similar to that used by Holbrook et al. (2001) the level descriptors were adjusted for more validity. The original rubrics assessed skill performances by capturing the extent to which students in a group participated in practicing a skill: if more students were actively involved the group got a higher rating. According to Jonsson and Svingby (2007) this (possibly) causes validity problems because this method fails to reveal the quality of students' individual performances. Because

© i-STEM 2015, j-stem.net

a well-validated rubric, matching all our skill dimensions, was not available a rubric was created by combining existing rubrics. For this, we used an available qualitative framework of criteria to guaranty an acceptable level of validity, because a more sophisticated approach, achievable within this study, is still in its infancy (Baartman, Bastiaens, Kirschner, & Van der Vleuten, 2006; Moskal & Leydens, 2000). In short, rubrics compromising the following properties were selected: applicability to a 5-point scale, level descriptors based on observable behavior of individuals, univocal descriptors that actually reflect the skill dimension, some degree of validation, development based on experiences, expert involvement and suitability for the target group. Based on the final rubric two experts rated students' skill competences individually whereupon, after establishing an acceptable Cohen's Kappa, the experts' average ratings were assigned as final scores. Differences between pre- and post-ratings were also tackled by paired samples t tests and Wilcoxon signed-rank tests.

To investigate the strength of the relationships between pre- and post-assessment variables the Pearson product-moment correlation coefficient was computed for all possible combinations of variables. It is particularly interesting to find out how the multiple choice test and concept map test are correlated because they both concern conceptual knowledge. Furthermore, it reveals which skills strongly interact with concept learning.

For questionnaire analysis, at first to categorize responds, we distinguished between positive and negative opinions on the process of conceptual learning. After this, within these categories common themes were grouped and tagged by a description resulting in sub-categories of impeding or stimulating properties. Finally, all questionnaires were re-read to make sure all responses were categorized properly. During the group interview all properties were discussed and, based on students' input, slightly customized or filled up. Finally, remarkable differences and correlations regarding assessment outcomes were accompanied by a uniform group opinion on how to interpret results. For theoretical underpinning of students' opinions scientific literature was searched through.

Results

Table 4 gives a complete overview of all pre- and post-assessment results per student including mean scores and standard deviations. For the multiple choice test the average Cronbach's alpha, based on individual objectives in Table 2, is 0.72 for the pre-test and 0.69 for the post-test. The linear weighted Kappa values for the concept map and performance assessment analysis are shown in Table 5. Thus, in case of all assessments the reliability is sufficient.

The conceptual learning gains in Table 4 for the multiple choice test are significant, $t(20) = -30.87$; $p < 0.001$, just as for the concept map test, $t(20) = -24.58$; $p < 0.001$. This is confirmed by the Wilcoxon signed-rank test that gives the same p value for both tests. The mean gain for the multiple choice test (0.68) significantly increased compared to Study 2 (0.49) and Study 1 (0.35) and exceeds conceptual gains (LBD) found by Holbrook et al. (2001) that revealed mean gains up to 0.40. Compared to a large previous survey of pre-post-test multiple choice data for physics courses (Hake, 1998), that showed maximum gains between 0.60 and 0.70, our students were equally successful. Remarkably, the highest gains found by Hake (1998) resulted from interactive engagement (IE) methods designed, similar to LBD, to promote conceptual understanding through heads- and hands-on activities contributed by (peer) feedback, collaboration and intensive teacher guidance.

Table 4

Results pre- and post-assessments (n = 21): multiple choice test, concept map test, skill performanceS

Stud.	Multiple choice test			Concept map test			Performance assessment[b]													
	Proportion score			Proportion score			Negotiations		Distr.effort/ tasks		Prior know.use		Prior know. adeq		Scient.reason-ing		Experimenta-tion		Self-checks	
	Pre	Post	Gain[a]	Pre	Post	Gain[a]	Pre	Post	Pre	Post	Pre	Pre	Post	Pre	Post	Pre	Post	Post	Pre	Post
1	0.37	0.80	0.69	0.40	0,72	0,57	2,00	3,00	2,50	3,00	2,00	3,00	1,50	4,00	2,00	3,00	3,00	4,00	1,50	4,00
2	0.37	0.78	0.66	0.40	0,64	0,43	3,00	4,00	3,00	5,00	1,00	2,50	1,00	3,00	1,00	2,50	2,50	3,50	2,00	3,00
3	0.39	0.78	0.64	0.41	0,70	0,53	2,50	3,00	3,00	3,00	1,50	2,50	2,00	3,00	1,50	3,00	3,00	4,50	2,00	4,00
4	0.57	0.91	0.80	0.58	0,79	0,56	1,00	2,00	2,00	3,00	4,50	5,00	4,00	5,00	3,00	4,00	3,50	4,00	2,00	4,00
5	0.46	0.83	0.68	0.49	0,67	0,38	1,00	2,00	1,50	3,00	3,00	4,00	3,00	4,50	3,00	4,00	2,50	4,00	1,00	4,00
6	0.35	0.78	0.67	0.36	0,63	0,44	3,00	3,00	2,00	2,50	2,00	3,00	2,00	3,00	2,00	4,00	3,00	3,50	2,00	5,00
7	0.28	0.70	0.58	0.33	0,56	0,37	2,50	3,50	3,00	3,00	2,00	3,00	1,50	3,50	1,00	2,50	3,00	5,00	2,00	3,50
8	0.41	0.78	0.63	0.43	0,70	0,51	2,00	4,00	2,50	3,00	3,00	4,00	3,00	4,00	2,50	4,00	3,50	4,50	2,00	4,00
9	0.41	0.78	0.63	0.41	0,69	0,51	2,00	2,50	2,00	3,00	2,50	4,00	2,50	4,00	2,50	4,00	1,50	4,00	1,00	4,00
10	0.28	0.85	0.79	0.29	0,66	0,55	1,00	2,00	1,50	3,00	1,00	3,00	1,50	4,00	1,00	3,00	2,00	3,50	2,00	3,50
11	0.33	0.72	0.58	0.33	0,60	0,43	2,00	4,00	2,00	4,50	2,00	4,00	2,00	3,50	2,00	3,00	2,00	3,00	2,00	3,00
12	0.46	0.78	0.60	0.43	0,67	0,45	2,50	3,00	2,00	3,50	2,00	3,00	2,50	3,50	3,00	3,50	2,00	3,00	1,50	4,00
13	0.33	0.76	0.65	0.38	0,61	0,41	2,00	3,50	3,00	4,00	2,00	3,00	1,50	3,00	1,00	3,00	1,50	3,00	2,00	3,00
14	0.35	0.74	0.60	0.44	0,64	0,38	3,00	4,00	2,00	4,00	2,00	4,00	2,00	4,00	2,00	3,50	2,50	3,50	2,00	4,00
15	0.39	0.83	0.71	0.39	0,68	0,51	2,00	4,00	2,50	5,00	2,00	4,00	2,00	4,00	1,50	3,00	3,00	4,00	1,00	4,00
16	0.39	0.83	0.71	0.42	0,71	0,54	2,00	2,50	2,00	3,00	2,00	4,00	2,00	4,00	2,00	3,00	3,00	4,00	2,00	5,00
17	0.46	0.87	0.76	0.39	0,75	0,64	3,50	4,00	3,00	4,00	3,50	4,50	3,00	4,00	3,00	3,50	3,00	3,00	2,50	5,00
18	0.24	0.72	0.63	0.28	0,59	0,46	2,00	3,50	3,00	4,00	1,00	3,00	1,00	4,00	1,50	2,50	3,00	4,00	2,00	4,50
19	0.22	0.80	0.75	0.30	0,61	0,48	2,50	3,00	2,00	4,00	1,50	3,50	1,00	4,00	1,00	2,00	3,50	4,50	3,00	3,00
20	0.39	0.83	0.71	0.36	0,69	0,54	2,50	3,00	3,00	3,50	2,00	4,00	2,00	4,00	1,00	2,50	3,00	5,00	2,50	4,00
21	0.37	0.83	0.72	0.42	0,73	0,58	2,00	4,00	3,00	4,00	2,00	3,50	2,00	4,50	1,50	3,00	2,50	4,00	3,00	5,00
Mean (SD)	0.37 (0.08)	0.80 (0.05)	0.68 (0.07)	0.39 (0.07)	0.67 (0.06)	0.49 (0.07)	2.19 (0.66)	3.21 (0.72)	2.40 (0.54)	3.57 (0.71)	2.12 (0.84)	3.55 (0.67)	2.05 (0.76)	3.83 (0.53)	1.86 (0.74)	3.17 (0.60)	2.69 (0.60)	3.88 (0.61)	1.95 (0.55)	3.98 (0.66)

SD = standard deviation; Stud. = student

a Gain: (g) = (post%-pre%)/(100-pre%); where: high gain: (g) ≥ 0.70, medium gain: 0.70 > (g) ≥ 0.30, low gain: (g)< 0.30 (Hake, 1998)

b Mean scores awarded by two experts per skill dimension based on a 5-point scoring rubric (See Appendix), with 5 being the highest rating.

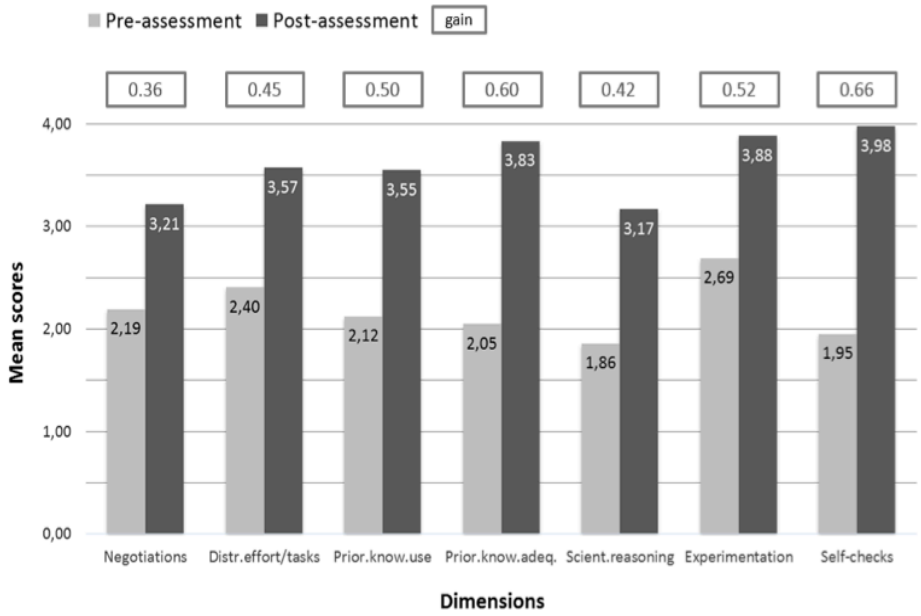

Figure 4. *Mean skill performances based on assessment results (Note: Dimensions and results are based on Table 4. Gain = (post-score – pre-score)/(100 – pre-score).*

Incidentally, a critical comment should be made because the concept map tests showed significantly lower ($p < 0.001$), but still substantial, gains (mean gain = 0.49; lowest gain = 0.37; highest gain = 0.64). This will be discussed in more detail later on.

Table 5

κw for concept map and performance assessment ratings

Test	kw pre-ratings	kw post-ratings
Concept map test	0.68 (lower limit = 0.62; upper limit = 0.74)	0.65 (lower limit = 0.58; upper limit = 0.72)
Performance assessment	0.62 (lower limit = 0.52; upper limit = 0.72)	0.66 (lower limit = 0.58; upper limit = 0.74)

Studying the performance assessment results in Table 4, shown graphically in Figure 4, it indicates that all skill dimensions show an increase in achievement level, where the highest progressions concern the adequacy of prior knowledge, experimentation skills and self-checks. However, all improvements are significant ($p < 0.001$) and fairly comparable to the performance assessment results found by Kolodner, Camp, et al. (2003). Those results showed scores between 2.00 and 3.00 for honors non-LBD students (the category befitting the students in our study at pre-testing) and scores up to 4.00 for typical LBD students (students exposed to LBD). Overall, students in this study reached, compared to previous LBD studies, much higher conceptual learning gainswhile advancement in skill performances was not hindered.

According to Table 6 there were strong (significant) positive correlations between the pre-scores of the multiple choice and concept map test, as well as for the post-scores. The gains of both tests showed a lower, but fair, correlation ($r = 0.683$, $n = 21$, $p < 0.01$) that can be explained by the fact that the mean gain for the concept map test was significant lower compared to the multiple choice test. According to Constantinou (2004) multiple choice test and concept map test results vary to a greater or lesser extent depending on which kind of learning is assessed through the multiple choice test (e.g. rote learning or meaningful learning). In general, Ruiz-Primo, Schulz, and Shavelson (1997) state that the correlation between both tests should be positive because they measure the same knowledge domain but the magnitude may differ. The interviewed students all agreed that the concept map test was more difficult because it stronger appealed to mastering well organized, relevant knowledge structures.

Furthermore, Table 6 shows moderate or strong positive correlations between the conceptual tests and three dimensions of the performance assessment (use and adequacy of prior knowledge and scientific reasoning) that also positively correlated with each other. Other positive or negative correlations between variables were not found or appeared to be weak or occasional. These findings correspond to previous findings: Schreiber, Theyßen, and Schecker (2016) found high correlations between conceptual tests and the preparation and evaluation of experiments, where prior knowledge and scientific reasoning are important, and low correlations with respect to conducting the experiment by following the rules for fair experimentation. Stone (2014) states that general skills, like collaboration and reflection, have a limited interconnectedness with more science-specific skills (practices) and the knowledge domain, where Zimmerman (2000) explicitly mentions the weak relation between conducting reception experiments and mastering conceptual knowledge and strong relations between conceptual knowledge, prior knowledge and scientific reasoning. All these insights perfectly reflect our findings where the interviewed students also emphasized the concept-free character (according to science knowledge) of collaboration, reflection and conducting a prescribed experiment. On the other hand students compared the use and adequacy of (prior) knowledge combined in combination with scientific reasoning to the mental activity important for creating a concept map, which reflects the mastering of knowledge structures.

Table 6

Positive Pearson product-moment correlations of pre- and post-assessment results (n = 21)

Variables		1	2	3	4	5
1. Multiple choice test	Pre	-				
	Post	-				
2. Concept map test	Pre	0.900**	-			
	Post	0.831**	-			
3. PA Prior knowl. use	Pre	0.790**	0.766**	-		
	Post	0.496*	0.500*	-		
4. PA Prior knowl. adequacy	Pre	0.883**	0.817**	0.920**	-	
	Post	0.568**	0.550**	0.725**	-	
5. PA Scientific reasoning	Pre	0.758**	0.689**	0.753**	0.834**	-
	Post	0.692**	0.426	0.447*	0.428*	-

Note: Based on pre-post-results in Table 4; PA = performance assessment; *p < 0.05; **p < 0.01 g

Table 7 shows the results of the questionnaire analysis where the amount of positive replies largely exceeds the negative ones. According to students, activities that directly appeal to underlying science (explicit teaching and experimentation) are invaluable for concept learning complemented by sufficient teacher and task guidance (feedback, clear instructions and transparency). These results are perfectly consistent with the results of Study 1 and 2. It is, however, surprising that in this study learning form peers is clearly more appreciated. Maybe because this study revealed less trial and error behavior and therefore students acted more like a role model or because the guided discussion approach, where the use of students' insights is directive, clarifies that peers are an important learning source. Although interviewed students seemed to confirm both statements a sloid validation failed to appear.

© i-STEM 2015, j-stem.net

Table 7

Positive and negative opinions about concept learning (questionnaires; n = 21)

Influences on concept learning			
Stimulating factors (N = 106)	Perc.	Impeding factors (N = 44)	Perc.
Explication of underlying science by the teacher	28	Fragmentation of addressed science	36
(Results of) conducted experiments and simulations	20	Fragmentation of the task	29
Learning from peers: collaboration, sharing information, peer feedback	16	Uncertainty and uncomfortability (because of the new educational setting)	13
Teacher feedback during the task (no direct explication of science)	13	Too little attention to deeper assimilation of addressed science.	11
Clear instructions and transparency of tasks and objectives	12	Other (e.g. false information sharing, task duration)	11
Other (e.g. reflection, information seeking, the design context)	11		

Note: Perc. = relative distribution (%) of all replies within each category on which the corresponding sub-categories were distracted. Descriptions were redefined based on the group interview.

Taking the impeding factors into account fragmentation is an issue. First, students experienced too little coherence in addressed science and, second, students described the number of stages and accompanying administration as disruptive to the ongoing learning process. Also some students missed assimilation of addressed science for anchoring. Compared to the preliminary studies, the initial problems of addressing an incomplete science domain and a lack of science explication seem to be tackled. However, despite the improvements, coherence still is an issue and the amount of administration is still disruptive.

Discussion and implications

The adjustments, deduced from two preliminary studies, to enhance concept learning by design challenges appear to be successful because this study reveals a solid improvement of conceptual learning gains without reducing a positive effect on skill performances. Especially when the multiple choice test results (the assessment form used in all studies) are taken into account: gain-indices increased from the lower limit of medium (> 0.30) up to the very margin of high (0.70) where the latter is more or less reserved for the most successful physics-related courses (Hake, 1998). Students' responses show, which can be considered as an important reason for improved concept learning, that in contrast with the preliminary studies little comments were made about a lack of (explicit) science teaching. It seems that a combination of backward design, guided discussion and informed design is an appropriate remedy to enhance concept learning by extending strongly task-driven concepts and further deepening of all concepts. This happens, first, by introducing additional teacher-driven concepts (weakly task-driven) to complement the knowledge domain; important for understanding individual concepts. Second, by explicating and deepening all addressed science (explicit teaching). Figure 5 illustrates how contributions to conceptual learning gains may possibly collude where, of course, nearly always room for further improvement is left.

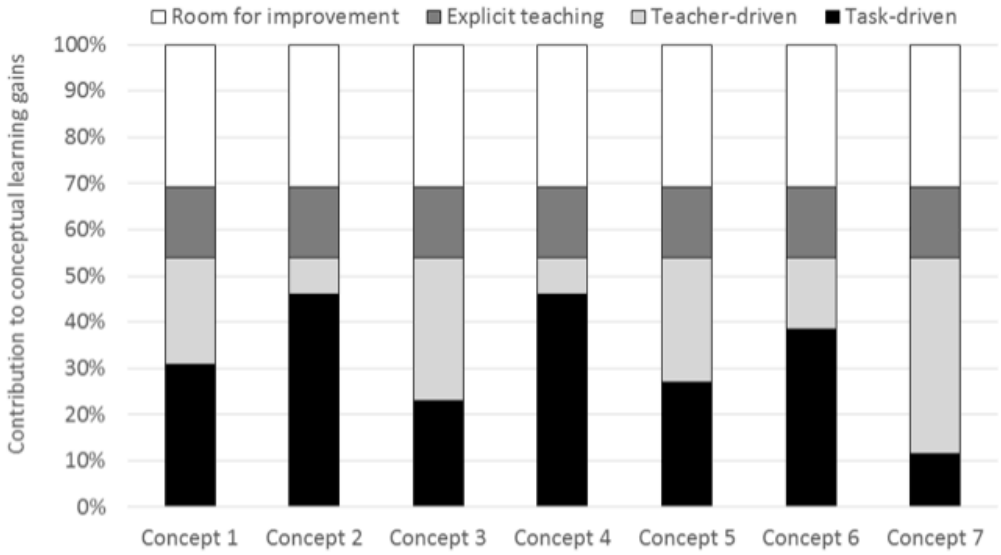

Figure 5. *Qualitative model for contribution to conceptual learning gains (Note: The (distribution of) percentages and amount of concepts are fictional and just used for illustration purposes where there is some resemblance to percentages found in the series of LBD studies. Concepts 1 through 7 represent the addressed knowledge domain.)*

This study provides some interesting clues where to search for further improvement. First, this study reveals significant positive correlations between students' conceptual performances, the use and adequacy of (prior) knowledge and scientific reasoning. Second, although students reached substantial conceptual learning gains the concept map test gains were significantly lower than multiple choice test gains. Third, students compared the use and adequacy of (prior) knowledge in combination with scientific reasoning to the mental activity important for creating concept maps. Fourth, students mentioned the fragmentation of addressed science and a lack of deeper assimilation of addressed science as important shortcomings. Thus, combining all four, more coherence of addressed science may be valuable because mastering explicit interrelationships between domain concepts enhances learning (Brandsford et al., 2003; Wiggins & McTighe, 2006). This may also improve the adequate use of knowledge and scientific reasoning and, with this, meaningful learning (important for concept mapping).

Beside fragmentation of addressed science, according to students, the same comment applies to the task itself. Although students experienced sufficient guidance and task transparency they described the number of stages and accompanying administration as disruptive to the ongoing learning process. Maybe a reduction of the number of (separate) stages and activities, through amalgamation, offers more coherence and less administration where guidance and scaffolding is shifted towards the ongoing process itself rather than breaking down into parts.

To conclude, both aspects of fragmentation, as discussed before, will be a main topic for further research. However, in general this study revealed some more interesting research themes. First, it is interesting to study the interaction between skill and concept learning in more detail because both types of learning are (partly) correlated and may strengthen each other. All the more because learning (STEM) skills is regarded as an important goal for modern education driven by a complex world economy that demands for those skills (ICF and Cedefop for the European Commission, 2015). Second, more insight is needed into the creation, use and validation of rubrics to assess skills. Third, correlations between multiple choice tests and concept map tests are often significant but widely spaced (Constantinou, 2004). Therefore it is necessary investigate this correlation in more detail and to find out how conceptual knowledge can be assessed properly depending upon the learning objectives.

© i-STEM 2015, j-stem.net

Acknowledgement

This work was supported by the Netherlands Organisation for Scientific Research (NWO) under Grant number 023.001.030.

References

Abdul Gafoor, K., & Akhilesh, P. T. (2013). Strategies for facilitating conceptual change in school physics. *Researches and Innovations in Education*, *3*(1), 34-42.

Aikenhead, G. (2006). *Science education for everyday life. Evidence-based practice*. New York: Teachers College Press.

Archer, A. L., & Hughes, C. A. (2011). Exploring the foundations of explicit instruction. In K. R. Harris & S. Graham (Eds.), *Explicit Instruction: Effective and Efficient Teaching* (pp. 1-21). New York: The Guilford Press.

Baartman, L. K., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2006). The wheel of competency assessment: Presenting quality criteria for Competency Assessment Programmes. *Studies in Educational Evaluation*, 32, 153-170.

Bamberger, Y. M., & Cahill, C. S. (2013). Teaching design in middle-school: Instructors' concerns and Scaffolding strategies. *Journal of Science Education and Technology*, *22*(2), 171-185.

Brandsford, J. D., Brown, A. L., Donovan, M. S., & Pellegrino, J. W. (2003). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.

Burghardt, M., & Hacker, M. (2004). Informed Design: A Contemporary Approach to Design Pedagogy as the Core Process in Technology. *Technology Teacher*, *64*(1), 6-8.

Bybee, R. W. (2013). *The case for stem education: Challenges and opportunities*. United States of America: National Science Teachers Association.

Cañas, A. J., Coffey, J. W., Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J., & Novak, J. D. (2003). *A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support,* 1-108. Retrieved from http://cmap.bradercomm.net:8001/rid=1186721000045_1491728413_5664/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitive guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementery School Journal*, *97*(1), 3-20.

Chan, C. (2009). *Assessement: Practical experiment*. Retrieved April 2014, from University of Hong Kong http://ar.cetl.hku.hk/am_pe.htm

Cobern, W. W. (1994). *Worldview theory and conceptual change in science education*. Paper presented at the National Association for Research in Science Teaching, Anaheim.

Constantinou, C. (2004). Concept mapping for performance assessment in physics. In P. A. M. Kommers (Ed.), *Cognitive support for learning: Imagining the unknown* (pp. 155-166). The Netherlands: IOS Press.

Crouch, M., & McKenzie, H. (2006). The logic of small samples in interview-based qualitative research. *Social Science Information*, *45*(4), 483-499.

Engelhardt, P. V., & Beichner, R. J. (2004). Students' understanding of direct current resistive electrical circuits. *American Journal of Physics*, *72*(1), 98-115.

Franker, K. (2015). *Elementary teamwork rubric.* Retrieved April 2014, from University of Wisonsin - Stout https://www2.uwstout.edu/content/profdev/rubrics/elemteamworkrubric.html

Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74.

Holbrook, J. K., Gray, J., Fasse, B. B., Camp, P. J., & Kolodner, J. L. (2001). *Assessment and evaluation of the learning by design physical science unit, 1999-2000: A document in progress,* Retrieved from http://www.cc.gatech.edu/projects/lbd/htmlpubs/progress.html

ICF and Cedefop for the European Commission. (2015). EU Skills Panorama (2014) STEM skills Analytical Highlight, 1-5. Retrieved from http://skillspanorama.cedefop.europa.eu/sites/default/files/EUSP_AH_STEM_0.pdf

International Technology Education Association. (2007). *Standards for Technological Literacy: Content for the Study of Technology (3 ed.)*. Reston, Virginia.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130-144.

Kolodner, J. L. (2002). Learning by design: Iterations of design challenges for better learning of science Skills. *Bulletin of the Japanese Cognitive Science Society*, *9*(3), 338-350.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., . . . Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design into practice. *The Journal of the Learning Sciences*, *12*(4), 495-547.

Kolodner, J. L., Gray, J. T., & Fasse, B. B. (2003). Promoting transfer through case-based reasoning: Rituals and practices in learning by design classrooms. *Cognitive Science Quarterly*, *3*(2), 1-28.

Kolodner, J. L., Hmelo, C., & Narayanan, N. (1996). Problem-based learning meets case-based reasoning. *The Journal of the Learning Sciences*, *12*(4), 495-547.

Lustig, F., West, E., Martinez, B., Staszel, M., Borgato, M. T., Iosub, I., & Weber-Hüttenhoff, U. (2009). *Experiences and results from the European project 'Integrated subject science understanding in Europe'.* Paper presented at the ESERA conference, Istanbul.

Meek, G. E., Ozgur, C., & Dunning, K. (2007). Comparison of the t vs. Wilcoxon Signed-Rank test for likert scale data and small samples. *Journal of Modern Applied Statistical Methods*, *6*(1), 91-106.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10). Retrieved from http://PAREonline.net/getvn.asp?v=7&n=10

National Science and Technology Council. (2013). *Federal Science, Technology, Engineering, and Mathematics (STEM) education: 5-Year strategic plan*. Executive Office of the President, Washington, D.C.

Office of the Chief Scientist. (2013). *Science, Technology, Engineering and Mathematics in the national interest: A strategic approach*. Australian Government, Canberra.

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*. Retrieved from http://www.nuffieldfoundation.org/science-education-europe.

Parliamentary Office of Science & Technology. (2013). *STEM education for 14-19 year olds.* Houses of Parliament, London.

Rennie, L., Venville, G., & Wallace, J. (2012). *Integrating Science, Technology, Engineering and Mathematics*. New York: Routledge.

Rhodes, T. L. (2010). *Scientific reasoning rubric.* Retrieved April 2014, from the Association of American Colleges and Universities https://www.aacu.org/sites/default/files/files/VALUE/VALUE_ScientificReasoningRubric.pdf

Rhodes, T. L., & Finley, A. (2013). *Using the VALUE rubrics for improvement of learning and authentic assessment*. Washington DC: Association of American Colleges and Universities.

Ruiz-Primo, M. A., Schulz, S. E., & Shavelson, R. J. (1997). *Concept map-based assessment in science: Two exploratory studies*. Retrieved from http://web.stanford.edu/dept/SUSE/SEAL/Reports_Papers/TECH436.pdf.

© i-STEM 2015, j-stem.net

Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics*, *102*(1), 33-44.

Schreiber, N., Theyßen, H., & Schecker, H. (2016). Process-oriented and product-oriented assessment of experimental skills in physics: A comparison. In N. Papadouris, A. Hadjigeorgiou, & P. C. Constantinou (Eds.), *Insights from research in science teaching and learning: Selected papers from the ESERA 2013 Conference* (pp. 29-43). Cham: Springer International Publishing.

Scott, D. (2008). *Critical essays on major curriculum theorists*. London: Routledge.

Sjöberg, S., & Schreiner, C. (2010). *The ROSE project: An overview and key findings*. Retrieved from http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf

SRI International Center for Technology in Learning. (1999). *Performance assessment links in science* (PALS). Retrieved Spring, 2015 http://pals.sri.com/tasks/tasks9-12.html

Stoddart, T., Abrams, R., Gasper, E., & Canaday, D. (2010). Concept maps as assessment in science inquiry learning – A report of methodology. *The International Journal of Science Education*, *22*(12), 1221-1246.

Stone, E. M. (2014). Guiding students to develop an understanding of scientific inquiry: A science skills approach to instruction and assessment. *CBE Life Sciences Educatio*n, *13*(1), 90-101.

Van Breukelen, D. H. J., De Vries, M. J., & Schure, F. A. (2016). Concept learning by direct current design challenges in secondary education. *International Journal of Technology and Design Education*, 1-24.

Van Breukelen, D. H. J., Van Meel, A. M. D. M., & De Vries, M. J. (2016). Teaching strategies to promote concept learning by design challenges. Manuscript submitted for publication.

Verhagen, P. (2011). *Reflectie met de STARR-methode Kwaliteit met beleid*. Bussum: Coutinho.

Wendell, K. B. (2008). *The theoretical and empirical basis for design-based science instruction for childre*n. Unpublished Qualifying Paper. Tufts University.

Wiggins, G., & McTighe, J. (2006). *Understanding by design. Upper Saddle River*. New Jersey: Pearson Education Inc.

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.

Zimmerman, C. (2000). *The development of scientific reasoning skills. Developmental Review*, *20*(1), 99-149.

Appendix: Scoring rubric for assessing skill dimensions

[1] Negotiations[a]				
1	2	3	4	5
Not at all	Rarely makes compromises to accomplish a common goal and has difficulty getting along with other group members.	Occasionally makes compromises to accomplish a common goal, and sometimes helps keep the group working well together.	Usually makes necessary compromises to accomplish a common goal.	Consistently makes necessary compromises to accomplish a common goal.
[2] Distributed efforts and tasks[a]				
1	2	3	4	5
Not at all	Performed little duties of assigned team role and contributed a little amount of knowledge, opinions, and skills to share with the team. Relied on others to do the work.	Performed a few duties of assigned team role and contributed a small amount of knowledge, opinions, and skills to share with the team. Completed some of the assigned work.	Performed nearly all duties of assigned team role and contributed knowledge, opinions, and skills to share with the team. Completed most of the assigned work.	Performed all duties of assigned team role and contributed knowledge, opinions, and skills to share with the team. Always did the assigned work.
[3] Using prior knowledge[b]				
1	2	3	4	5
Not at all	Identifies connections between prior events and experiences or prior concepts that relate to the context.	Identifies multiple prior experiences and events or prior concepts that relate to the context.	Prior events, experiences and concepts are identified and applied to the problem.	Prior events, experiences and concepts are routinely recalled that assist in problem solving.
[4] Adequacy of prior knowledge[b]				
1	2	3	4	5
Not at all	At least one of the mentions of prior knowledge is followed up on and is useful	A few of the mentions of prior knowledge were appropriate to the problem and used during the process.	Multiple mentions of prior knowledge were effectively selected and followed up on and were important for succeeding.	Nearly all mentions of prior knowledge are routinely followed up on and were important for succeeding and understanding.
[5] Use of science terms, scientific reasoning[c]				
1	2	3	4	5
Not at all	States ambiguous, illogical, or unsupported conclusions without proper use of science concepts. Demonstrates no ability to distinguish between causal and correlational relationships.	States general conclusions by using science terms, but beyond the scope of the inquiry findings limitations and implications. Demonstrates limited ability to distinguish between causal and correlational relationships.	States scientifically formulated conclusions focused solely on the inquiry findings. The conclusion arises specifically from and responds to the inquiry findings. Demonstrates appropriate ability to distinguish between causal and correlational relationships.	States a scientifically formulated conclusion that is a logical extrapolation from the inquiry findings limitations and implications. Demonstrates advanced ability to distinguish between causal and correlational relationships.

[6] Experimentation and fair testing[d]				
1	2	3	4	5
Not at all	Experimentation was done chaotically, without (proper) knowledge of how to control variables. At least one variable was tested.	Only some of the experimental tasks were done satisfactorily; some understanding of fair testing was present by controlling at least two (dependent) variables.	Most of the experimental tasks were done neatly and satisfactorily; an acceptable amount of understanding of fair testing and controlling for variables was present.	Experimental tasks were done in an organized and effective way by understanding for fair testing and controlling all variables.
[7] Self-checks[d]				
1	2	3	4	5
Not at all	A little (serious) self-checks took place without serious consideration for improvements.	A few self-checks took place by questioning several aspects of the procedure but were not properly addressed (without serious consideration)	(Some) self-checks took place by questioning several aspects of the procedure but were not optimally addressed.	Self-checks took place by questioning several aspects of the procedure and affected in improvements while performing the task

a Adapted from Franker (2015).
b Adapted from Rhodes and Finley (2013).
c Adapted from Rhodes (2010).
d Adapted from Chan (2009).

© i-STEM 2015, j-stem.net