

PAPER



Cite this: *Chem. Educ. Res. Pract.*, 2020, 21, 1173

Examining the effect of lab instructions on students' critical thinking during a chemical inquiry practical†

Marion E. van Brederode, ^{*ab} Sebastiaan A. Zoon ^a and Martijn Meeter ^b

Developing students' critical thinking skills is often seen as an important educational goal for inquiry assignments. In this study, we investigated to what extent pre-laboratory activities of a chemical inquiry assignment influence students' independent critical thinking. We compared two forms of pre-laboratory activities that are frequently used in educational practice to prepare students for their inquiry assignments: on the one hand paved road pre-laboratory activities that lead students with sensemaking preparatory questions and on the other, critical-thinking pre-laboratory activities in which students start with the development of an experiment plan using provided information and criteria for a good experimental design. We conducted this study two years in succession in senior year Dutch high school chemistry classes during an inquiry assignment that involved the study of the relation between reaction kinetics and molecular reaction mechanisms of organic nucleophilic substitution reactions (S_N1/S_N2). We focused on aspects associated with critical thinking, such as the desire to understand what is observed and to be able to adjust an existing method or model on the basis of experimental data. The results show that the design of pre-laboratory activities strongly influence the critical thinking exhibited by students during their inquiry activities, whereby students who perform critical thinking pre-laboratory activities are more motivated to think more deeply about the meaning of their measurements than students that perform paved road pre-laboratory activities.

Received 19th January 2020,
Accepted 4th June 2020

DOI: 10.1039/d0rp00020e

rsc.li/cerp

Introduction

Research goal

A laboratory is a complex learning environment in which students often have to deal with an overwhelming amount of written and verbal instructions about the functioning of instrumentation, safety, underlying theory, and input from the experiment itself (Johnstone, 1997; Agustian and Seery, 2017; Seery *et al.*, 2019). These instructions affect how students approach their laboratory assignments. It has been noted often that step-by-step instructions, sometimes called cook-book or traditional instructions, that guide students through an experiment have limited learning effects (Kirschner, 1992). These kind of instructions can often lead to situations in which students only start to think about the meaning of the laboratory activity when they are writing their assignment reports. In chemistry education research and practice different approaches have been worked out to prevent this situation (Agustian and Seery, 2017).

In this study, we investigated how the design of pre-laboratory activities elicit pre-university education students' critical thinking during laboratory work. For this, we made a direct comparison of two different, but equally comprehensive, designs of pre-laboratory activities that we implemented in parallel in four chemistry classes of senior-year secondary school students. These we named paved road- and critical thinking pre-laboratory activities. Both forms of activities are based on elements that are often recognizable in educational practice to prepare students for laboratory work.

With the paved road pre-laboratory activity, students were making pre-laboratory questions that guide them through the design of the assignment. These questions focused the students on the research question and how the presented experimental method is well suited to answer the research question.

With the critical thinking pre-laboratory activity, students investigate the same pre-specified research question but they were making their own design for the experiment according to given criteria for a good experimental set-up and provided information (hints).

The way we investigate and define critical thinking is described in the literature section and was inspired by a recent

^a Stedelijk Gymnasium Haarlem, Haarlem, The Netherlands.

E-mail: m.vanbrederode@sghaarlem.nl

^b Vrije Universiteit Amsterdam, LEARN! Research Institute, Amsterdam, The Netherlands

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0rp00020e

study on learning physics students critical thinking during lab work (Holmes *et al.*, 2015) and by the definitions for critical thinking presented in (Facione, 1990). In summary, we focused on aspects of critical thinking that are perceptible as the desire to understand what can be observed in experimental measurement data, the ability to deal with unexpected observations and the ability to adjust one's opinion.

We investigated these aspects of critical thinking by looking at the data analysis in students' research reports. Furthermore, we looked at students' perceptions on their own critical thinking and inquiry work as expressed in answers on a digital questionnaire. Below we first review the literature on the design of laboratory learning and on critical thinking in relation to such laboratory learning and relate this to the investigated educational designs used in our study.

Educational design for laboratory-learning

In the field of chemistry education, the rationale, development, and experience of various new instructional models for laboratory/inquiry learning have been described in many research articles (Domin, 1999; Quattrucci, 2018; Walker *et al.*, 2011; Seery *et al.*, 2019; Farrell *et al.*, 1999; Ditzler and Ricci, 1994; Rodriguez and Towns, 2018). Even though these instructional models set different accents there also appear common elements like the laboratory instructions are not recipe-like instructions, the assignment is more student-centered, cooperative learning is promoted, and the preparatory phase plays an important role.

Logically, a preparatory phase should support the final educational goal of the assignment. In chemistry education, these goals tend to shift from learning specific subject content and using laboratory equipment to skills like experimental design, scientific argumentation, connecting science with social contexts and the development of critical thinking skills. (Agustian and Seery, 2017) reviewed 60 research studies on pre-laboratory activities that have been designed to support students learning in the chemistry laboratory. They distinguish three rationales for pre-laboratory activities: guiding the understanding of chemical concepts, the learning of laboratory techniques, and addressing affective dimensions to enhance learner confidence and motivation for laboratory work.

New instructional designs are often compared to a situation in which there was a less advanced preparatory phase. Such studies virtually all report a noticeable increase in students' appreciation and involvement (Ditzler and Ricci, 1994; Farrell *et al.*, 1999; Chase *et al.*, 2013) (Keys *et al.*, 1999; Cooper and Kerns, 2006; Chase *et al.*, 2013; Mistry *et al.*, 2016; Chase *et al.*, 2017; Agustian and Seery, 2017; George-Williams *et al.*, 2018; Seery *et al.*, 2019).

In addition, the quality of students reports, according to rubrics that evaluate the level of scientific argumentation and presentation, have been reported to increase with instructional designs like Argument-Driven-Inquiry (Walker and Sampson, 2013) and Science Writing Heuristics (Gupta *et al.*, 2015; Oliver-Hoyo, 2003).

It has also been investigated that the openness of inquiry assignments (Fay *et al.*, 2007) has a strong effect on students' conversations and actions in the laboratory environment (Xu and Talanquer, 2013b). These appear to shift with increasing openness from task-oriented, *i.e.* focused on managing and completing lab work towards discussions in which ideas about the inquiry are expressed. Higher levels of inquiry also favoured episodes in which experimental work was approached in a more exploratory (*versus* procedural) manner. These observations could also be found in the written reflections in the student reports which shifted from focussing on factual knowledge to procedural evaluations and metacognitive knowledge (Xu and Talanquer, 2013a).

A frequently used method in educational practice is to guide the thinking and acting of students during their inquiry work with laboratory questions. This is the approach we also followed in the design of the paved road pre-laboratory activity. A recent article describes the criteria that these kind of -laboratory questions must meet to promote engagement in critical thinking (Rodriguez and Towns, 2018). These authors recognized that with traditional pre-laboratory questions students are often asked to do a calculation or answer a question without assigning any meaning to the answer which did not involve engagement in science. In their modified form of the pre-laboratory questions, students were directly asked to think about the meaning of certain aspects of the investigation like, "...explain what method we are using in this laboratory experiment..." and "...explain why these methods are used and how the data generated from these methods answer the questions above...".

Another way to promote students' investigative attitude during laboratory work is to have students start by making their own experimental plan and discuss this with others, (see for example Cooper and Kerns, 2006; Walker *et al.*, 2011). This is also present in a study on critical thinking during physics students' laboratory work. (Holmes *et al.*, 2015). This educational design in this study is based on the idea that the critical thinking of scientists involves an iterative process of repeated comparisons and decisions, *i.e.* comparing new data to existing data and/or physical models and then deciding how to act on those comparisons by improving the experiments and/or adjusting applied physical models. By explicitly instructing students to mimic this process by making their own comparisons and decisions about their experimental methods and data-analyses, students get used to do these actions independently and think critically during laboratory work. This has been tested in a large-scale study by slowly fading away the explicit instructions during a laboratory course. It appeared that at the end of the laboratory course, when there were no longer explicit instructions for making comparisons, students performed much better in terms of critical thinking (see next section) than students from previous years and these improvements persisted into a subsequent course taken the following year.

We have implemented aspects of this instructional design in the critical thinking pre-laboratory activities.

Critical thinking and laboratory learning

In previous chemistry education studies on critical thinking, the emphasis was put on the level of scientific argumentation and the general quality of research products (Oliver-Hoyo, 2003; Gupta *et al.*, 2015).

However, when critical thinking is included as a learning objective for inquiry learning, it is also essential that students by using analytical skills recognize aspects of experimental data that are important for their conclusions, but that they may not have initially considered in the systematic design of their research. This open and independent investigative attitude is an element of critical thinking characterized by a desire to understand what is observed and the ability to adjust an existing method or a vision (Facione, 1990). This is a skill that in retrospect often appears to be of crucial importance for scientific developments, where crucial moments are often initiated by noticing something peculiar in experimental data that is then further investigated. (Hofstein and Lunetta, 2004). In this way, not waving away, but trying to explain or investigate, unexpected aspects in experimental data is a clear indication of critical thinking.

To recognize these aspects of critical thinking in students' laboratory work Holmes *et al.* developed a rubric in which the level of critical thinking was investigated in terms of students identifying and correcting for disagreements between data and models and students acting on these disagreements by further investigations (Holmes *et al.*, 2015). This aspect of critical thinking is further defined as the process by which one decides what to believe of experimental data, evidence and models (Holmes *et al.*, 2017). The anchor points in the rubric used in this study to investigate critical thinking were based on the rubric used in (Holmes *et al.*, 2015).

Current study

Here, we investigated the effect of the two forms of pre-laboratory activities described above on critical thinking. We named our two conditions the paved road condition and the critical thinking condition. After our study was done, the design of the pre-laboratory questions in our paved road condition appeared to be similar to the modified pre-laboratory questions of Rodriguez and Towns, 2018 in that the questions focus the students on thinking about the goal of the inquiry and the experimental method. The pre-laboratory activities in our critical thinking condition were based on the work of Holmes and coworkers in that the students were guided to make their own well-considered decisions to address a pre-specified research goal experimentally. A detailed description of our paved road and critical thinking pre-laboratory activities are given below and in Appendix 1 and the ESI.†

The requirements for data analysis, follow-up experiments, and the report students wrote at the end (which we used to measure critical thinking) were identical for both conditions. In this way, we could assess the effect of the design of the preparatory activity on the students' critical thinking, while a possible influence of specific guidelines for writing the report and/or data analysis on the student work was excluded.

Experimental setting

Investigated student groups

The study was performed on senior-year secondary school students taking chemistry classes (age 17/18) within one pre-university school in the Netherlands (VWO – the highest track in Dutch secondary education), which has consistently been in the top 5% of pre-university education on national exams.

These students have already performed several biology, chemistry, and physics inquiry projects (with different levels of guidance and openness) during their school career. The students were therefore, at least to some degree, experienced and capable to perform measurements and to incorporate experimental data into inquiry reports.

The first two authors were chemistry teachers of all students included. The students were aspecific and randomly assigned at the start of each year to chemistry classes without input from the teachers. Both cohorts consisted of four chemistry classes, two of which were assigned to the paved road condition and two to the critical thinking condition. For each teacher, one group was assigned to the paved road condition and one group to the critical thinking condition. Classes were divided over the two conditions by matching average chemistry grades in the preceding year, which resulted in equivalent average grades for each condition.

The work of in total 86 (cohort 1) and 80 students (cohort 2) Dutch pre-university students (age 17/18) was investigated. The groups' size varied between 20 and 24 students. Students worked in pairs so that the sample size for lab data is half the number of students.

At the end of the school year we could compare all final exam scores (combined school exams and national exams). These results confirm that the average chemistry level of the students in the two conditions was equivalent (and remained equivalent) [cohort 2: average scores and standard deviation critical thinking-condition 7.2 ± 1.3 and paved road condition 7.1 ± 1.0 ; on a 1–10 scale, $t(79) = 0.34$ $p = 0.74$]. For the cohort 1, the four groups had identical average scores at the national chemistry exam (7.0).

The research was conducted in line with the ethical guidelines of the Faculty of Behavioral and Movement Sciences of Vrije Universiteit Amsterdam, which at the time did not mandate informed consent for educational research. In the study, students performed a lab assignment which was part of the normal curriculum in the school, and a survey on that assignment – which was not. Participation in the survey was voluntary. Students were invited by email and informed about the goal of the survey, and that participation was voluntary and anonymous. Care was taken to formulate questions in such a way that students could not be identified from their answers. Student work was collected in the course of normal classwork and was anonymized before being coded into a research database.

Inquiry experiment

The students performed a project that involved the study of reaction kinetics of organic chemistry reactions to investigate the molecular reaction mechanism (S_N1/S_N2). The project

started three weeks after the summer break. The total duration of the inquiry assignment was 9 lessons of 45 minutes. During the first three school weeks, all groups followed the same program, in which students had worked through several chapter sections on chemical reaction mechanisms learning about displaying the transfer of electrons in molecular reaction mechanisms, the relation between rate-determining steps and energy diagrams/observed kinetics, different kind of reaction mechanism (ionic and radical mechanisms) and we had worked out S_N1 and S_N2 mechanisms.

For all experiments in the project, we applied the paved road and critical thinking preparatory activities as described in the introduction and below. To investigate the level of critical thinking in the students' report we focused on an experiment that allowed us to identify the level of critical thinking as presented in their research report particularly well.

The goal of this experiment was to distinguish between two different reaction mechanisms (S_N1 or S_N2) for the substitution reaction shown in Fig. 1.

Based on Fig. 1 and on what they had learned, the students argue that an increase of $[OH^-]$ would either increase the reaction rate (S_N2 mechanism) or would have no effect (S_N1 mechanism), depending on whether OH^- is involved in the slowest, rate-determining step. An increase in the concentration 2-chloro-2-methylpropane (2Cl2MP) would increase the reaction rate in both reaction mechanisms.

An important issue that the students in both conditions seldom noticed during the preparatory activity can be explained as follows. The students followed the progress of the reaction by recording the time for a pH indicator to change color. In this way, when varying $[2Cl2MP]$ a change in the (relative) reaction rate is reflected in a change in the reciprocal of the reaction time. However, in the case $[OH^-]$ is varied the substance that is varied is the same as the substance that is recorded.

When the students performed the experiment, they observed that by increasing the (excess) concentration of 2-chloro-2-methylpropane, the reaction time indeed decreased, which was interpreted as an increased reaction rate. In contrast to this, when the students increased the concentration of OH^- in their experiments they observed that the reaction time increased (suggesting to them a decrease in the reaction rate).

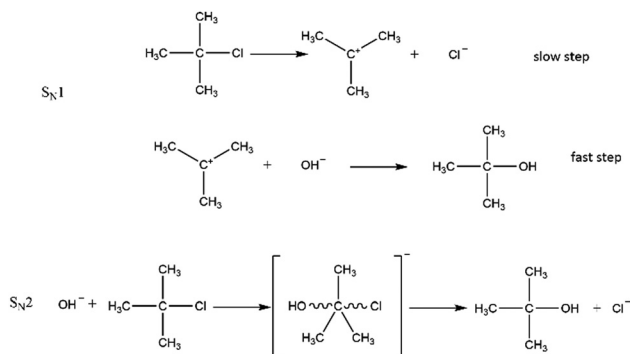


Fig. 1 S_N1 and S_N2 reaction mechanisms of the nucleophilic substitution reaction of 2-chloro-2-methylpropane with hydroxide ions.

This was an effect that was not predicted beforehand in both conditions, even though we did give some hint on this by writing in all student instructions: “Remember that a measure of the reaction rate is inversely proportional to the time it takes to convert a certain amount of substance”. However, no further emphasis was placed on this by the teachers. Probably the students were biased because in previous typical school experiments in which comparisons of relative reaction rates were made, this was done by looking at the appearance of product formation (*i.e.* formation of precipitations/gas development, color changes *etc.*), but it is uncommon that the substance for which the change is recorded is also the (starting) substance of which the concentration is varied. For the experiment in which $[2Cl2MP]$ was varied the same trend was observed when displaying the reciprocal of the reaction time as when displaying the calculated reaction rate, because the $[OH^-]$ was constant in that experiment.

That even university students can easily overlook such issues in experimental data on chemical kinetics has also recently been shown. (Rodriguez *et al.*, 2019).

It is important to mention that we could easily have organized the instructions in such a way that almost all students analyse the data in the correct way by simply giving them more prompts how to analyse the data in the analysis section. However, for this study we wanted the students to encounter something during the experiment they had not expected beforehand (*i.e.* the reaction time increases upon higher concentration of OH^-), but where it would be relatively easy for students who thought critically about the meaning of their data to notice this ‘hidden trap’ and propose a correct way of analysing the data (*i.e.* by dividing $[OH^-]/\text{reaction time}$). However, students who are not concerned about the meaning of their experimental data either fail to notice the apparent discrepancy between their experimental data and the models they had learned about or are not bothered by it. This can be interpreted as an indication of a lack of critical thinking during the inquiry assignment.

This hidden trap of the relation between reaction time, chemical rate and molecular concentrations was therefore used to obtain marker points for critical thinking aspects as “the desire to understand what is observed and to be able to adjust a method or model on the basis of experimental data” (Facione, 1990; Holmes *et al.*, 2015).

Paved road and critical thinking pre-laboratory activities

A translation of the investigated assignment in the paved road and critical thinking condition is given in the ESI.†

Students in both conditions received the same number of lessons, the same theoretical subject information, and the hints for the design of the experiment in the critical thinking condition were given to the students in the paved road condition as information for answering the pre-lab questions.

Therefore, the major difference between the paved road pre-laboratory activity and the critical thinking pre-laboratory activity was the way how the information was processed, *i.e.* by answering pre-laboratory questions or by designing an

experimental plan, respectively. Besides this, students in the paved road condition were also provided with compact laboratory instructions on how to perform the experiment. However, this difference was marginal since the students in the critical thinking condition also did get a hint on a starting combination of substance concentrations for which the experiment should work. The instructions to finish the project after completion of the experiment were identical for both conditions and no hint was given to the students about how to process the experimental data in this evaluation phase.

After the students had performed their experiments, students in both conditions were encouraged by the teachers and written instructions to consider if it was necessary to do additional measurements to answer the research question. The instructions for completing the project were identical for both groups; thus, the paved road condition also had the possibility to improve/repeat their experiment or adjust their model after the execution of the first experiment. Both groups received the same guidelines for writing the report. These instructions were not specific and identical for both the paved road and critical thinking conditions. For example, '*analyze the results and display all results in tables and graphs*', '*indicate to what extent you have been able to answer the research question*'.

Difference between cohort 1 and cohort 2

We conducted this study for two years in succession. Based on our findings in the study with cohort 1, the designs of both the paved road and critical thinking condition has been slightly adjusted for cohort 2 (see Appendix 1 (ESI[†]) information for a detailed description of the differences).

In summary, the cohort 2 students in the paved road conditions were more focused on the research goal and the cohort 2 students in the critical thinking condition were given more direction for the set-up of the experimental plan, relative to the students in cohort 1. For both cohorts we paid close attention that both conditions received the same amount of hints and the extra direction/hints given in cohort 2 for one condition was also presented to the other condition. The laboratory conditions were also optimized for cohort 2. The results of the studies on both cohorts, both on four student groups (two for each condition), are presented in this article.

Defining level of critical thinking

In Table 1 a rubric of four levels is depicted that was used to rank the different levels of critical thinking in the students

reports. This idea for this rubric was originally derived from the rubrics of (Holmes *et al.*, 2015).

The rating was done by the first two authors individually, with all reports rated independently by both to establish the reliability of the ratings.

Their ratings were the same in 85% of cases, with a difference of just one level for the remaining cases. When there was doubt after discussion of the differences, for the paved road condition the more positive ranking was granted and for the critical thinking condition the more negative ranking. The presented results therefore provide a conservative estimate on the difference between conditions. Examples of student answers at the four different levels are provided in Appendix 1 (ESI[†]). It is important to note that critical thinking as assessed by this tool is sensitive to instructions and hints that students receive. Holmes *et al.* (2015) used it to follow the educational process in a course in which students were explicitly instructed to make comparisons, and to investigate learning at the end of the course when explicit instructions had faded away. Any effect of instructions on critical thinking with this tool can only be interpreted in settings with control groups that perform similar experiments but with slightly different instructions.

Digital questionnaire

We invited all students *via* school email in the cohort 2, approximately 1 week after they had submitted their reports (but before they had received their grades), to fill out an anonymous digital questionnaire to reflect on their experience during the inquiry assignment. The difference between responses given by students in the two conditions was analyzed using a series of two-sided independent *t*-tests. Using *t*-tests presupposes a ratio or interval scale, while our Likert scales are arguably at an ordinal level. We therefore repeated the analyses using Mann-Whitney *U* tests that are appropriate for ordinal scales. As the analyses of the questionnaire were explorative, we did not correct for multiple testing. The magnitude of any difference between the two conditions was expressed by using Cohen's effect size, *d*, which indicates by how many standard deviations the averages of the two measurements are separated (see Appendix 1, ESI[†]). Descriptors for the magnitude of *d* are that the under limit for a small, medium and large and very large effect size are *d* = 0.2; *d* = 0.5; *d* = 0.8 and *d* = 1.2, respectively (Cohen, 1992).

Table 1 Assessment rubric critical thinking level. Checkpoints for identification critical thinking level in students' reports

Level 0	No thoughtful comparison between data and models.
Level 1	A comparison between experimental data and kinetic models without realizing the real meaning of the data. <i>E.g.</i> , interpretation of increased reaction time as a decrease of the reaction rate and the notion that a decreased reaction rate upon increased [OH ⁻] does not match with any of the two suggested kinetic models.
Level 2	Interpretation of the experimental data and determination of the extent to which the data correspond to the theoretical model/hypothesis. <i>E.g.</i> , the notion that the trend in the (reciprocal of the) reaction time is not a direct measure of the reaction rate since more OH ⁻ has to react before the pH indicator changes color.
Level 3	Quantitative interpretation of the experimental data, and of the extent to which the data correspond to the theoretical model/hypothesis. <i>E.g.</i> , a quantitative approach of the experimental data by which the dependence of the reaction rate on [OH ⁻] can be evaluated.

Results

Level of critical thinking

Fig. 2 presents the students' critical thinking level identified by specific marker points in their reports as depicted in Table 1. These critical thinking levels represent the extent to which the students had thought through about the meaning of their measurement. Examples of students' answers at the different levels are provided in Appendix 1 (ESI[†]).

The results are presented separately for cohort 1 and cohort 2, since the students' instructions were slightly adjusted for cohort 2 (see experimental setting and Appendix 1, ESI[†]).

To test whether the level of critical thinking was different between our two conditions, we performed a nonparametric Mann–Whitney *U* test, fitting for the ordinal measurement level of critical thinking. Over the two cohorts, students in the critical thinking condition reached a higher level than that of the paved road condition (Mann–Whitney *U* test: $p = 0.006$).

In the cohort 1 study, 57% of the students in the paved road condition did not exceed level 0. Some of these students wrote surprisingly haphazard statements in their reports. At this level, students often just stated that the reaction rate depends on the concentration of both reactants and did not elaborate on how this relates to the kinetic models. Students' work was also classified at this level when the students indicated that they saw a trend in the data that was clearly not there or when they ignored the experimental data when writing the discussion. This outcome contrasts with the answers to the pre-lab questions of these students from which it became clear that most of

them had thought about the expected relation between the reaction rate and the concentration of the reactants in the two molecular models and understood this. Apparently, by the time students in the paved road condition were writing the report, they had forgotten about it, or they did not understand (or care) in practice which aspects from the experimental data were important for making this comparison.

To avoid a similar situation, we sharpened the assignment instructions of the paved road for the cohort 2 by making the research goal more explicit by mentioning that it will be checked experimentally whether the reaction rate is dependent on both concentrations and to use this information to experimentally distinguish between the S_N1 and S_N2 mechanism (and the same text was added to the cohort 2 critical thinking assignment).

This resulted in student pairs in the paved road condition (and critical thinking condition) generally receiving higher critical thinking ratings than those in cohort 1. In cohort 2, only 5% of the paved road pairs were classified at level 0 and the largest fraction was classified at level 1.

The students in level 1 interpret the increase of the reaction time upon increasing $[OH^-]$ as a decreased reaction rate and note that this does not match with the expected effect of any of the two investigated kinetic mechanisms. After this notice, some of the following remarks can typically be read: 'the obtained results are thus unreliable'; 'since the reaction rate appears to be decreasing (it is thus certainly not increasing) and therefore the S_N1 mechanism fits best to the experimental data since with the S_N2 mechanism the reaction rate would increase upon increasing $[OH^-]$ and that is certainly not happening'.

For the critical thinking condition, it was already rare in the cohort 1 study that a student pair's work was classified at level 0. In the cohort 1 study 36% of the students in the critical thinking condition were classified at level 2, for only 5% of the paved road condition. When rated at level 2 students did note that the trend in the (reciprocal of the) reaction time is not a direct measure of the reaction rate, since upon increasing $[OH^-]$ the reaction must proceed longer before the pH indicator changes color. This was then followed by remarks such as: 'the reaction rate thus does not decrease but (probably) has remained constant (without checking for this by doing a calculation on the data) or the statement that they cannot identify the effect of $[OH^-]$ on the reaction rate since they think that in this experiment two things are varied at the same time'.

At level 3 students first reach the same conclusions as at level 2, and then think about how they can quantitatively correct for this. Usually, they reach the solution of dividing the concentration of $[OH^-]$ by the reaction time of the pH indicator to change color. In the cohort 2, more critical thinking condition pairs were rated at level 3 than level 2. Some student pairs at level 3 noticed in their experimental data that even this average reaction rate decreases slightly with increasing $[OH^-]$ concentration, and a few even realized that this could be explained by the decreasing concentration of the excess 2Cl2MP during the course of reaction (because it is not a perfect initial rate experimental setting in which the decline of the 2Cl2MP is negligible), and the concentration of 2Cl2MP

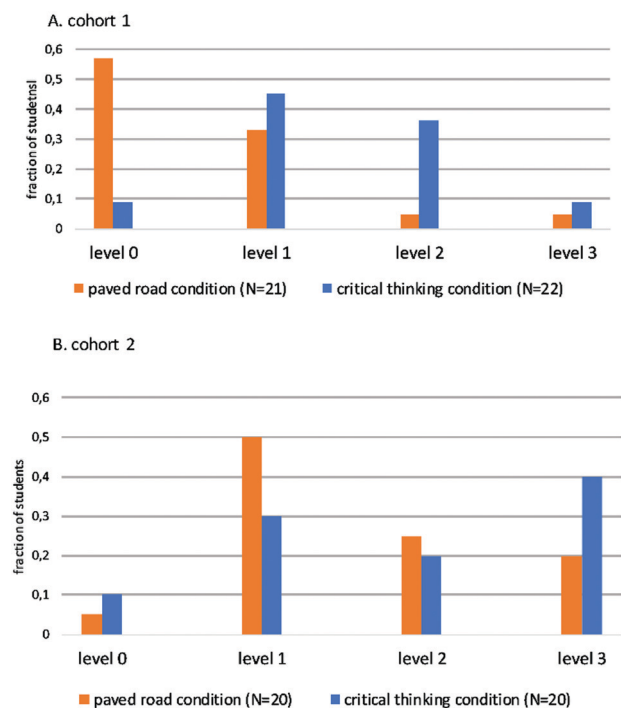


Fig. 2 Ranking of the critical thinking of students pairs, according to the assessment rubric presented in Table 1, split out for the cohorts and conditions as described in the text.

does affect the rate of the substitution reaction rate in both kinetic models. In both the cohort 1 and cohort 2 study the fraction of students that reached level 3 was smaller for the paved road condition than for the critical thinking condition.

The shift from students who are assigned level 2 to level 3 in the critical thinking condition between cohort 1 and cohort 2 is potentially related to the subtle modification of the written instructions for cohort 2, by which the students were guided more in the direction of varying concentrations of starting substances with the design of the experimental plan. In this respect, it should be noted that in the cohort 1 critical thinking conditions, in which the students received less direction in coming up with the experimental method, some student pairs opted for a different experimental approach. Sometimes these students were clearly highly motivated to understand the meaning of their measurement data and expressed a lot of ‘original’ thinking over their measurement data in their experimental reports but did not succeed to interpret the effect of $[\text{OH}^-]$ as depicted in Table 1 correctly. Initially, we ranked this kind of “original” argumentation as a higher level of critical thinking, similar to the rubric in Holmes *et al.*, 2015 “*Synthesis level: a written reflection statement that synthesizes multiple ideas, tool analyses, or reflections to propose a new idea*”. However, we finally chose to uniformly judge the level of critical thinking with the use of strict marker points as depicted in Table 1, since this enabled us to identify consistently the levels of critical thinking. This more strict approach caused that the ‘original’ work of some student pairs (mainly from the critical thinking

condition in the cohort 1 study) was classified at a lower level than when we used a less strict rubric, which makes that the differences between the students’ critical thinking in the two conditions slightly decreased (in cohort 1 we ranked 18% at level 3 with the original rubric).

Digital questionnaire

Fifty-nine students in the cohort 2 (74% response rate) filled out the anonymous digital questionnaire to reflect on their experience during the inquiry assignment (see Fig. 3 and Table 2). In the questionnaire, the students were asked to rate their own involvement with the assignment, their judgment of their level of preparation for the experiment, the level/content/usefulness of the assignment, and statements that can be associated with the level of critical thinking during the performance of the experiments. The statements for which average scores both showed a significant difference ($p < 0.05$) and at least a medium effect size ($d > 0.5$) are highlighted in italic. All of these questions also showed significant differences when analyzed with Mann-Whitney U tests except the last one (statement 17), where the difference between the two conditions just failed to reach the threshold for significance ($p = 0.065$), though it did reach significant with the t -test ($p = 0.038$).

No significant differences were seen for statements that refer to the content, level or usefulness of the assignment (statements 4, 8, 9 and 16). Statements 1–3 touch on topics on how well the students felt prepared for the assignment. On statement 1 (how to execute the experiment) and statement 2

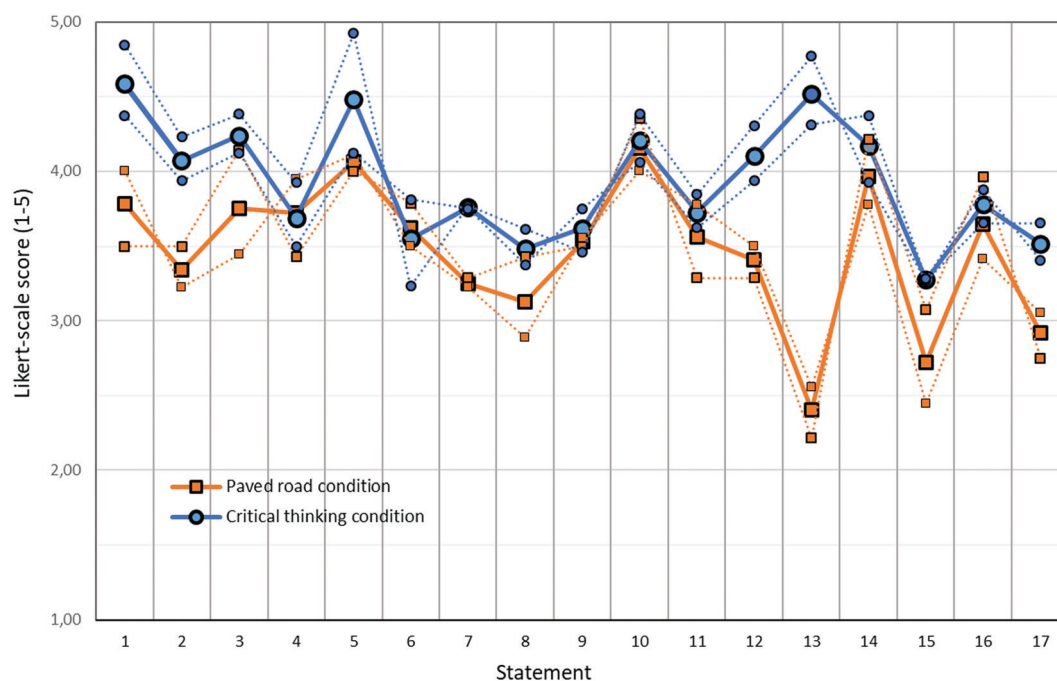


Fig. 3 Average scores on 1–5 Likert scale for statements 1–17 (see Table 2 for the statements) for students in the critical thinking condition (solid blue line with large blue circles, $N = 28$, response rate 70%) and in the paved road condition (solid red line with large red squares, $N = 31$, response rate = 78%). Statistical evaluations of the observed differences between the two conditions are depicted in Table 2. Thin lines and small symbols represent the average scores of the two chemistry classes within each student condition. Statement 15, 16 and 17 are evaluated on a 1–10 scale; these responses were divided by 2 to match the scale of the other questions.

Table 2 Statements and statistical evaluation of differences between the students' responses in the critical thinking- and paved road condition depicted in Fig. 3. Probability (p) value and Cohen's effect size (d) of the difference between the Likert-scale scores in the critical thinking and paved road condition. The statements for which average scores both showed a significant difference ($p < 0.05$) and at least a small effect size ($d > 0.2$) are italicized

Statement	p	d	Statement	p	d
<i>1. In this project, prior to the execution of the experiment, I thought about how I could conduct the experiment.</i>	4×10^{-4}	0.94	10. In this project, I searched for reasons and causes if the results were different than expected.	0.83	0.05
<i>2. In this project, I thought about how I could analyze the measurement data before the experiment was carried out.</i>	3×10^{-3}	0.79	11. In this project, I have reviewed and adjusted my vision on the outcomes of the experiment.	0.54	0.15
3. During this experiment, I was aware of the research question we were investigating during the experiment.	0.08	0.44	<i>12. With this assignment, I tried to understand the results and to value the meaning of the results.</i>	0.04	0.76
4. In this project, we investigated to what extent our measurement data were corresponding to a theoretical model.	0.93	-0.02	<i>13. In this project, we determined ourselves how to conduct the experiment.</i>	$< 10^{-5}$	1.92
5. In this project, I have consulted well with my partner about the significance of our measurement data.	0.14	0.38	14. In this project, I learned a lot about subject content	0.40	0.21
6. During the execution of the experiment, I was curious about the results.	0.77	-0.07	<i>15. I enjoyed making the pre-lab questions/ the experimental plan.</i>	0.02	0.59
7. During the execution of the experiment, I gained new insights into the measurement method we were using.	0.05	0.50	16. I think to make the pre-lab questions/ the experimental plan was useful.	0.56	0.15
8. Finishing this project gave me a sense of personal achievement.	0.22	0.31	<i>17. During this project, I learned a lot by discussing with my project partner.</i>	0.03	0.53
9. This project was challenging.	0.69	0.10			

(thinking beforehand about data-analysis) students within the critical thinking condition gave significantly higher ratings than those in the paved road condition. For statement 3, (being aware of the research question during the execution of the experiments) the difference between the two conditions did not reach significance ($p = 0.08$), but it is interesting to note that percentage of students that indicated that they were not aware of the research question while doing the experiment (Likert-scale response of 1 or 2) did differ significantly between the paved road and critical thinking condition. In both groups of the paved road condition, a fairly large portion of the students gave such a response (21% and 35%), whereas no student in both groups in the critical thinking condition gave a score lower than a neutral score of 3. Statements that address to students' critical thinking level are statements 7, 10, 11 and 12. Statements 7 (gaining new insights), 10 (adjustment of one's opinion) and 11 (search for reasons and causes of discrepancies) showed no significant differences between the two conditions. However, when it comes to estimating the value of measurement data and understanding the experimental results (statement 12), students in the critical thinking condition gave significantly higher ratings than the paved road condition. Apparently, students in both conditions had a similar experience of how their understanding developed during the inquiry (statement 7, 10, 11). Only, when it comes to estimating the value and the meaning of the measurement data (statement 12), students in the critical thinking condition felt that this applied more than students in the paved road condition did. Apparently, review and adjustments of one's vision (statement 11) and looking for reasons and causes when things are different as expected (statement 10) might also be something that can occur in other ways than by carefully looking at data (statement 12), for example by noting that other students interpret their results in a different way and then adopting their point of view.

Students in the critical thinking condition experienced a strong autonomy in determining how they carried out the experiment, which was generally not the case in the paved road condition (statement 13). Also, students in the critical thinking condition reported more enjoyment during the pre-laboratory activity (statement 15) than students in the paved road conditions and reported more learning from discussions with one's project partner (statement 17).

The course of the lessons as experienced by the teachers

In general, students in both conditions seemed to appreciate the assignment and to enjoy doing the experiments. The students in the critical thinking condition were convinced that they first had to prepare an experimental plan before they could implement it. We also looked at the reflections students wrote at the end of their reports. Students indicated that they appreciated that they were given time to sort things out themselves together with their classmates. Students took the making of the experiment plan seriously.

In the paved road condition, the teachers sometimes had to pull harder to get all students to work during the preparatory activities. After the students had performed their first experiments, it became clear that due to the condition of conducting the experiment according to a prescription, the students in the paved road condition were not inclined to repeat or alter the experiment. They argued with the teacher that they had all the measurement data since they had reached the end of the prescription. Since that went fine in their opinion, it should suffice for writing the report. They were focused more on finishing the project than on the interpretation of the meaning of their results and thinking about further experimentation. In the cohort 1 study, 43% of the students in the paved road condition presented more than one dataset compared to 77% for the critical thinking condition. In the cohort 2 study, these

percentages were 10% and 70% for the paved road and critical thinking condition respectively.

The first author (MEvB) furthermore counted whether the reflections of students mentioned specific points like collaboration with project partner, thinking/discussing about the experiment, adjusting the experiment, autonomy/planning of the process, proud/satisfaction about the process/product, externalizing setbacks, gaining subject knowledge, thanking lecturers, and developing research skills. For the cohort 1, with the most open approach in the critical thinking condition, a significantly larger fraction ($p < 0.05$) of the students in the critical thinking condition mentioned that they were very proud of their achievement, that they had discussed/been thinking a lot about the experiments and that they practiced their inquiry skills. For some other statements, differences were also present, always in favor of the critical thinking condition, but these were not statistically significant.

Discussion and conclusion

In this study, we investigated two years in succession the effect of pre-laboratory activities on students' scientific reasoning and acting during their inquiry assignment. We focused on aspects associated with critical thinking, such as the desire to understand what is observed and to be able to adjust an existing method or a vision. We compared a paved road pre-laboratory activity that prepared students thinking about the experiment with preparatory questions, to a critical-thinking pre-laboratory activity in which students were instructed to develop an experimental plan for a specific research goal.

In both years students that had worked with critical thinking pre-laboratory activities reached a much higher level of critical thinking in their reports than the students that had worked with paved road pre-laboratory activities. Students that followed paved road pre-laboratory activities more often did not notice clear discrepancies between model and data or did not take those discrepancies seriously and because of this could not reach a valid interpretation of their measurements. Such deficient reasoning about data was much less common for students that had performed critical thinking pre-laboratory activities.

For cohort 2, in the paved road condition we made it more explicit what data should be used to answer the research question and perhaps as a result we observed that more students interpreted their measurement data with more depth. This may be interpreted as that the level of critical thinking reached in inquiry learning is largely a function of the direction given in the guidance. However, this is probably not the case and these (our) teacher actions are probably exemplary of what often occurs in research assignments with paved road instructions. Once teachers note that students miss a specific point in their inquiry task, the instructions will be modified the following year in a form that encourages the students not to overlook this specific point. This might be promoted by experiences of teachers, as also recently emerged from a meta-analysis of studies on guided inquiry instructions that more specific

guidance during inquiry assignment results in higher quality learning products than inquiry learning that is guided by less specific instructions (Lazonder and Harmsen, 2016). This is unsurprising as with more specific guidance the students are better informed on what is expected from them or what they are assessed on.

In this study we observed that with the same information as in the paved road condition, critical thinking pre-laboratory activities, with instructions to design, compare and make decisions, elicit more critical thinking skills among the students. This became evident in the subsequent laboratory work in which students' in the critical thinking condition more often noticed an unexpected observation and tried to understand and clarify it instead of waving it away.

According to students' self-evaluation, the students in the critical condition value their own experimental data more and thought with more depth about their measurements' meaning. The critical thinking instructions also resulted in an experienced better collaboration in- and between student pairs. A general factor in the higher appreciation for the assignment may well be that in the critical thinking instructional model the three key precursors of motivation in self-determination theory, namely autonomy, competence and relation (Deci *et al.*, 1991), seem to be present to a higher degree. Probably this results in students and teachers having positive experiences during inquiry activities.

It is important to note that in neither condition there was explicit critical thinking skills training. The differences we observed are therefore unlikely to be the result of students learning to think critically. Instead, they result from the extent to which the instructions elicit critical thinking skills that students had already acquired – either in preceding lessons or in informal settings outside of the school.

However, since most learning starts with doing, it is to be expected that if certain instructions stimulate students to think critically during an assignment, students' critical thinking skills will be trained. Therefore, it might be expected that frequent use of this type of assignment will result in students' critical thinking skills increasing, these skills being used more naturally and frequently in all kinds of situations.

Limitations

The current study also has a number of limitations. The sample size is not very large, and has therefore limiting power. Moreover, the study was conducted at one only one school with relatively high average exam scores and may not be representative of all other secondary schools in the Netherlands.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Onno Peeperkorn for the practical preparation of all chemical experiments and his help with the supervision of the

students during the conduction of their experiments. M. E. van Brederode was financially supported by a research grant for teachers with a doctoral degree (the postdoc-VO project) which was financed by the ministry van Education Culture and Science (OCW). This program is organized by the Freudenthal Institute for Didactics at Utrecht.

Notes and references

- Agustian H. Y. and Seery M. K., (2017), Reasserting the role of pre-laboratory activities in chemistry education: a proposed framework for their design, *Chem. Educ. Res. Pract.*, **18**, 518–532.
- Chase A., Pakhira D. and Stains M., (2013), Implementing process-oriented, guided-inquiry learning for the first time: adaptations and short-term impacts on students' attitude and performance, *J. Chem. Educ.*, **90**, 409–416.
- Chase A. M., Clancy H. A., Lachance R. P., Mathison B. M., Chiu M. M. and Weaver G. C., (2017), Improving critical thinking via authenticity: the CASPiE research experience in a military academy chemistry course, *Chem. Educ. Res. Pract.*, **18**, 55–63.
- Cohen J., (1992), A power primer, *Psychol. Bull.*, **112**, 155–159.
- Cooper M. M. and Kerns T. S., (2006), Changing the laboratory: effects of a laboratory course on students' attitudes and perceptions, *J. Chem. Educ.*, **83**, 1356–1361.
- Deci E. L., Vallerand R. J., Pelletier L. G. and Ryan R. M., (1991), Motivation and education - the self-determination perspective, *Educ. Psychol.*, **26**, 325–346.
- Ditzler M. A. and Ricci R. W., (1994), Discovery chemistry - balancing creativity and structure, *J. Chem. Educ.*, **71**, 685–688.
- Domin D. S., (1999), A review of laboratory instruction styles, *J. Chem. Educ.*, **76**, 543–547.
- Facione P., (1990), Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report). Educational Resources Information Center (ERIC).
- Farrell J. J., Moog R. S. and Spencer J. N., (1999), A guided inquiry general chemistry course, *J. Chem. Educ.*, **76**, 570–574.
- Fay M. E., Grove N. P., Towns M. H. and Bretz S. L., (2007), A rubric to characterize inquiry in the undergraduate chemistry laboratory, *Chem. Educ. Res. Pract.*, **8**, 212–219.
- George-Williams S. R., Soo J. T., Ziebell A. L., Thompson C. D. and Overton T. L., (2018), Inquiry and industry inspired laboratories: the impact on students' perceptions of skill development and engagements, *Chem. Educ. Res. Pract.*, **19**, 583–596.
- Gupta T., Burke K. A., Mehta A. and Greenbowe T. J., (2015), Impact of guided-inquiry-based instruction with a writing and reflection emphasis on chemistry students' critical thinking abilities, *J. Chem. Educ.*, **92**, 32–38.
- Hofstein A. and Lunetta V. N., (2004), The laboratory in science education: foundations for the twenty-first century, *Sci. Educ.*, **88**, 28–54.
- Holmes N. G., Kumar D. and Bonn D. A., (2017), Toolboxes and handing students a hammer: the effects of cueing and instruction on getting students to think critically, *Phys. Rev. Phys. Educ. Res.*, **13**, 010116.
- Holmes N. G., Wieman C. E. and Bonn D. A., (2015), Teaching critical thinking, *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 11199–11204.
- Johnstone A. H., (1997), Chemistry teaching - Science or alchemy? 1996 Brasted lecture, *J. Chem. Educ.*, **74**, 262–268.
- Keys C. W., Hand B., Prain V. and Collins S., (1999), Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science, *J. Res. Sci. Teach.*, **36**, 1065–1084.
- Kirschner P. A., (1992), Epistemology, practical work and academic skills in science education, *Sci. Educ.*, **1**, 273–299.
- Lazonder A. W. and Harmsen R., (2016), Meta-analysis of inquiry-based learning: effects of guidance, *Rev. Educ. Res.*, **86**, 681–718.
- Mistry N., Fitzpatrick C. and Gorman S., (2016), Design your own workup: a guided-inquiry experiment for introductory organic laboratory courses, *J. Chem. Educ.*, **93**, 1091–1095.
- Oliver-Hoyo M. T., (2003), Designing a written assignment to promote the use of critical thinking skills in an introductory chemistry course, *J. Chem. Educ.*, **80**, 899–903.
- Quattrucci J. G., (2018), Problem-based approach to teaching advanced chemistry laboratories and developing students' critical thinking skills, *J. Chem. Educ.*, **95**, 259–266.
- Rodriguez J. M. G., Bain K., Hux N. P. and Towns M. H., (2019), Productive features of problem solving in chemical kinetics: more than just algorithmic manipulation of variables, *Chem. Educ. Res. Pract.*, **20**, 175–186.
- Rodriguez J. M. G. and Towns M. H., (2018), Modifying laboratory experiments to promote engagement in critical thinking by reframing prelab and postlab questions, *J. Chem. Educ.*, **95**, 2141–2147.
- Seery M. K., Jones, A. B., Kew W. and Mein T., (2019), Unfinished recipes: structuring upper-division laboratory work to scaffold experimental design skills, *J. Chem. Educ.*, **96**, 53–59.
- Walker J. P. and Sampson V., (2013), Argument-driven inquiry: using the laboratory to improve undergraduates' science writing skills through meaningful science writing, peer-review, and revision, *J. Chem. Educ.*, **90**, 1269–1274.
- Walker J. P., Sampson V. and Zimmerman C. O., (2011), Argument-driven inquiry: an introduction to a new instructional model for use in undergraduate chemistry labs, *J. Chem. Educ.*, **88**, 1048–1056.
- Xu H. Z. and Talanquer, V., (2013a), Effect of the level of inquiry of lab experiments on general chemistry students' written reflections, *J. Chem. Educ.*, **90**, 21–28.
- Xu H. Z. and Talanquer, V., (2013b), Effect of the level of inquiry on student interactions in chemistry laboratories, *J. Chem. Educ.*, **90**, 29–36.