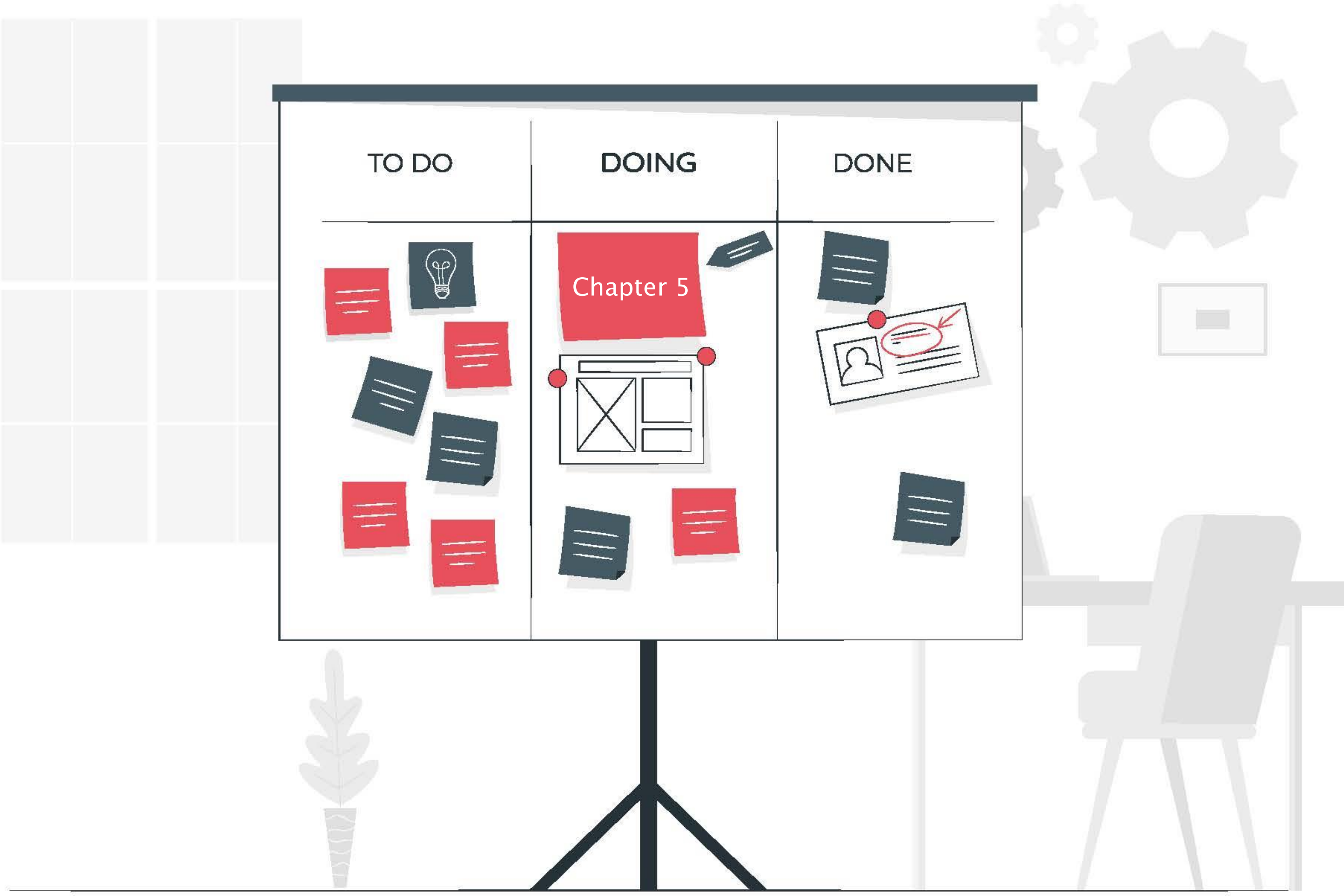
Author: Vogelzang, J.
Title: Scrum in secondary chemistry education: A methodology to support teachers and to scaffold students
Issue Date: 2020-11-10

Chapter 5

Scrum methodology in context-based secondary chemistry classes: Effects on students' achievement and their affective and metacognitive outcomes



This chapter is based on:

Vogelzang, J., Admiraal, W. F., & van Driel, J. H. (*submitted*). Scrum methodology in context-based secondary chemistry classes: Effects on students' achievement and their affective and metacognitive outcomes.

Abstract

Teaching with Scrum methodology includes ceremonies, roles and artefacts supporting students in planning, monitoring and directing their learning process. It scaffolds students' learning in complex and sometimes overwhelming context-based learning environments.

Effects of the implementation on both students' learning outcomes and self-reported perceptions of six affective and metacognitive outcomes were investigated.

Six teachers implemented Scrum methodology in a context-based secondary chemistry course on Green Chemistry, forming the experimental group. Based on the quality of the implementation, teachers of the experimental group were subdivided into top-teachers and growth-teachers. The comparison group, which did not use Scrum methodology, consisted of four teachers. A pre-test post-test control group design was used to study its effect on students' achievements and self-reported affective and metacognitive outcomes.

Students of both experimental groups outperformed students of the comparison group with a large effect-size (top-teachers); and medium effect-size (growth-teachers) on learning outcomes. Findings on affective and metacognitive outcomes revealed medium and small effects of Scrum methodology.

Despite the fact that the implementation is challenging for teachers, it appears that Scrum methodology has positive effects on students' achievement and their affective and metacognitive outcomes.

Keywords: context-based approaches, Scrum methodology, secondary chemistry education.

5.1 Introduction

Context-based learning approaches are developed and introduced to address several challenges in secondary chemistry education (Childs et al., 2015; Sevian et al., 2018). It has been suggested that students perceive the chemistry curriculum as fragmented, overloaded and as irrelevant for their personal lives (Gilbert, 2006). Moreover, transfer of chemistry concepts to new situations turns out to be difficult (Pilot & Bulte, 2006a). In general, a context-based approach starts with a relevant real-world question, stimulating students to use chemistry concepts to explore, experience and evaluate a problem extracted from real life (Taconis et al., 2016). Furthermore, educational research revealed that context-based approaches, in general, have positive effects on both students' motivation, and their perception of relevance of chemistry for their personal lives (Bennett, 2017). Effects found on students' achievements are ambiguous, with generally similar or slightly better results than more traditional approaches (Savelsbergh et al., 2016).

However, students might experience context-based approaches as rather complex and overwhelming (King & Ritchie, 2012; Quintana et al., 2004). Although real-world questions are usually attractive and motivating for students, their complexity and open-endedness might evoke frustration and confusion. Moreover, the real-world question might conceal the underlying concepts (Parchmann et al., 2006), causing uncertainty among students about what to learn for their end-of-term exams. In addition, answering a complex and ill-structured real-world question sets high demands on self-regulation processes, such as monitoring progress, and adjusting learning strategies to optimize performance (Dori & Avargil, 2015; Panadero, 2017). Unfortunately, students do not always monitor and regulate these processes during learning with context-based approaches, which limits its potential and effectiveness as educational method to enhance their learning (Azevedo, Behnagh, Duffy, Harley, & Trevors, 2012).

Therefore, implementing a scaffold to guide students through their learning process and to monitor their conceptual development might enhance students' learning. Scaffolding makes complex tasks more manageable and accessible for students (Hmelo-Silver et al., 2007). An interesting approach would be to implement a project management framework in a context-based learning environment and explore whether students experience a context-based approach as less complex. Project management frameworks are used in business to guide employees working on complex projects. In addition, a project management framework consists of features to monitor the project's progress and to adjust to changing circumstances. It seems reasonable to expect that the implementation of a project management framework might be beneficial for students working on a rather complex real-world question.

For that reason, we decided to implement Scrum methodology as framework to scaffold students' learning in context-based learning environments. Scrum methodology provides ceremonies, roles and artefacts that invite students to reflect on their learning progress (Vogelzang, Admiraal, & van Driel, 2019) and aims to scaffold students' self-regulation processes (Pope-Ruark, 2012). As yet, Scrum methodology is seldom used in educational settings, although there are some exceptions (Cook, 2017; Parsons & MacCallum, 2019; Pope-Ruark, 2012). Therefore, insights in the effects of the implementation of Scrum methodology on students' learning are still lacking. The current study can be characterised as an exploratory study, to find out whether the use of Scrum methodology in an educational context might increase students' learning in rather complex learning environments such as context-based chemistry education.

5.2 Context-based learning

In general, context-based approaches aim at various student outcomes, including cognitive, metacognitive and affective ones (Bennett, 2017; Dori & Avargil, 2015). First, context-based approaches intend to promote students´ cognitive development and increase their achievement by connecting a recognizable real-life issue to chemistry concepts. Instead of transmitting knowledge to them by a teacher, students are invited to relate concepts (De Putter-Smits, Taconis, & Jochems, 2013) and to construct meaningful knowledge structures themselves (Krajcik & Shin, 2014). These knowledge structures consist of interconnected concepts rather than as isolated facts. Students construct and use these knowledge structures to interpret, analyze, and answer the real-world question (Hmelo-Silver, 2004). There is widespread consensus in the field of learning sciences that such student-centered approaches contribute to deep learning and sound understanding (Sawyer, 2014).

Secondly, context-based approaches require students to use and develop their metacognitive skills. Answering an open real-world question demands skilled use of self-regulation processes, including planning, monitoring, reflection, and control of learning strategies (Azevedo et al., 2012, p. 171; Dori & Avargil, 2015, p. 123). Self-regulation of a learning process involves analysing the learning context, setting learning goals, determining learning strategies, reflecting whether the strategies are effective, and monitoring learning progress (Azevedo et al., 2012).

Thirdly, context-based approaches intend to improve students' attitude towards chemistry. This affective aim focuses on how students feel about the chemistry they do (Bennett, 2017, p. 23). This is of utmost importance given the fact that in many countries students' attitude towards chemistry is generally negative (Potvin & Hasni, 2014; Savelsbergh et al., 2016). By creating a learning environment in which students are rather autonomous in how they arrange their work, the hope is that this might engage students and influence their feelings of ownership (Katz & Assor, 2007). In addition, the context-based approach might contribute to a more realistic picture of chemistry by showing how it is connected to their personal lives. The aims of context-based approaches show that its nature is constructivist.

Obviously, context-based approaches share many similarities with student-centered learning environments, such as problem-based learning environments, in which students: 1) work collaboratively on real-world issues; 2) take ownership of their learning; 3) reflect explicitly on their learning process and 4) use higher order thinking skills (Land et al., 2012, p. 11). However, educational research revealed that students of all ages often perceive difficulties in such complex learning environments and students should be scaffolded to engage successfully in student-centered learning environments (Azevedo et al., 2012). This study focuses on Scrum methodology as scaffold of students' learning in context-based learning environments.

5.3 Scrum methodology

Scrum methodology is a project management framework, frequently used in business and industry to manage and monitor complex projects (Schwaber & Sutherland, 2012). It consists of ceremonies, roles and artefacts that provide overview, and keeps employees on track to achieve the ultimate objectives of the project. Only recently Scrum methodology was adjusted for educational purposes (Cook, 2017; Parsons & MacCallum, 2019). Cook (2017) introduced Scrum methodology in a project-based learning unit on climate change and found that her students showed initiative, were self-directed and collaborated closely with their team mates (Cook, 2017, p. 90). Furthermore, Parsons and MacCallum (2019) edited a book with experiences of the implementation of Scrum methodology in a variety of fields, including software development courses, mathematics and chemistry. In general, the experiences suggest that the use of Scrum methodology might increase metacognitive aspects of students' learning, such as students' planning skills and mutual collaboration. In addition, the experiences seem to suggest that cognitive learning outcomes as well as affective aspects of students' learning, such as attitude towards subject, are influenced positively. However, empirical research into the effects of the implementation of Scrum methodology on cognitive, metacognitive and affective aspects of students' learning is lacking.

A concise overview of Scrum methodology including a short description of how it was used by the teachers in this study, is presented in Table 1. Moreover, its features have been connected to motivational scaffolds. As can be seen in Table 1, Scrum methodology aims at stimulating a classroom climate in which students work together, deploy and develop their personal qualities. In addition, their self-regulatory skills, their self-efficacy and their self-reflection, might be scaffolded by the different ceremonies, roles and artefacts of Scrum methodology. Therefore, we might hypothesize that Scrum methodology provides a framework that scaffolds students' ownership, stimulates discussion and reflection, and contributes to a learning environment in which students think more deeply about chemistry concepts, are more engaged with the subject of teaching and regulate their learning process more, compared to regular teaching. In the current study, effects of Scrum methodology were examined on students' learning outcomes and their self-reported affective and metacognitive skills.

Two research questions guided this study:

1. *What is the effect of the use of Scrum methodology on students' learning outcomes? (RQ1)*
2. *What is the effect of the use of Scrum methodology on students´ perceptions of affective and metacognitive aspects of their learning? (RQ2)*

Table 1. Concise overview of the ceremonies, roles and artefacts of Scrum methodology.

Step	Ceremony	Roles	Artefacts	Description
1	Introduction	Product owner		The teacher, as product owner, introduces the real-world question. Students are encouraged to ask questions about the objectives of the context-based approach.
2	Team forming	Scrum master		Groups of approximately four persons are formed based on qualities they want to share with their team, such as taking leadership, perseverance or clear writing etc. Each team has a Scrum master, the first among equals. This student takes responsibility for the team's progress and is the contact point for the product owner.
3			Product backlog	The product owner provides all teams with a product backlog, which consists of all learning goals, tasks, exercises and experiments necessary to answer the real-world question.
4	Definition of done Definition of fun			Before starting to work on all assignments, every group formulates its own definition of done (i.e.: an answer to the question: "When are we satisfied with our achievement?") and its own definition of fun: (i.e.: "How to create a learning environment in which we feel happy?").
5	Valuing items			Students discuss in their group the relative weight of all tasks and assignments. A small task is awarded 1 point, a difficult task 5 points. This ceremony provides students with a clear picture of how many points they have to 'burn' during a sprint (see step 8).
6	Sprint planning			All tasks and assignments are written on Post-Its, which are used to visualize students' planning on their Scrum board.
7			Scrum board	A Scrum board comprises three columns 'to do', 'doing' and 'done' and a 'burndown chart'. It provides overview of what tasks should be executed, are in progress or have been finished. The burndown chart shows how many points should be 'burned' during a lesson and reveals if students' progress is satisfactory.
8	Sprint			After the planning phase, teams are ready to start their first sprint, which typically consists of a two-week-period, with approximately six lessons of fifty or sixty minutes.
9	Stand-up			At the start of every lesson all team members gather around their Scrum board for a stand-up ceremony. They discuss (1) what they have done in the last lesson, (2) what they will contribute in the upcoming lesson, and (3) if they experience any problem.

10	Sprint release	Intermediate product	At the end of a sprint, every team releases an intermediate product, which can have several forms (a report, a graphic design).
11	Sprint review	Formative assessment	During the review the quality of the intermediate product or students' conceptual development is checked. An intermediate product, which, in general, has physical properties, is verified if it meets with the objectives. Students' conceptual development is reviewed by using a formative assessment, which sheds light on their understanding of concepts covered during the sprint and reveals misconceptions in an early stage.
12	Sprint retrospective		Teams reflect on the learning strategies they used, assessing and discussing whether the strategies were effective in meeting the learning goals. In addition, they reflect on their motivational and affective state and determine how much social support may be needed to perform a new sprint cycle.
13	Summative assessment or final product		After finishing several sprint cycles, students deliver their final product in which they answer the real-world question. (In this study, each group releases a written advice in which they argue what synthesis route for adipic acid is preferable). In addition, every student performs a summative assessment which focuses on the concepts covered during the context-based approach.

5.4 Method

5.4.1 Context of the study

A context-based approach was implemented in a module on Green Chemistry in secondary chemistry classrooms (grade 11) (Jansen-Ligthelm, Scheffers-Sap, Verhofstad, & Van der Reijt, 2010). Central theme in this module is the need for new sustainable routes to produce innovative chemicals that do not harm the environment, are less hazardous and use renewable resources (Sjöström, Rauch, & Eilks, 2015). Students become aware that Green Chemistry can be seen as the response of the chemical field to issues concerning our planet, (Lozano & Watson, 2013, p. 186) and that impacts people's lives in general and their personal lives in particular. During the course the twelve principles of Green Chemistry (Bodner, 2015) are presented (Table 2). For instance, students calculate the atom efficiency and reaction enthalpy of chemical reactions (resp. number 2 and 6) and discuss the effect of a catalyst as well as the toxicity of reagents (resp. number 9 and 10).

Table 2. Green Chemistry principles.

1. **P**revent wastes	7. **P**ractice with renewable feedstocks
2. **E**fficient use of atoms	8. **L**imited number of reaction steps
3. **O**mit hazardous synthesis	9. **A**uxiliary catalysts
4. **P**roducts are degradable and benign	10. **N**on-toxic precursors and products
5. **L**ow risk and benign solvents	11. **E**valuation of process real time to prevent pollution
6. **E**nergy efficient design	12. **T**o minimize safety risks

The 12 principles of Green Chemistry converge during the final assignment of this module. Teams are invited to compare two different routes for the synthesis of adipic acid, an important precursor to produce nylon-6,6-polyamide. Every team produces a written advice in which they substantiate what route is preferable considering the principles of Green Chemistry.

Ten teachers of secondary schools all over the Netherlands (grade 11) implemented the Green Chemistry module in their classrooms as part of regular chemistry classes. Six of these teachers, from three different schools, used Scrum methodology as scaffold for students' learning. Together with their students they formed the experimental group. The other four teachers did not use Scrum methodology and formed the comparison group. Differences and similarities with regard to different learning phases are presented in Table 3.

Table 3. Differences and similarities between the experimental and comparison group.

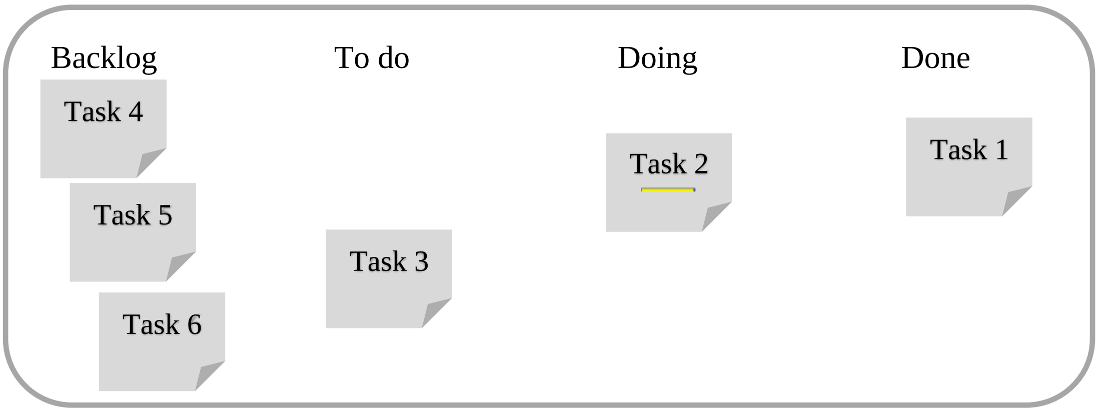
Learning phase	Experimental group (Scrum methodology)	Comparison group
Introduction	The teacher introduced the module on Green Chemistry and shared the ultimate objective of the project.	Similar approach: the teacher introduced the module and shared the ultimate objective.
Planning	Students formed groups using a *special ceremony, including formulating of a definition of done and a definition of fun*. Groups received the module as well as a *product backlog*. In addition, students value all assignments and fill in their Scrum board.	Students formed groups themselves and groups received the module. No special attention on executive regulation functions such as planning and monitoring of progress.
Performing & Evaluating	Students work collaboratively in *sprints* and start every lesson with a *stand-up*. At the end of a sprint, students *review* their conceptual development with a formative assessment and reflect on their learning process with a *retrospective*.	Students work collaboratively and use their own approach. No systematic check moments on conceptual development or quality of learning process. The teacher had a coaching and facilitating role: answered questions and stimulated students to proceed with their learning process.
	Final assessment	Final assessment

Clearly, the main difference between the experimental and comparison group is found in the systematic attention on executive regulation functions, including planning and monitoring of work. Furthermore, within the experimental group a continuous alternation of performing and evaluating is found, whereas the performing phase and evaluating phase are separated in the comparison group.

5.4.2 Implementation of the Green Chemistry module using Scrum methodology

At the start of the module, which lasted six weeks with two or three lessons of 50 or 60 minutes a week, the teacher introduced the central theme of the module as well as its ultimate goal. Thereafter, students entered a comprehensive planning phase, starting with a ceremony in which they formed groups based on their personal qualities (see Table 1). Secondly, the teacher, or in Scrum terms, the product-owner, provided every group with a product backlog in which the learning goals, tasks, and experiments were formulated. However, before the students started to work on the assignment, the teacher asked every group to describe their own *definition of fun* and their own *definition of done*. It seems reasonable to expect that formulating your own *definition of fun* and your own *definition of done* contributes to a learning environment in which students feel both comfortable and responsible. Subsequently, students discussed the relative weight of all assignments. A small and easy assignment was awarded with 1 point, whereas a difficult task was awarded with 5 points. Clearly, this ceremony helped students to plan all tasks and assignments over the time available. They visualised their planning on a Scrum board, which basically consists of three columns (to do, doing, and done) (Figure 1). Hence, they wrote their tasks on Post-Its, and redirected them to the right column when they accomplished a task.

Figure 1. Simplified example of a Scrum board.



The Scrum board provided both the teacher and the students with an overview of students' progress. The planning phase was followed by a sprint in which students performed the assignments and experiments. A sprint took two weeks and consisted of five or six lessons of 60 minutes. At the start of every lesson each group gathered around their Scrum board and had a small stand-up meeting (5 minutes), in which they discussed what they had done in the last lesson, what they intended to do in the current lesson, and whether they experienced problems or not. The first sprint focused on the 12 principles of Green Chemistry and was finished with a formative assessment in which students reviewed their knowledge of the chemistry concepts involved. In addition, they discussed the quality of their collaboration and reflected on how they could improve their learning in the next sprint. In the second sprint, they studied additional concepts such as reaction enthalpies and block diagrams. Afterwards, students' conceptual development was measured again with a formative assessment. In the last sprint, the students transferred the knowledge they had acquired to a new situation. They were asked to develop all necessary tasks themselves, to create a substantiated advice on the greenest synthesis of adipic acid. Clearly, the students were rather autonomous in how they planned and organized their work. Moreover, the role of the teacher was still important, although its focus changed from transmitting knowledge to students to a more facilitating role. For instance, she/he introduced the module and provided Scrum boards, product backlogs and formative assessments, and, furthermore discussed conceptual issues with students on request.

5.4.3 Implementation of Green Chemistry in regular teaching

The four teachers of the comparison group implemented the Green Chemistry course without using Scrum methodology. In their classes, students formed groups themselves. Similar to the way in which the teachers of the experimental group introduced the module, teachers of the comparison group explained the objective of the module.

A typical lesson in the comparison group started with a short explanation (i.e., what is atom efficiency and how to calculate its value?) and afterwards students were allowed to work in groups on the tasks, experiments and assignments as presented in the module. In general, teachers had a facilitating role in this phase of the lesson, which means that they stimulated their students, discussed conceptual issues with groups and answered questions on demand.

However, they did not provide scaffolds or special clues to their students with regard to how they might plan and monitor their work. Furthermore, teachers of the comparison group used neither formative assessments to review students' learning progress nor retrospectives to stimulate their students to reflect on their learning approach.

5.4.4 Participants

Participating teachers and students gave their consent to use their responses for research purposes. Students were informed that their responses were anonymised and therefore had no influence on their grades. All students came from grade 11, and were told that their participation was voluntary. They were informed that they had the opportunity to opt out in any stage of the study. All participating teachers graduated from university with a degree in chemistry as well as chemistry education. Nine teachers had taught chemistry for more than 15 years, whereas one teacher, participating in the experimental group, had 5 years of teaching experience. Research clearance was received by the Ethics committee of Leiden University. In total, 320 students answered the questions of the pre-test and post-test of the Green Chemistry test (GCCT) (54%). In addition, due to time constraints, only 256 students completed both the pre-test and post-test about their affective and metacognitive perceptions (Table 4). The teachers of the experimental group followed a professional development program to get familiar with all the Scrum ceremonies, roles and artefacts (Table 1). They attended five sessions of 4 hours on Scrum methodology in which: (1) the principles underlying Scrum methodology were explained; (2) they practised with its ceremonies, roles and artefacts; (3) shared both their positive and negative experiences with Scrum methodology and (4) provided each other with feedback.

Teachers of the experimental group were split into two groups, based on a scale support by teacher, which was part of the QAM-questionnaire (see Data section below). This scale included three items on how the students evaluated teachers' guidance and instruction during the Green Chemistry module (with 1=completely disagree and 5= completely agree; Cronbach's α= 0.79). Analyses of variance with the six teachers from the experimental group as factor showed that the teachers significantly differed in perceived support ($F(5,189) = 22.731$; $p < .001$; $\eta^2 = .382$). Scheffé post-hoc analyses showed (with $p< 0.05$) that two of the six participating teachers had considerable higher scores. Their average score was 4.49 (with SD= 0.58), whereas two other teachers showed considerably lower scores. Two teachers did not show a clear difference with either group and were - based on their mean scores - added to the low scoring group of teachers. These four teachers with lower scores showed a mean score of 3.70 (with SD= 0.89). The two teachers with high scores were labelled as top teachers and the other four teachers as growth-teachers. Teachers of the comparison groups (from three schools) showed similar scores on support by teacher (with a means of 4.08 and SD of 0.60) and were therefore not split into separate groups. Consequently, the design of this study can be characterised as quasi-experimental with two separate experimental groups and one comparison group.

Table 4. Number of students participating in the study.

		***n*/(female)**	***n*/(female)**
		QAM (pre-test/post-test)	GCCT (pre-test/post-test)
Experimental group	Top-teachers (2 teachers with 2 classes)	100 (54)	109 (57)
	Growth-teachers (4 teachers with 1 class)	90 (36)	109 (46)
Comparison group	(4 teachers with 1 class)	66 (30)	102 (57)
	total	256	320

5.4.5 Data

Two types of data were gathered. First, students' learning outcomes were measured using a pre-test - post-test quiz. Before starting the context-based approach students answered twelve open questions focusing on principles of Green Chemistry. The same test was used as post-test and took place after the lesson series. Second, students completed a *questionnaire* concerning their perceptions of *affective* and *metacognitive* aspects (QAM) of their learning at the start and just after finishing the context-based course.

5.4.5.1 Students' learning outcomes.

Students' learning outcomes were measured with the Green Chemistry concept test (GCCT). The GCCT was based on the concepts covered in the module on Green Chemistry (Jansen-Ligthelm et al., 2010) and consisted of twelve open questions, with scores varying from 1 to 3 points (maximum score = 23 points). The questions assessed students' understanding of the principles of Green Chemistry (e.g.: E-factor, yield, atom-efficiency), as well as calculation of reaction energy using heats of formation. Content and face validity of the GCCT were checked by two secondary chemistry teachers, who were familiar with the module. The GCCT was piloted with 25 students, not participating in the present study. Answers of the pilot were scored independently by two teachers and resulted in an interrater reliability score, Cohen's Kappa > 0.98. Overall scores on pre-test as well as post-test were converted in a percentage, ranging from 0% - 100%.

Examples of questions:

- A manufacturer wants to produce a specific chemical. It turns out that there are two different synthesis routes available. Method 1 has an atom efficiency of 50%, whereas method 2 has an atom efficiency of 75%. Explain which method is preferable (2 points).
- Methyl-tert-butyl ether (MBTE, $C_5H_{12}O$) is added to petrol to increase its anti-knock rating. MTBE is synthesized from methyl propene (C_4H_8) and methanol (CH_3OH). Calculate the E-factor. Assume that the yield of the reaction is 100% (3 points).

5.4.5.2 Affective and metacognitive aspects of students' learning.

Students completed a Likert-type Questionnaire with items on their perceptions of Affective and Metacognitive dimensions of their learning (QAM). The questionnaire also included three items concerning how students perceived the support provided by their teachers (see Participants). The questionnaire was developed for this study. Several sources were used to develop appropriate items. QAM consisted of items on students' attitude about chemistry, inspired by Bennett (2017), other items focused on students' belief of their capabilities (Ajzen, 2002) to understand chemistry, and yet other items concentrated on whether the students perceived that they developed their personal qualities (Perry, Lundie, & Golder, 2019). Items concerning metacognitive dimensions intended to measure how students experienced their mutual collaboration (Schraw, Crippen, & Hartley, 2006, p. 120), and their self-regulation (Panadero, 2017). In addition, QAM consists of items about how students perceived the learning environment in general (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010) and the guidance provided by the teacher in particular (Joseph, 2009). Before QAM was piloted with one class (26 students), not participating in this study, its face validity was checked by two independent teachers, who suggested a few textual improvements.

A Principal Component Analysis (PCA) with Varimax rotation, using SPSS was executed on 34 items of the questionnaire. Three additional questions, focusing on teachers' behaviour, were excluded before executing the PCA, and formed an independent scale. The Kaiser-Meyer-Okin tests revealed KMO-values of respectively 0.929 and 0.941 for the pre-test and post-test. Six components had eigenvalues above the Kaiser's criterion of 1.0, explaining a total variance of 61.9 % (pre-test) and 64.7% (post-test). These components were selected and fitted in two separate clusters, which could be characterized as respectively an affective dimension and a metacognitive dimension of students' learning (Bennett, 2017). All 34 items were included in the 6 components.

Based on the item factor loadings the individual components were labelled as 1) self-efficacy (pre-test: 31.73%; post-test: 35.86%), 2) self-regulation (pre-test: 10.91%; post-test: 11.03%), 3) classroom climate (pre-test: 6.70%; post-test: 6.28%), 4) personal development (pre-test: 5.42%; post-test: 4.80%), 5) attitude towards chemistry (pre-test: 3.64%; post-test: 3.41%), 6) collaboration (pre-test: 3.48%; post-test: 3.33%). The six components found during the Principal Component Analysis are presented in Table 5, including an example item and the reliability in terms of Cronbach's α. All analyses were performed in IBM SPSS Statistics version 25.

Table 5. Scales, their definition and examples of Likert-type questions (score =1: completely disagree; score =5: completely agree).

Component	Definition	Cronbach's α *Number of items* (pre-test)	Cronbach's α *Number of items* (post-test)	Examples of statements used
Affective dimensions	*Students' feelings and beliefs about chemistry and their belief of their own capabilities.*			
Attitude towards chemistry	Beliefs held about school chemistry contributing to students' motivation.	0.85 *(5)*	0.89 *(5)*	I really enjoy chemistry.
Classroom climate	A social, emotional, and intellectual learning environment in which students learn, determined by interacting factors including student-student and teacher-student interactions.	0.80 *(6)*	0.79 *(6)*	The learning atmosphere in chemistry classes is excellent.
Self-efficacy	Students' beliefs in their own capabilities to successfully execute the behaviour to achieve a desired aim.	0.90 *(6)*	0.91 *(6)*	I am convinced that I will master the chemistry concepts involved in this module.
Metacognitive dimension	*Monitoring, control and regulation of learning.*			Examples of statements used
Collaboration	Working jointly with other students in a rather complex learning environment.	0.71 *(4)*	0.74 *(4)*	I assist my team mates during my chemistry class.
Personal Development	Skills and qualities students deploy to achieve the ultimate learning objective.	0.89 *(5)*	0.91 *(5)*	During the chemistry lessons I develop my personal skills.
Self-regulation	Students adjusting their learning behaviour towards the requirements of the specific assignment.	0.86 *(8)*	0.88 *(8)*	I keep myself on track during the chemistry lessons.

5.4.6 Data analysis

Two teacher groups, top teachers and growth teachers, can be understood as two separate conditions. In order to answer the research questions, differences between the experimental group and comparison group will be analysed separately for top teachers and growth teachers. In order to examine the effects on students' achievement (RQ1) two univariate analyses of covariance were performed with teacher group (one with top-teachers vs comparison-group, and one with growth-teachers vs comparison group) as factor, post-test achievement score as dependent variable, and pre-test achievement score as covariate.

In order to examine the effects on students' affective and metacognitive outcomes (RQ2), two multivariate analyses of covariance were performed with teacher group (one with top teachers vs comparison group, and one with growth teachers vs comparison group) as factor, the set of six affective and metacognitive outcomes as dependent variables, and their pre-test scores as covariate. The affective and metacognitive outcomes showed significant, though moderate, correlations (see Table 6). All dependent variables showed a normal distribution.

Table 6. Correlations for the affective and metacognitive outcomes (post-test; *n* = 256; p< .001).

	Self-efficacy	Self-regulation	Classroom climate	Personal Development	Attitude to chemistry	Collaboration
Self-efficacy	1	.269*	.400*	.585*	.689*	.200*
Self-regulation		1	.502*	.375*	.387*	.330*
Classroom climate			1	.560*	.472*	.423*
Personal Development				1	.666*	.252*
Attitude to chemistry					1	.207*
Collaboration						1

5.5 Results

5.5.1 Scrum methodology and students' achievement

Mean scores and standard deviation of the pre- and post-test of the three groups are presented in Table 7. Students from the two top teachers generally achieved significantly higher scores on the post-test compared to students from the comparison group, after controlling for the pre-test scores ($F(1,209) = 56.447$, $p < .001$, $\eta^2 = .213$). This can be understood as a large effect of the intervention (Cohen, 1988). Students from the four growth teachers generally performed significantly better on the post-test compared to students from the comparison group, after controlling for the pre-test scores ($F(1,209) = 18.232$, $p < .001$, $\eta^2 = .081$). This can be understood as a medium effect of the intervention.

Table 7. Mean scores on test (maximum 100 points) of students taught by respectively top-teachers, growth-teachers and students of the comparison group.

		Top teachers (*n* = 109)		**Growth teachers (*n* = 109)**		**Comparison (*n* = 102)**	
		M	*SD*	*M*	*SD*	*M*	*SD*
Green Chemistry	pre-test	25.73	14.0	25.45	15.0	25.19	11.3
	post-test	72.04	16.7	64.86	18.7	55.07	17.7

5.5.2 Scrum and affective and metacognitive aspects of students' learning.

Mean scores and standard deviations on the six dependent variables from QAM are presented in Table 8.

Table 8. Comparison of pre-test and post-test scores of 6 affective and metacognitive components of students' learning for students taught by respectively top-teachers, growth-teachers and students participating in the comparison group (minimum score =1; maximum score =5).

			Top teachers (*n* = 100)		**Growth teachers (*n* = 90)**		**Comparison (*n* = 66)**	
			M	*SD*	*M*	*SD*	*M*	*SD*
1	Self-efficacy	pre-test	3.95	0.84	3.51	0.93	3.57	0.96
		post-test	3.86	0.94	3.74	0.75	3.58	0.92
2	Self-regulation	pre-test	3.43	0.62	3.30	0.67	3,13	0.77
		post-test	3.55	0.60	3.24	0.66	3.17	0.72
3	Classroom climate	pre-test	3.85	0.50	3.17	0.68	3.22	0.53
		post-test	3.92	0.50	3.31	0.62	3.24	0.55
4	Personal Development	pre-test	3.32	0.74	2.82	0.81	2.91	0.75
		post-test	3.42	0.87	2.87	0.80	2.93	0.69
5	Attitude towards chemistry	pre-test	3.64	0.85	3.06	0.77	3.11	0.98
		post-test	3.58	0.89	3.18	0.83	2.98	0.87
6	Collaboration	pre-test	3.28	0.78	3.60	0.79	3.41	0.93
		post-test	3.80	0.73	3.66	0.76	3.44	0.98

Comparison of the student scores of the two top teachers and the comparison group revealed that students from the top teachers showed significantly higher scores on *classroom climate* ($F(1,164) = 17.968$, $p < .001$, $\eta^2 = .099$), *self-regulation* ($F(1,164) = 6.222$, $p < .05$, $\eta^2 = .037$); *attitude towards chemistry* ($F(1,164) = 4.894$, $p < .05$, $\eta^2 = .029$) and *collaboration* ($F(1,164) = 9.895$, $p < .005$, $\eta^2 = .057$), after controlling for pre-test scores on each relevant variable. The difference on classroom climate can be understood as a medium effect size of the intervention; the other three as small effect sizes. No significant differences were found for self-efficacy and personal development.

Comparison of the student scores of the growth teachers and the comparison group after controlling for pre-test scores on each relevant variable revealed that students from the four growth teachers showed significantly higher scores on *self-efficacy* ($F(1,154) = 6.554$, $p < .05$, $\eta^2 = .041$) and *attitude towards chemistry* ($F(1,154) = 7.001$, $p < .01$, $\eta^2 = .044$). These effects of the intervention can be understood as small effects. No significant differences were found for the other four variables.

5.6 Discussion

In this study, the effects are examined of the use of Scrum methodology on students' learning outcomes and perceptions of affective, and metacognitive aspects of students' learning.

5.6.1 Students' learning outcomes

The results on learning outcomes show that students taught by top-teachers as well as students taught by growth-teachers outperform students of the comparison group. This finding suggests that the ceremonies, roles and artefacts of Scrum methodology guide students through, and simultaneously scaffold, their learning process.

Therefore, implementing Scrum methodology in context-based learning environments might respond to criticism that unguided approaches often fail (Kirschner, Sweller, & Clark, 2006). However, providing a unique and straightforward explanation for the observed effect is impossible. Scrum methodology itself comprises a variety of components. Nevertheless, essential in Scrum is that students are invited systematically and explicitly to use, think and rethink the concepts involved in the context-based course. For example, the recurrent stand-up ceremony enforces students explicitly to share what they have done for their team and what problems they encountered. The stand-up ceremony contributes to a learning environment in which students experience that they are mutually interdependent. With this ceremony each team member is brought systematically in a situation in which he or she has to take responsibility for their team effort. Consequently, it ensures that each student engages in the project individually and focuses on its underlying concepts. During the review ceremony students have to give serious thoughts to the concepts they worked on during the sprint. Embedding a formative assessment within a review ceremony evokes both individual reflection and additional discussions about concepts and stimulates mutual peer feedback in a natural way. Educational research has shown that applying formative assessments might contribute to students' achievement (Andrade & Heritage, 2017; Black & Wiliam, 2009; Vogelzang & Admiraal, 2017).

Therefore, the review ceremony might explain the observed effects on learning outcomes. In contrast, students of the comparison group were not exposed to systematically organized moments in which they are challenged to rethink the concepts involved. In conclusion, the review ceremony allows students to gain an understanding of key concepts and helps them to elucidate the learning objectives. Additionally, during the review ceremony, in which students perform a formative assessment, both students and teacher diagnose students' conceptions and might become aware of misconceptions (Treagust, 2012), which, in turn, creates opportunities for teachers to adapt their teaching to students' specific needs. The ceremonies function as focus points and bring students back to the essentials of the project. Students of the comparison group do not have these focus points, which enhances the chance that their learning is superficial. Remarkably, despite the fact that growth-teachers reported organizational issues, and in some cases even some resistance against Scrum methodology, their students still outperformed students of the comparison group, suggesting that the ceremonies and artefacts of Scrum methodology play a key role in the observed, positive effects on students' learning outcomes.

That said, the findings on learning outcomes clearly show that the teachers themselves have a substantial impact. Students taught by top-teachers outperform students of growth-teachers as well as students of the comparison group.

5.6.2 Affective and metacognitive aspects of students' learning

Findings for the effects of the use of Scrum methodology on metacognitive and affective dimensions of students' learning are less pronounced than those on learning gains. In general, the results suggest that the teacher plays a key-role, when Scrum methodology is used to support students' learning. There are several clues that support this claim.

First of all, for *self-regulation* only a small effect-size is found when students taught by top-teachers are compared with students of the comparison group. Although the recurring parts of Scrum methodology, including stand-up and retrospective ceremonies, intend to promote students' self-regulation (see Table 1), no significant effect of teaching with Scrum methodology is found when students taught by growth-teachers are compared with students of the comparison group. Especially noteworthy is that students taught by growth-teachers also did not show an increase in their self-reported perception of their self-regulation. Given the fact that growth-teachers reported resistance towards Scrum ceremonies among their students, whereas top-teachers did not notice resistance, this might be an indication that the way they mentored their students plays a crucial role (Perry et al., 2019).

Findings on *self-efficacy* confirm the importance of teachers' role. In general, high levels of self-efficacy are related positively to students' learning outcomes (Boz, Yerdelen-Damar, Aydemir, & Aydemir, 2016). Given the fact that high effect-sizes on students' learning outcomes were found, comparable effect-sizes on self-efficacy were expected. However, rather unexpected, the self-efficacy of students taught by top-teachers decreased slightly. This might be explained by several reasons. Students taught by top-teachers appreciated their support, and reported a high average score on self-efficacy compared to students of the other groups, suggesting that they did not see any need to change their learning strategies. Leaving behind rather successful learning strategies, used for a long time, and interiorizing a new learning approach might be an inconvenient process and might have a negative impact on students' self-efficacy. What is also striking, students taught by growth-teachers reported an increase in self-efficacy with a small effect-size compared to students of the comparison group, suggesting that, despite their resistance, they benefited from certain ceremonies of Scrum methodology.

The most important clue that the teacher plays a key role can be drawn from students' perception of the *classroom climate*. Students of classes taught by top-teachers appreciated their classroom climate more than students of the comparison group (medium effect-size). Comparing classes taught by growth-teachers and the comparison group revealed no difference. In general, the teacher has a determining role in creating a classroom climate which energizes students' learning, in which students feel comfortable, work hard, feel free to contribute to discussions, develop their skills, and provide and receive feedback etcetera (Ambrose et al., 2010). Apparently, according to their students, the participating top-teachers were able to create such a positive classroom climate. Findings on *personal development* are in line with these results. Compared to students of the comparison group, students taught by top-teachers reported a stronger increase that they developed the personal qualities they deployed during Scrum lessons. Developing personal qualities requires a learning environment in which students receive adequate feedback from their teacher and team mates. Scrum methodology creates these opportunities systematically (stand-up, review, retrospective). Solely implementing Scrum methodology in a context-based classroom does not guarantee that students develop their qualities. Again, the teacher, implementing Scrum methodology, matters.

Findings on students' *attitude towards chemistry* revealed small effect-sizes for both students taught by top-teachers and growth-teachers, compared to students of the comparison group. There is empirical evidence that both context-based approaches and teachers' behaviour are effective in encouraging students to develop more positive attitudes towards science (Lee & Erdogan, 2007). Findings on students' attitude, as presented in this study, are in line with results reported in review study by Savelsbergh et al. (2016). They calculated effect-sizes for context-based approaches intended to stimulate students to reflect on chemistry concepts, and to connect these concepts to their personal lives (Marks & Eilks, 2009). Teachers provide feedback, challenge their students to discuss and thus promote a reflective attitude among them. Apparently, the observed effect-sizes suggest that the Scrum methodology contributes to a positive attitude towards chemistry.

The average scores on *collaboration* of students taught by top-teachers, growth-teachers and students of the comparison group increased substantially during the intervention (Table 8). However, a statistical difference was only found when scores of students taught by top-teachers were compared with students participating in the comparison group. No difference was found between students taught by growth-teachers and students of the comparison group. This result suggests that the ceremonies (such as forming of groups, stand-up) scaffold mutual collaboration in an appropriate classroom climate. Moreover, this result is an indication that the impact of Scrum methodology is profoundly influenced by teachers who are able to create an atmosphere in which students feel comfortable to work and learn.

5.7 Limitations and future directions

Although we distinguished between two qualities of implementation (top-teachers and growth-teachers), the present study did not discriminate between the different components of Scrum methodology. Therefore, it is impossible to explain to what extent the different parts (e.g. the review ceremony, the retrospective or the stand-up ceremony) account for the effects found. A follow-up study might manipulate the various components of Scrum methodology, in particular the review ceremony (with or without a formative assessment). Results on learning outcomes, affective and metacognitive dimensions of students' learning were gathered during the first implementation of Scrum methodology by the participating teachers.

Thus, findings are in part due to the inexperience of the teachers, and their students. If teacher and students are more familiar with Scrum methodology, this might positively impact students' self-efficacy, their self-regulation and their perception that they develop their personal qualities. This requires a longitudinal study of the effect of Scrum methodology on students' learning.

5.8 Conclusions and implications

Although participating teachers and their students were unfamiliar with Scrum methodology at the start of the study, results reveal that it can be a powerful and useful tool in rather complex, context-based learning environments. Obviously, effects of the Scrum methodology on students' learning outcomes depend on the quality of its implementation. Besides, especially in classes taught by top-teachers, positive effects of teaching with Scrum have been found on affective and metacognitive aspects of students' learning such as their self-regulation, classroom climate, personal development and their attitude towards chemistry.

Yet, several participating teachers, both top-teachers and growth-teachers, mentioned that implementing Scrum methodology in a classroom is not an easy job. At the start, it requires more preparatory work, more planning and many other organizational issues, which is common when a new instructional approach is introduced in a classroom (Prince & Felder, 2007). Other factors to be taken into consideration are student resistance and collaboration issues, not to mention that teachers themselves might perceive feelings of uncertainty.

The present study shows that Scrum methodology might strengthen students' learning by providing a structure with clear ceremonies, roles and artefacts that guides students in their learning process. However, as always, the teacher matters. Following a professional development program to become familiar with the tenets of Scrum methodology is not enough. Attention should be given to teachers' behaviour when they implement and, subsequently, how they mentor their students during their lessons, for example, by filming and visiting lessons. Participating teachers as well as the course instructor of the professional development program should act as critical friends to strengthen teachers' mentoring when they guide their students through the Scrum ceremonies. If a teacher is capable to create a classroom climate in which students feel free to reflect, to provide and receive feedback, the classroom becomes a place where students 'learn as they Scrum'.