Scaffolding a Complex Task of Experimental Design in Chemistry with a Computer Environment

Isabelle Girault · Cédric d'Ham

Abstract When solving a scientific problem through experimentation, students may have the responsibility to design the experiment. When students work in a conventional condition, with paper and pencil, the designed procedures stay at a very general level. There is a need for additional scaffolds to help the students perform this complex task. We propose a computer environment (copex-chimie) with embedded scaffolds in order to help students to design an experimental procedure. A pre-structuring of the procedure where the students have to choose the actions of their procedure among pre-defined actions and specify the parameters forces the students to face the complexity of the design. However, this is not sufficient for them to succeed; they look for some feedback to improve their procedure and finally abandon their task. In another condition, the students were provided with individualized feedbacks on the errors detected in their procedures by an artificial tutor. These feedbacks proved to be necessary to accompany the students throughout their experimental design without being discouraged. With this kind of scaffold, students worked longer and succeeded better to the task than all the other students.

Keywords Experimental design · Scaffolding · Technology-enhanced-learning (TEL) systems · Feedback · Praxeologies · Scientific inquiry

I. Girault · C. d'Ham
Laboratoire d'Informatique de Grenoble, MeTAH Team, University of Grenoble, Grenoble, France

I. Girault (✉) · C. d'Ham
ENSE3 Ampère, Domaine Universitaire, 11 rue des Mathématiques, BP 46, 38402 St Martin d'Hères Cedex, France
e-mail: Isabelle.girault@imag.fr

Introduction

Experimental Design

Learners engaged in an inquiry sequence may be confronted with the task of experimental design. Typically, when learners have to solve a scientific problem through experimentation, it may be their responsibility to design or fine-tune the experiment. Koretsky et al. (2008); Neber and Anton (2008) observe higher-order cognitive activities of students facing such a task. Apedoe and Ford (2010) stress the importance to help students acquire an empirical attitude by making them design experiments. Karelina and Etkina (2007) find that, when students design their own experiments, they engage in behaviours that are much closer to the ones of scientists than did students working in traditional laboratories, because they spend more time "making sense", *i.e.* in discussions about physics concepts, experimental design, and data analysis. Arce and Betancourt (1997) find that, in the exams, students show a better understanding of concepts related to the experiments they design themselves, while Séré (2002) suggests that experimental designs might be helpful to acquire procedural knowledge. Etkina et al. (2010) find that when students are used to design experiments, they perform similarly on exams than students who did not design experiments, while they develop further scientific abilities (*i.e.* the most important procedures, processes, and methods that scientists use when constructing knowledge and solving experimental problems).

Experimental design is a complex task for students (Séré and Beney 1997), which may be part of the reason why it is difficult for a teacher to let students carry on such tasks (Girault et al. 2012). Several difficulties encountered by students have been reported, including correctly analysing the issue, putting the experimental procedure into words which relates

to difficulties in writing a text (Marzin and De Vries 2008), taking into account the question of measurement accuracy (Girault et al. 2007), and using the necessary conceptual knowledge they should master (Laugier and Dumon 2003). Thus, if we expect students to design experiments, they have to be scaffolded in order to reduce cognitive load: "design activities, when embedded in an inquiry cycle and appropriately scaffolded and supplemented with reflection, can promote the development of scientific abilities that are an important part of scientific practice" (Etkina et al. 2010).

Scaffolding Inquiry-Based Science Education

Computer environments can support learner needs in inquiry-based science education by scaffolding student understandings and skills. In simple terms, scaffolding means that support structures (scaffolds) are provided when novices or learners cannot work unassisted and require support to accomplish a task (Wood et al. 1976). In a formal learning situation, scaffolding ought to support students in achieving intended learning goals and tasks (Hmelo-Silver et al. 2007). The scaffolding strategy depends on the goals for a given task or subtask: just doing, learning how to do it, or learning why it should be done that way. Reiser's categories (2004) emphasize the dual nature of scaffolding while considering doing and learning: structuring the task and problematizing aspects of subject matter. The first category is mainly driven by facilitation purposes in order to enable the learner to achieve the task (reduce the complexity, maintain direction, and provide additional structure, etc). On the contrary, the problematizing mechanism is "to make some aspects of students' work more problematic. […] This may actually add difficulty in the short term, but in a way that is productive for learning". According to Reiser, one must "look for an optimal balance with the tension between structuring the task and problematizing".

Hmelo-Silver et al. (2007) consider the scaffolding capabilities of computer environments for structuring complex tasks: (a) a task can be structured in ways that allow the learner to focus on aspects of the task that are relevant to the learning goals; (b) the software can restrict the options available to students. This is coherent with Quintana et al. (2004) who envisioned software itself as the scaffolding rather than using "help functions" and other options within software as the scaffolding.

Only few studies really take into account the evolution of individuals because it requires long-lasting experiments with the same population. There is an obvious balance between artificial and human tutors. This is discussed in the specific and deeply study of Graesser et al. (2005) who are convinced that artificial tutors are necessary. Indeed, "one-on-one scaffolding is not always a viable option in a classroom with a single instructor" (Morgan and Brooks 2012).

A limitation underlined by Reiser (2004) arises from "the limited ability of most scaffold tools to individualize their support". Providing appropriate support remains a major challenge. Appropriate means that the support is adapted to one student's need, that it does not give more help than necessary, and that it is provided at the moment it is needed: the scaffolding process must assess "the learner's actual state of knowledge" (de Jong and van Joolingen 1998). This requires a proper and on time identification of each student's features: "the individual student's needs, predilections, interests, and abilities" have to be fully considered (Lipscomb et al. 2012). The tutor (human or artificial) has therefore to collect all the required information. This means that the interaction between the tutor and the learner is not only composed of elements meant to support the learner but also combines support actions with collecting information for the tutor. De Jong and van Joolingen (1998) suggest that scaffolding tools should be used as unobtrusive measures. This point has also to do with the fading processes when considering the individual evolution of learners' knowledge and skills: the complex integration of scaffolds for complex tasks requires appropriate fading of the scaffold as learning outcomes are achieved (Pea 2004).

Purpose of the Study

We developed a computer environment (name to be added in the final version), which scaffolds the activity of experimental design. It has been conceived to help learners to design a specific experiment in Chemistry, with several embedded scaffolds. The purpose of this study is to check whether this computer environment facilitates the task of experimental design. More precisely, we want to test different conditions of scaffolding and answer the following research question: under which scaffolding condition(s) do the students succeed in their experimental design?

In the next sections, we first describe the computer environment with its different types of scaffolds. Then, we present the research methods and the data obtained from 39 first-year university students working with paper and pencil or with two different configurations of the computer environment. The results concern the difficulties the students encounter when they design an experiment under the three scaffolding conditions. These results are discussed under the point of view of the types of scaffolding, our goal being to generalize our results beyond the computer environment described here.

Scaffolding a Design Task with Copex-chimie, a Computer Environment

Following the literature, copex-chimie embeds different types of scaffolds: scaffolds for helping the learners to achieve the

task or scaffolds for problematizing the task; generic scaffolds or individualized scaffolds. The first scaffold included in our computer environment is generic and related to the pre-structuring of the task and aims to help the students to achieve the task. The two other scaffolds are individualized scaffolds, based on the evaluation of the procedure either on the experimental aspect or on the knowledge aspect: one is constituted by the experimental simulated results given to the learners, according to their production; the other type of scaffold is provided through the feedback messages of an artificial tutor that give information on the learners' errors. We describe the strategy used to produce these messages, based on the analysis of the task and knowledge with the help of the praxeology model.

Copex-chimie: A Computer Environment for Experimental Design

This computer environment is a Web application (http://copex-chimie.imag.fr) in which the learners have to determine the concentration of the red dye in grenadine syrup by spectrophotometric titration. To attain this goal, students must write an experimental procedure that can be read by the application in order to simulate the experimental results. This chemistry laboratory work ("determine the concentration of a substance in a solution") refers to design problems (Apedoe and Ford 2010) where the

question is already given to students, and there is no need to formulate hypothesis. The students have to focus on the design of the experiment and the analysis of the data to conclude. In our situation, they do not have to perform the experiment since they obtain their data by simulation, according to the procedure they have described.

Scaffolding by Pre-structuring the Procedure

A model of experimental procedure has been previously described (Girault et al. 2012) following Leont'ev's model of activity (Leont'ev 1978). In this model, a complete procedure includes two types of tasks: steps and actions, the latter being described with parameters. To scaffold the activity of experimental design, the computer environment pre-structures the procedure at two levels, according to this model. At the highest level, the procedure must be written following three steps imposed to the learner: (1) prepare the standard solutions, (2) obtain the points of the standard curve, and (3) determine the concentration of the dye. At a lower level, the actions constituting the procedure must be chosen among a given list of eight actions (e.g. prepare a solution by dilution, measure an absorbance, etc). For each action added in the procedure, the parameters describing the action (two to five parameters) have to be set by the learner. For example, when choosing the action "prepare a solution by dilution", the learner has to set the following parameters: name of the new solution, volume of the
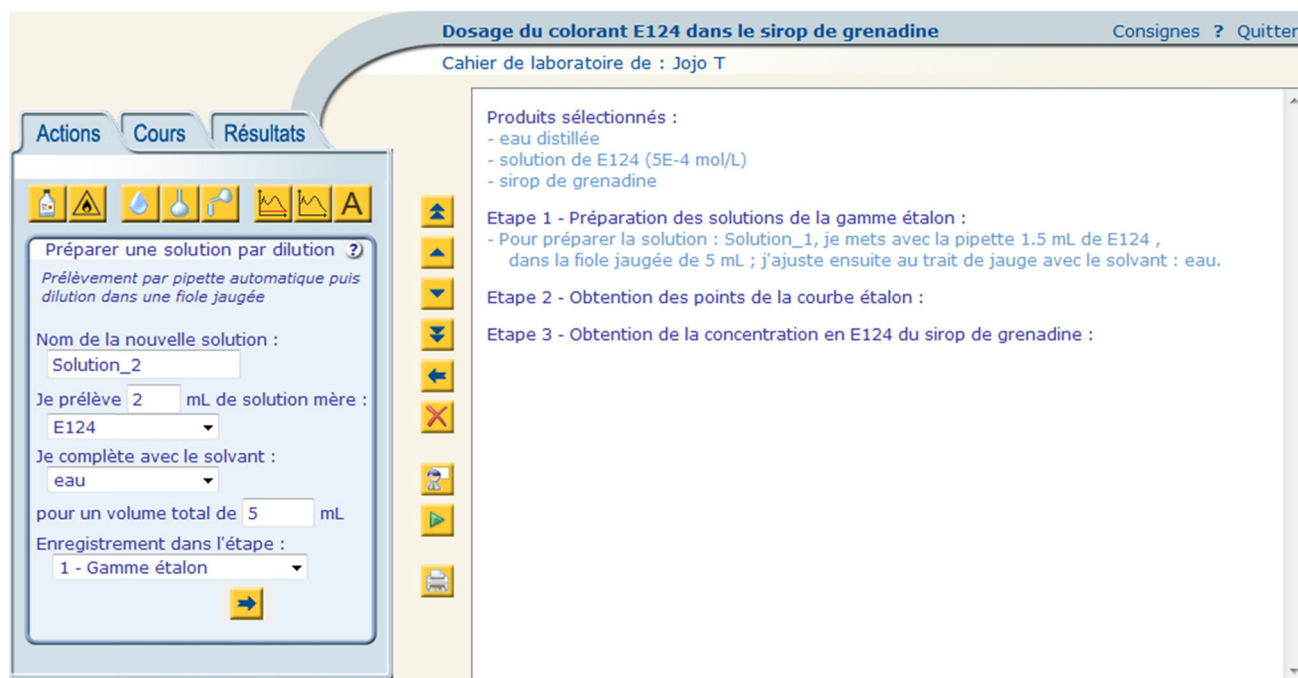

Fig. 1 Copex-chimie at the beginning of the learners work: on the right, stands the learner's experimental procedure pre-structured by three pre-defined steps (*darker lines*); on the left, is the frame used by the learner to define a new action by filling the boxes corresponding to

its parameters. Once the parameters are set up, the action automatically appears in the procedure on the right side under the chosen step (*bullet lines*)

parent solution, nature of the parent solution, solvent, and total volume of the new solution (see Fig. 1). Within this system, a valuable procedure is composed by at least 26 ordered actions with their adequate parameters' values.

The pre-structuring scaffold corresponds to the first mechanism of scaffolding in software tools described by Reiser (2004): "one way to help learners is to use the tool to reduce complexity and choice by providing additional structure to the task". In their scaffolding design framework, Quintana et al. (2004) also propose a guideline named "provide structure for complex tasks and functionality". The design process students are involved in is similar to what Ohlsson (1996) describes as "sequential choice tasks". When an action is chosen, the subsequent task requires a sequence of discrete actions (*i.e.* set up the value of each parameter of the action). When there are multiple options, each action is the result of a choice among competing alternatives.

Scaffolding by Providing Feedbacks with Simulated, Empirical Results

Another strategy of scaffolding embedded in the computer environment is constituted by the feedback given to the user. This feedback is made of two kinds of information: (1) the experimental results corresponding to the procedure and (2) the errors detected in the procedure by an artificial tutor. In this section, the first kind of feedback is described.

At any point of their work and without any limitation, learners can ask the system to provide the empirical results corresponding to the experiment they have, so far, designed. These results are simulated by a spectrophotometry simulation. Absorbance values and absorbance spectrums can be simulated (see Fig. 2). The calculation of these data is possible due to the pre-structuring of the procedure with actions. The simulation is able to select the values of the actions' parameters that are needed for the calculations. Due to the diversity of the procedures that can be produced by learners, the simulation cannot simulate any results. In fact, the artificial tutor (see next section) determines whether the procedure fits the domain of validity of the simulation. If it does, results are provided to the learners.

The experimental results have to be treated to answer the initial question ("determine the concentration of the red dye in the grenadine syrup"), and during this treatment, the learners have to evaluate the validity of their results.

Scaffolding by Providing Feedbacks on the Learner's Errors

The Diagnostic System for Detecting the Errors in the Procedure

A feedback, provided by an artificial tutor, is accessible on demand to the learner. The tutor evaluates the three steps of the procedure with a set of constraints, as described by
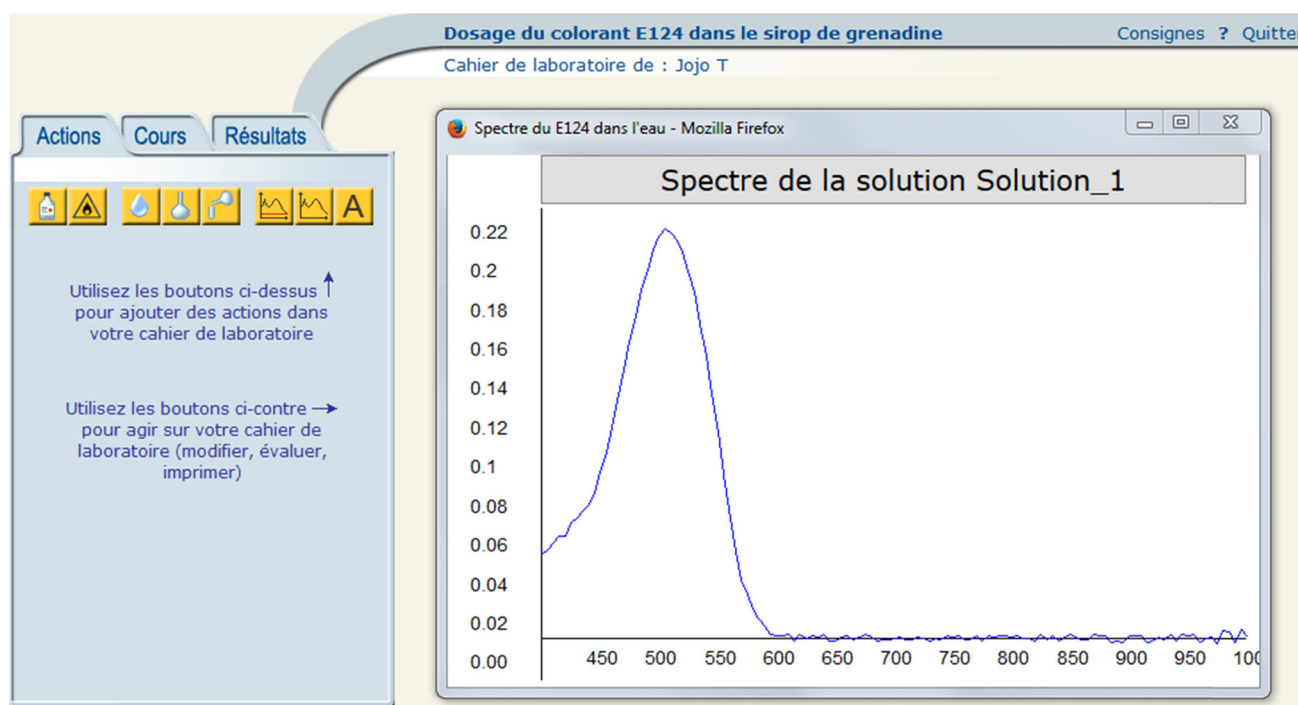

Fig. 2 Copex-chimie with a result window opened. The window displays an absorbance spectrum simulated by the computer environment, based on the actions' parameters of the learner's procedure

Ohlsson ([2002](#)): each constraint has a satisfaction condition and a relevance condition. The satisfaction condition determines whether the error is present in the procedure. For example, a satisfaction condition used in copex-chimie is "Are the standard solutions made with the compound to measure and the adequate solvent?" The relevance condition determines for which states of the procedure a constraint should be verified. For example, the previous satisfaction condition should be explored only if the learner has prepared solutions in the first step: "prepare the standard solutions". With this tutoring system, all the usual errors made by students are detected, even if there is not a unique procedure that can be considered as correct. Furthermore, only the state of the procedure is used for the diagnosis, without recourse to the process that has been necessary to create it. Thus, this tutoring system is completely adapted for the student's exploratory process with no pre-defined strategy of resolution and with no unique correct solution, which characterizes design tasks.

The Feedback Strategy in the Light of the Praxeology Model

Once the diagnosis is made, the artificial tutor points out to the learner the errors detected in the procedure (see Fig. [3](#)). The teacher can initially fine-tune the feedback provided by the artificial tutor in two ways: the total number of accesses to the tutor can be limited during the session, and the level of detail for describing the errors to the user can be adjusted to three options with more or less details concerning the errors: the level 1 is a global level and points out the achievement for each step via a gauge; the level 2 provides the amount of errors in each category for each step (there are six categories: objective, practical problems, washing, series of standard solutions, homogenizing, and spectrophotometry); finally, the level 3 provides details for each error. It is possible to give the students access to all the three levels or limit to level 1 or to level 1 and 2. For example, if a learner does not choose an appropriate solution to wash a volumetric flask, the artificial tutor can give, on request, the following message at level 2: "you have one error related to the washing" (in this case, the category of the error is the washing), and at level 3: "evaluate the influence of your washing solution on the solution prepared by dilution". Furthermore, the artificial tutor provides links to pages of information related to the detected errors. It has to be noted that the learner can also freely access to these pages from a menu in the application.

The content of the messages given at level 3 is based on the Anthropological Theory of the Didactic (Chevallard [1999](#); Rodriguez et al. [2007](#)). The general epistemological model provided by the Anthropological Theory of Didactic

Fig. 3 Copex-chimie with the artificial tutor frame opened below the experimental procedure: a global evaluation (level 1) is given step by step with gauges (*left*), and details are given for each error (level 3) on demand to the learner (*right*). Level 2 is not visible in this figure

proposes a description of both activity and knowledge in terms of praxeologies whose four main components are tasks, techniques, technologies, and theories. A task is a problem or a subproblem the learner has to solve. A technique is the process that the learner follows in order to accomplish a task. The technology is the discourse that the learners use to describe, explain, and justify a technique. The theory is a more general discourse that is necessary to understand the concepts used in the technology. The praxeologies consist of a practical block or "know-how" (the praxis) integrating tasks and techniques, along with a theoretical block or "knowledge" (the logos) integrating both the technological and the theoretical discourses.

This model of activity allows the analysis of the chemical knowledge during the activity of experimental design in copex-chimie. Table 1 illustrates an example of a praxeology that describes a task dedicated to the learners. In this example, the task is "choose the appropriate solution in order to wash a volumetric flask". The appropriate technique to solve this task is to consider the usage of the volumetric flask (preparation of a solution by dilution) and thus to choose the dilution solvent as washing solution. The technology that justifies this technique is, briefly, the following: after washing the volumetric flask with the dilution solvent, if there are still drops of solvent in the volumetric flask, it would not have any incidence for the final solution, as the flask will finally be filled by solvent. The theory is the broader discourse that explains the concepts of solution and concentration of solutes.

The praxeology modelling of the activity and knowledge for copex-chimie proved to be helpful to produce the content of the detailed feedback messages about the errors. For each error, we characterize it with two criteria. Looking at the activity aspect, we determine from the procedure whether the error corresponds to the absence of consideration of a task (unconsidered task) or whether it corresponds to a task executed with an incorrect technique. Looking at the knowledge aspect, we determine whether

the task and technique related to the error are considered as learning goals for the laboratory session. With these two criteria, we adopt the following strategy for writing the messages constituting the feedback to the learner (see also Table 2): for an unconsidered task, the message makes a reference to its type (e.g. "You didn't prepare any solution in this step"); for a task executed with an incorrect technique (e.g. the value of a parameter in the procedure is inadequate), we consider its importance in term of learning. If the task is a learning goal, the feedback is given at the technology level in order to make students think about the rationale of the task (e.g. "Evaluate the influence of your washing solution on the solution prepared by dilution"). If the task is not a learning goal, the feedback is given at the technique level in order to facilitate the success of the task (e.g. "You need to homogenize and transfer your solutions immediately once you have prepared them, since there is only one volumetric flask").

The principles that have been followed to create the feedback messages are consistent with Ohlsson's recommendations (1996) for error correction. This author recommends that built-in features and devices should tell the trainee when he or she does the wrong thing. Ohlsson (1996) adds that the formulation of verbal instructions is important and that these instructions need to focus on the features of the decision situation, as opposed to the action itself. This is the position we adopted in copex-chimie for errors related to learning goals, where the tutor does not tell what should have been done, but makes the students think about why an action is inappropriate with a message situated at the technology level. If we refer again to Reiser's work (2004), messages given by the tutor at the technology level correspond to the second mechanism of scaffolding he describes "problematize subject matter", while messages given at the technique level aim to facilitate the task of the learner. An interesting feature of this feedback focusing on the learner's errors is that it individualizes the support in relation to the individual needs.

Table 1 Detailed example of a praxeology corresponding to the task "choose the appropriate solution to wash a volumetric flask"

Praxeology		
PPRAXIS: know-how	Task	Choose the appropriate solution in order to wash a volumetric flask
	Technique	Use the solvent of dilution to wash the flask
LOGOS: knowledge	Technology	Using any other solution would change either the concentration of the solute or the composition of the solution
	Theory	All the concepts around solutions and concentrations of solutes

Table 2 Strategy adopted for producing the feedback messages following the characteristics of the error and the praxeology analysis of the activity and knowledge

Characteristics of the error	Content of the feedback message
Unconsidered task	Focus on the type of the unconsidered task
Incorrect technique/ learning goal	Express the task whose technique is incorrect and provide hints at the technology level
Incorrect technique/not a learning goal	Express the task whose technique is incorrect and provide hints at the technical level

Table 3 Three students' groups have different scaffolding conditions: (+) the scaffold is present; (−) the scaffold is absent

Scaffolds	Groups		
	Without copex-chimie (no-copex) 9 students	Copex-chimie without tutor (copex-no-tutor) 16 students	Copex-chimie and full tutor (copex-full-tutor) 14 students
Pre-structuring: steps level	+	+	+
Pre-structuring: actions level	−	+	+
Feedback: simulated, empirical results	−	+	+
Feedback: errors	−	−	+

Methods

Experimental Setting

The trial has been performed at the University of Grenoble, France, in January 2010 with 39 first-year university students enrolled in a science curriculum. The context is an interdisciplinary course focused on laboratory work, with eight laboratory sessions designed around the theme of water, as well as pre- and post-laboratory sessions, involving five disciplines. We study one of the pre-laboratory sessions that lasts 120 min. During this session, the students are asked to write individually an experimental procedure to titrate the E124 dye in grenadine syrup, using a spectrophotometric method. The teacher briefly presents the work to be done and then lets the students work independently, without answering their questions related to the content. The students already have some knowledge about the titration by spectrophotometry since they have studied it during the previous year and some reminders are given in a previous session.

We test the experimental design situation with three different conditions.

The first one corresponds to 9 students working without copex-chimie. Documents and information are given to theses students, similar to what they could find in the computer environment:

– The detailed goal
– The principle of the method
– The available material and products
– An order of magnitude of the molar extinction coefficient
– A scientific handbook: procedural and theoretical information
– A pre-structuring of the experimental procedure in steps (select products, prepare the standard sample solutions, obtain the calibration curve points, and obtain the E124 concentration in the grenadine syrup.)

In the second group, 16 students are working with the computer environment without the artificial tutor (group "no-tutor"). The scaffold corresponds to the pre-structuring provided by copex-chimie at the steps and actions levels and the feedback on simulated data.

The third group corresponds to 14 students working with copex-chimie while the tutor is set up with no limitation (group "full tutor"). The students can access to the tutor as many times as they want and they get the three levels of feedback. The scaffold corresponds to the pre-structuring of the procedure provided by the computer environment and to the feedback on errors and simulated data.

Table 3 summarizes the different conditions of experimental design for the three groups regarding scaffolding.

Data Collection

Indicators Extracted from the Log-Files

For the students working with the copex-chimie, the log-files describing their activity during the session are recorded. The log-files are composed of the sequence of events describing the interaction of the users with the computer environment. Each event is described by a time code, a user code, the name of the event, and the values of its parameters. The activity is recorded until students complete the task or during the whole session for the ones who do not complete the task.

From the log-files, indicators are extracted to describe the students' work with copex-chimie and particularly their interaction with the provided scaffolds. For the group "no-copex", we manually record similar data. The indicators are the following:

– The duration of the students' work: some students stop working before the end of the session, either because they consider they managed to do their work, or because they abandon.
– The success: it is a score out of 20 that corresponds to a global success. In copex-chimie, the tutor automatically calculates this score in order to display the global evaluation gauges. For the group "no-copex", the

written procedures are analysed manually following the algorithm used in the computer environment.

The following indicators only concern the two groups of students working with copex-chimie:

– The number of accesses to the simulation.
– The actions students perform after they request and consult (if any) simulated results.

Analysis of Students' Final Procedures

In order to determine the difficulties that students encounter when they design the experiment, we analyse their final procedures, either retrieved from a paper document for the "no-copex" group, or from the computer environment for the other groups. The analysis of the procedures is a manual process, as described below.

From the praxeology analysis of the task, seven subtasks related to learning goals have been selected. We search for the difficulties of the students within these seven tasks (numbered from T1 to T7):

T1. Choose the nature of the standard solutions for the calibration curve
T2. Choose the nature of the sample to measure
T3. Choose the concentrations of the standard solutions, with regard to the measurement range of the spectrophotometer
T4. Choose the appropriate solution to wash the volumetric flask
T5. Choose the appropriate solution to wash the cuvette
T6. Choose the reference solution for the measures of absorbance
T7. Choose the wavelength for the measures of absorbance by analysing a spectrum

For each task, a difficulty can be expressed in two ways: it can be the absence of consideration of the task by the student (unconsidered task) or the use of an erroneous technique to complete the task.

This diagnosis is obtained from the content analysis of the procedures: unconsidered tasks are detected through missing actions or missing parameters, and erroneous techniques are detected through inadequate values of some action's parameters.

Limitations of the Study

Our research intends to explore the impact of the different scaffolds on the students' work. With the current experimental setting, we cannot evaluate separately the effects of the low-level pre-structuring scaffold (actions) from the effects of the empirical feedback (simulated results)

scaffold. These two scaffolds are simultaneously added to the group "copex-no-tutor" in comparison with the group "no-copex". In order to get some insights into this question, we have explored the use of the simulated empirical results by the students in the two groups using the computer environment.

Results

Duration, Success

For students of the three groups, we recorded the time spent for writing their experimental procedure and their associated success score. The average results per group are given in Table 4.

The more students have scaffolds, the longer they work until they stop working. It can be noted that there is an important disparity among students of the "copex-no-tutor" group regarding the time spent on their design.

The success score slightly improves from the "no-copex" to the "copex-no-tutor" condition and can be attributed to the effect of two scaffolds (the pre-structuring of the actions and the simulated results). However, the success for the group "copex-no-tutor" is still low, since students from the group "copex-full-tutor" succeed much better.

Students' Difficulties

Characterization of the Difficulties with Criteria

In order to characterize the students' difficulties at a higher level, we use a set of criteria proposed for the evaluation of student-written procedures (Girault et al. 2012). These criteria are of three types: communicability, relevance, and executability (Table 5). The relations between the difficulties and the criteria are then discussed.

Some of the criteria used to evaluate the experimental design are not concerned in this experimental design: the structure of the procedure (communicability) is given to the

Table 4 Duration of the students' work to design their experiment and associated success score

	No-copex (9 students)	Copex-no-tutor (16 students)	Copex-full-tutor (14 students)
Duration (min)	69 (s.d.7)	77 (s.d.22)	98 (s.d.6)
Success (out of 20)	8.8 (s.d.2)	10.8 (s.d.4)	17.1 (s.d.3)

Table 5 Criteria to evaluate the students' procedures after their experimental design activity (extract from Girault et al. 2012)

A. *Communicability: the description of the experiment (a target reader needs to be specified)*

    Structure (temporal or logical organization of the sequence of actions/easy to be read by the target)

    Completeness (does the procedure give enough detail for the target reader to execute the experiment?)

B. *Relevance: the function of the experiment*

    External relevance between the hypothesis and the quantities to measure

    Internal relevance: measurement strategy (choice of the methods and materials)

    Quality of data acquisition: trueness and precision

C. *Executability: the experiment in the laboratory conditions*

    Adequacy between the samples and the domains of validity of the measurement methods and materials

    Observation of material constraints (availability, cost, feasibility, and hazard control)

    Observation of temporal constraints

students in the three experimental groups in the form of pre-defined steps; the external relevance between the hypothesis and the quantities to measure does not apply in this situation since the students do not formulate hypotheses; the material constraints (executability) are not relevant since a list of material is given to the students; the temporal constraints are not to be considered as the empirical results are given instantly by the simulation. Thus, our results address only the four following criteria:

– Communicability: completeness
– Relevance: internal relevance
– Relevance: quality of data acquisition
– Executability: adequacy between the samples and the domains of validity of the measurement methods and materials

Unconsidered tasks are problematic to characterize from the analysis of the student's procedure. They can either be attributed to the criterion of communicability completeness (a student does not write down some actions because he/she considers that these actions are obvious and will be treated during the manipulation), or to one of the three other criteria (a student does not intend to carry out the missing actions). By only analysing the written procedures, we are not able to determine accurately to which criteria an

unconsidered task should be attributed. Without additional information, we choose to relate the unconsidered tasks difficulties to the completeness criteria. Thus, for the communicability-completeness criteria, we consider the presence of the actions in the procedure, and when the actions are present, we check the presence of the parameters needed to execute the experiment. All the tasks (T1–T7) are concerned with the completeness.

For the relevance and executability criteria, we give the number of tasks that have an incorrect technique. Depending on the tasks, an incorrect technique can be related to different criteria: T1, T2, and T7 are connected to the "internal relevance"; the "quality of data acquisition" corresponds to T4, T5, and T6, while an incorrect technique for the task T3 can be explained by the criterion "adequacy between the samples and the domains of validity of the measurement methods and materials".

Difficulties Related to Unconsidered Tasks: Communicability Criterion

Table 6 includes the results associated with the communicability-completeness criterion. We count the unconsidered tasks, detected in the procedure as missing actions or missing parameters inside an action. We consider the seven tasks for all the students in each group.

The results show that there are less unconsidered tasks under the condition "copex-no-tutor" (28 %) than the condition "no-copex" (54 %). The number of unconsidered tasks even decreases with the group "copex-full-tutor" (11 %).

Difficulties Related to Incorrect Techniques: Relevance and Executability Criteria

Table 7 displays the amount of difficulties in relation to the relevance and the executability criteria. The results are not detailed: only a global number for each criterion is given that includes all the concerned tasks. The maximum amount of tasks is less important than in Table 6, since each criterion does not concern all the seven tasks on one hand, and we do not count the unconsidered tasks on the other hand. For example, for the internal relevance criterion, 17 tasks are counted for the "no-copex" group. This number corresponds to the tasks T1, T2, and T7, so 27

Table 6 Difficulties related to unconsidered tasks, compared to the maximum number of tasks per group (7 tasks multiplied by the number of students in a group)

	No-copex (9 students)	Copex-no-tutor (16 students)	Copex-full-tutor (14 students)
Communicability			
Unconsidered tasks (tasks T1–T7)	34/63 (54 %)	31/112 (28 %)	11/98 (11 %)

tasks (3 tasks multiplied by 9 students) to which were removed 10 tasks unconsidered by the students.

The amount of difficulties associated with the "internal relevance" criterion does not seem to be correlated with the increase in scaffold. Regarding the "quality of data acquisition", we do not consider the results of the no-copex group since the number of tasks concerned is too small to give a significant result. However, there is an impact of the tutor scaffold since 17 % of the tasks have an incorrect technique for the "copex-no-tutor" group, whereas none are incorrect for the group "copex-full-tutor".

Regarding the executability criterion, the only difference appears with the group "copex-full-tutor" whose students perform better than students of the two other groups.

Use of the Simulated Results

We expected students with simulated results to improve their technique for the task T3 associated with the executability. Indeed, a simulated spectrum can inform the students that a prepared solution is too concentrated or too diluted, which corresponds to an executability problem regarding the domain of validity of the spectrophotometer. Unexpectedly, students from the group "copex-no-tutor" have similar results than students from the group "no-copex", regarding T3 and the executability of their procedures (Table 7).

In order to get more insight into the use of the simulation by the students, we extract from the logs some indicators regarding the use of simulation and the subsequent actions. These results (Table 8) only concern the two groups using copex-chimie who can access to the simulated results.

On average, the students without tutor request the simulated results much more than the students with tutor, but a simulated result is less often provided to the students of the group "copex-no-tutor" (14 %) compared to the students of the group "copex-full-tutor" (60 %). When a simulated result is not provided, the students have a message

indicating that their procedure is not adapted to calculate simulated results.

After a simulation request, the students of the group "copex-no-tutor" mainly modify their procedure (70 %), and in a less extent, they search for information (30 %) either in the scientific content or in the detailed goal.

On the other hand, after requesting the simulation, the students from the group "copex-full-tutor" mainly ask for an evaluation from the tutor (45 %) and in a less extent modify the procedure (32 %) or search for information (22 %).

Discussion

We discuss the results regarding the impact of the different scaffolds on the duration, the success, and the students'

Table 8 Use of the simulation by the two groups of students using the computer environment

	Copex-no-tutor (16 students)	Copex-full-tutor (14 students)
Number of simulation requests per student	9.7	3.7
Number of times the simulation provides a result compared to the total number of simulation requests	22/155 (14 %)	31/52 (60 %)
Number of times the students modify, add, or delete an action immediately after requesting the simulation	105 (70 %)	16 (32 %)
Number of times the students consult information immediately after requesting the simulation	44 (30 %)	11 (22 %)
Number of times the students request the tutor evaluation immediately after requesting the simulation	N.a.	22 (45 %)

Table 7 Difficulties related to tasks that have an incorrect technique in the students' final procedures, compared to the corresponding amount of considered tasks. These difficulties are organized according to the criteria of relevance and executability

Difficulties related to:	No-copex (9 students)	Copex-no-tutor (16 students)	Copex-full-tutor (14 students)
Relevance			
Internal relevance: measurement strategy (tasks T1, T2 and T7)	5/17 (29 %)	13/30 (43 %)	5/33 (15 %)
Quality of data acquisition: trueness and precision (tasks T4, T5 and T6)	2/3	6/35 (17 %)	0/40 (0 %)
Executability:			
Adequacy between the samples and the domains of validity of the measurement methods and materials (task T3)	5/9 (56 %)	8/16 (50 %)	4/14 (29 %)

difficulties. The first scaffold to be analysed is the scaffold based on the provision of empirical simulated results.

Scaffolding by Providing Feedbacks with Simulated, Empirical Results

In a first analysis, we expected that the students who have access to empirical results would increase the executability of their procedures: simulated spectrums or absorbance values being out of range should help the students to reflect on the adequacy between their samples and the domain of validity of the spectrophotometer (executability criterion). Comparing the results of the students working with the simulation ("copex-no-tutor" group) and without ("no-copex" group), it appears that the difficulties related to the executability criterion do not evolve with the access to simulated results (Table 7). The log-files show that the students in the "copex-no-tutor" group request about 10 times the simulation during the session (Table 8), but they obtain a simulated result only 14 % of the times they ask for it (Table 8). In average, each student of the "copex-no-tutor" group receives only one to two simulated results during his session. For this group, it seems that simulated results are searched by the students with a validation purpose in the flow of their actions, probably in a trial and error manner: 70 % of the actions following a simulation's request are made to modify the procedure (Table 8).

Without access to the tutor, the students in the "copex-no-tutor" group frequently ask for feedbacks from the simulation, but their procedure is not good enough to obtain the expected simulated results. Thus, it appears that the simulated results, by themselves, do not provide a significant help to the students. The differences observed between the students working without the computer environment and the students from the "copex-no-tutor" group might principally be the consequence of the pre-structuring scaffold.

Scaffolding by Pre-structuring the Procedure

The effect of the high-level pre-structuring scaffold (the three given steps) was not evaluated, since this scaffold is provided to the three groups of students.

The low-level pre-structuring scaffold corresponds to the given list of actions with parameters available in copex-chimie. This list of actions is not given to the students working without copex-chimie. To study the impact of the pre-structuring scaffold, we can thus compare the results from the groups "no-copex" and "copex-no-tutor", having said previously that the simulation is not a great help for the students of the group "copex-no-tutor".

Students with the pre-structuring scaffold ("copex-no-tutor") spend a little more time for designing their procedure than students without this scaffold ("no-copex") (Table 4). This result can correspond to the time they need to get accustomed to the software. However, this cannot explain the disparity in time among the students from the group "copex-no-tutor". We explain this disparity by the fact that some students of the "copex-no-tutor" group have enough scaffolds to keep working while others are discouraged. The success score (Table 4) makes us believe that the pre-structuring scaffold is not enough to succeed, since the success is low (10.8 out of 20) even if the score slightly improves compared to the group "no-copex".

Regarding the completeness of the procedure, there are more unconsidered tasks in the "no-copex" group than for the students using the software (Table 6). This means that the list of actions helps the students to think about what could be needed to write their procedure. This has to be compared to the results of Jordan et al. (2011) who find that "the tools [a list of available equipment] made available to the novice students appeared to strongly guide their experimental design". The drawback is that it can limit their creativity and the students tend to be "driven (…) by task completion" (Jordan et al. 2011).

The pre-structuring scaffold also has an impact on "internal relevance" criteria. We could have expected fewer difficulties for the students of the group "copex-no-tutor" than for those working without copex-chimie; however, the results show the opposite. In fact, students working with paper and pencil avoid dealing with some complex subtasks, such as T7 "Choose the wavelength for the measures of absorbance by analysing a spectrum". This explains why their procedures show more unconsidered tasks and also less incorrect techniques for the internal relevance criteria. For the group "copex-no-tutor", the pre-structuring scaffold helps the students to take into account such a complex task but they have strong difficulties and incorrect techniques appear when they give more details.

Scaffolding by Providing Feedbacks on the Learner's Errors

The increase in time spent on the experimental procedure for students of the "copex-full-tutor" group can be explained by the time needed to explore the feedbacks of the tutor, since these students request on average 32 times the help of the tutor. Furthermore, students of the "copex-full-tutor" group seem not to be discouraged, as the other students could be. They tend not to abandon their goal and succeed fairly well in writing an experimental procedure (Table 4).

Their success is coherent with the fact that many students of the group "copex-full-tutor" manage to overcome their difficulties. The tutor seems to have a positive impact on each criterion. The decrease in the number of

unconsidered tasks (Table 6) and the smaller number of incorrect techniques (Table 7) for the group "copex-full-tutor" compared to the group "copex-no-tutor" can be attributed to the effect of the tutor and corroborate the need for feedbacks on top of the other tested scaffolds. This scaffold has the power to individualize the feedback (Reiser 2004), which is a real challenge for an artificial tutor. We chose to make this feedback unobtrusive, since the students decide when they want a feedback from the tutor. Consequently, the strategy of the students can vary with the use of the tutor. In another study (data to be published), we want to check how the students adapt their strategy when they have a limited or unlimited access to the tutor.

Furthermore, the students from the group "copex-full-tutor" take more advantage of the simulation in combination with the tutor evaluation, since they very often obtain simulated results (Table 8). This has a positive impact on the executability criterion (Table 7) since the students of the group "copex-full-tutor" have fewer difficulties with the associated task T3 than the students of the group "copex-no-tutor". Thus, the simulation is a scaffold that is mainly useful in combination with the tutor that helps the learners to deal with the validity domain of the simulation.

Conclusion

Our results show how experimental design is a complex task for students, even if the knowledge at stake has been studied before. When facing this task in conventional conditions, working with paper and pencil and few scaffolds (some scientific information, a pre-structuring of the procedure with three steps), we observe that students finish quickly their job but the result is not the expected one. The designed procedures stay at a very general level: they tend to be an overview of the experiment to come than a real procedure that would be helpful to carry on the manipulation. Many parameters are missing, and even whole parts of the procedures are eluded. In fact, the task seems to be too complex for the learners. They do not give details, either for avoiding the complexity of the reasoning, or because they cannot grasp this level of complexity by themselves.

We propose a computer environment with embedded scaffolds in order to help students to design an experimental procedure. The first level of scaffold is the pre-structuring of the procedure at a low level: the students have to choose the actions of their procedure among pre-defined actions. This scaffold forces students to face the complexity of the design. Firstly, the given actions help the students to think about some aspects of their procedures, and secondly, as the actions have to be defined with

parameters, it forces the students to choose a value for these parameters. As it is shown by our results, this scaffold does not allow the students to succeed in their design much better than students working without the computer environment. The students working with the computer environment and the pre-structuring scaffold seem to look desperately for some feedback from the simulation that they do not get because of the poor quality of their procedures. Most of the students seem to experience failure and they abandon their task.

In a third condition, the students working with the computer environment were provided with individualized feedbacks on the errors detected in their procedures by an artificial tutor. These feedbacks proved to be necessary to accompany the students throughout their experimental design without being discouraged. With this kind of scaffold, students worked longer and succeeded better to the task than all the other students. The provided feedbacks helped them to improve their procedure and thus to get some simulated results. We reach a similar conclusion than Etkina et al. (2010) saying that design activities must be appropriately scaffolded and supplemented with reflection in order to be positively implemented in the classrooms.

Since some of the individualized feedbacks given by the tutor are situated at the technology level (the rationale level), we expect the students not only to succeed in their tasks but also to understand what they are doing and improve their learning. In a subsequent work, we need to analyse the learning outcomes of students using the artificial tutor. We expect that this strategy of scaffold will increase learning, as stated by Ohlsson (1996): "If an action is incorrect, the knowledge structure (…) that generated that action must be faulty. (…) To correct an error is to improve future performance by revising the relevant knowledge. (…) Error correction is a mental process that results in some improvement in the performer's knowledge about the task".

For experimental purpose, there were no interventions of the teacher during the sessions, since we wanted to analyse how the students deal by themselves with the design of an experiment under the three experimental conditions. In a regular class situation, the students work with the computer environment and interact with their teacher. In this case, the scaffolding is not only provided by the computer environment but also by the combination of the computer environment and the teacher. This synergy that we call "co-scaffolding" seems to be relevant when students are facing tasks as complex as experimental design proved to be. Future work needs to be conducted on this idea of co-scaffolding for helping students to achieve complex tasks.

References

Apedoe X, Ford M (2010) The empirical attitude, material practice and design activities. Sci Educ 19:165–186

Arce J, Betancourt R (1997) Student-designed experiments in scientific lab instruction. J Coll Sci Teach 27(2):114–118

Chevallard Y (1999) L'analyse des pratiques enseignantes en théorie anthropologique du didactique. Recherches en Didactique des Mathématiques 19(2):221–266

De Jong T, van Joolingen WR (1998) Scientific discovery learning with computer simulations of conceptual domains. Rev Educl Res 68(2):179–201

Etkina E, Karelina A, Ruibal-Villasenor M (2010) Design and reflection help students develop scientific abilities: learning in introductory physics laboratories. J Learn Sci 19:54–98

Girault I, Cross D, d'Ham C (2007) Students' adaptation to a new situation: the design of an experimental procedure. In Proceedings of ESERA conference. Malmö, Sweden, 21–25 August

Girault I, d'Ham C, Ney M, Sanchez E, Wajeman C (2012) Characterizing the experimental procedure in science laboratories: a preliminary step towards students experimental design. Int J Sci Educ 34(6):825–854. doi:10.1080/09500693.2011.569901

Graesser AC, McNamara DS, VanLehn K (2005) Scaffolding deep comprehension strategies through point and query, autotutor, and iSTART. Educ Psychol 40(4):225–234

Hmelo-Silver CE, Duncan RG, Chinn CA (2007) Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). Educ Psychol 42(2):99–107. doi:10.1080/00461520701263368

Jordan RC, Ruibal-Villasenor M, Hmelo-Silver CE, Etkina E (2011) Laboratory materials: affordances or constraints? J Res Sci Teach 48(9):1010–1025. doi:10.1002/tea.20418

Karelina A, Etkina E. (2007). Acting like a physicist: student approach study to experimental design. Phys Rev ST Phys Educ Res 3:12 (online Journal: http://prst-per.aps.org/)

Koretsky MD, Amatore D, Barnes C, Kimura S (2008) Enhancement of student learning in experimental design using a virtual laboratory. IEEE Trans Educ 51(1):76–85

Laugier A, Dumon A (2003) Résolution de problème et pratique expérimentale : analyse du comportement des élèves en début de seconde. Chem Educ Res Pract 4(3):335–352

Leont'ev AN (1978) Activity, consciousness, and personality. Prentice-Hall Englewood Cliffs, Nj

Lipscomb, Swanson, West (2012, January 26). Scaffolding—emerging perspectives on learning, teaching and technology. Retrieved from http://projects.coe.uga.edu/epltt/index.php?title=Scaffolding

Marzin P, De Vries E (2008) How can we take into account student conceptions of the facial angle in a palaeontology laboratory work? In International Conference on Learning Science (p. 567). Utrecht, Netherlands, June, 23–28

Morgan K, Brooks D (2012) Investigating a method of scaffolding student-designed experiments. J Sci Educ Technol 21(4):513–522. doi:10.1007/s10956-011-9343-y

Neber H, Anton M (2008) Promoting pre-experimental activities in high-school chemistry: focusing on the role of students' epistemic questions. Int J Sci Educ 30(13):1801–1821

Ohlsson S (1996) Learning from error and the design of task environments. I J Educ Res 2(5):419–448. doi:10.1016/S0883-0355(97)81236-0

Ohlsson S (2002) Constraint-based student modelling. I J Artif Intell Educ 3:429–447

Pea RD (2004) The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. J Learn Sci 13(3):423–451

Quintana C, Reiser BJ, Davis EA, Krajcik J, Fretz E, Duncan RG, Soloway E (2004) A scaffolding design framework for software to support science inquiry. J Learn Sci 13(3):337–386

Reiser BJ (2004) Scaffolding complex learning: the mechanisms of structuring and problematizing student work. J Learn Sci 13(3):273–304. doi:10.1207/s15327809jls1303_2

Rodriguez E, Bosch M, Gascon J (2007) An anthropolical approach to metacognition: the "study and research courses". In Philippou DP-PG (Ed.), Fifth Congress of the European Society for Research in Mathematics Education. Larnaca, Cyprus, http://ermeweb.free.fr/CERME5b/

Séré MG (2002) Towards renewed research questions from the outcomes of the European project labwork in science education. Sci Educ 86(5):625–643

Séré MG, Beney M (1997) Le fonctionnement intellectuel d'étudiants réalisant des expériences: observation de séances de travaux pratiques en premier cycle universitaire scientifique. Didaskalia 11:75–102

Wood D, Bruner JS, Ross G (1976) The role of tutoring in problem solving. J Child Psychol Psychiatry 17:89–100