An integrated way of using a tangible user interface in a classroom

**Sébastien Cuendet[1] · Jessica Dehler-Zufferey[1] ·
Giulia Ortoleva[2] · Pierre Dillenbourg[1]**

Abstract Despite many years of research in CSCL, computers are still scarcely used in classrooms today. One reason for this is that the constraints of the classroom environment are neglected by designers. In this contribution, we present a CSCL environment designed for a classroom usage from the start. The system, called TapaCarp, is based on a tangible user interface (TUI) and was designed to help train carpenter apprentices. A previous study (Cuendet and Dillenbourg 2013) showed that the tangible nature of TapaCarp helped integrate it in the classroom environment, but that this did not guarantee a meaningful learning activity. In this article, we describe the process that led us to design a new learning classroom activity for the particular context of dual carpentry apprenticeships. One innovative aspect of the activity is that TapaCarp is used only for a small part of it. This contrasts with the mainstream CSCL approach that assumes that the system must be used from beginning to end of the activity. This new activity was used in a classroom study with 3 classes of carpenter apprentices over two days for each class. Despite its many steps, the activity proved usable and fostered many connections to the workplace, which was one of its main purposes. The teacher and the students were positive and showed high engagement in the activity. The learning gain results were mixed: the performance of the students improved from day 1 to

✉ Sébastien Cuendet
sebastien.cuendet@gmail.com

Jessica Dehler-Zufferey
jessica.dehler@epfl.ch

Giulia Ortoleva
giulia.ortoleva@unige.ch

Pierre Dillenbourg
pierre.dillenbourg@epfl.ch

[1] CHILI Lab, EPFL, Lausanne, Switzerland

[2] TECFA, University of Geneva, Geneva, Switzerland

day 2, but the learning gain measured with a pre-test/post-test mechanism did not show any significant difference compared to that of a control group.

Introduction

Technology for classroom orchestration

Although numerous projects have shown the potential benefits of technology for learning, and in particular for collaborative learning, implementations of these technologies in authentic teaching settings are few and far between (Roschelle et al. 2009; Dillenbourg et al. 2012). In fact, classrooms using computer technology are today the exception rather than the majority. This is not for the lack of potential of computer technology for learning. Computers can indeed support fundamental characteristics of learning, such as active engagement, participation in groups, frequent interaction and feedback, and connections to real-world contexts (e.g. Roschelle et al. 2000).

Dillenbourg and Jermann (2010) argued that the success of a learning technology in a classroom environment goes beyond the potential of the technology to deliver a learning outcome. The authors present six constraints that a teacher faces when managing a classroom: curriculum, assessment, time, effort, space, and safety. Others (Moraveji et al. 2011; Roschelle et al. 2009) have also observed the curriculum and assessment constraints, and added the following ones: pedagogical guidance; logistic support for teachers; coping with different levels of expertise among the students; and providing task-specific context to the teacher.

These extrinsic constraints were often ignored in the early days of CSCL, during which research tended to focus more on the learning outcome by satisfying primarily the constraints linked to the core pedagogical task (intrinsic constraints). Many scholars in the CSCL community now defend the idea that the extrinsic constraints should be given more attention. This has led to the recent focus on classroom orchestration (Dillenbourg et al. 2009). The core idea of classroom orchestration is that to be successfully used in a classroom, a learning technology should address both the constraints linked to the learning context and the intrinsic constraints (Dillenbourg et al. 2012).

Although the focus on classroom orchestration is quite recent, some learning systems have successfully been integrated in the classroom. ClassSearch (Moraveji et al. 2011), a tool for web search skills acquisition, was used in classrooms of 11-14 years old students. It reduced the teacher's orchestration load by displaying a search history on every student's screen and by encouraging students to stay on task. OneMouse Per Child (Alcoholado et al. 2011) is a single display groupware where each student controls a mouse linked to a single computer. It was used to train third grade students for arithmetic, geometry, and grammar. The teacher can easily monitor the activity of each pupil, which decreases the orchestration load. One of the main successes of classroom integration is Group Scribbles (Dimitriadis et al. 2007), a graphical user interface (GUI) that enables students and a teacher to scribble contributions on sheets similar to post-it notes, and to jointly manage the movement of these electronic notes within and between public and private spaces. The system increases awareness of students' actions for the teacher, making it easier to orchestrate the classroom.

This article follows up on the classroom orchestration and the integration of learning technology in the classroom. It is distinctive from previous work in that it focuses on a specific educational context, vocational education and training (VET), and that is uses tangible user interfaces (TUIs). Both are described in the following sections.

Educational context: Vocational training

The dominant approach to VET in Switzerland, where this study was conducted, is the dual one, where apprentices split their time between school and work. They spend one day per week in a professional school where they acquire profession specific knowledge. The practical knowledge is developed during the remaining four days spent in a company and during intensive practical courses that generally occur once a year over 2 to 3 weeks.

The dual system is ideal on paper: the firsthand practical experience acquired in the workplace is complemented by the theoretical lessons given at school. In practice, though, things are a bit more complex. Companies evolve in a fast and competitive environment and are required to constantly adapt their practices based on technological progress and industry standard changes. Schools, for their part, favor continuity based on a rarely changing curriculum and teachers are often not active in the profession anymore. For apprentices, this leads to a gap between the practical work in the company and the theoretical teaching at school. To be successful, a learning technology for VET training should therefore, in addition to satisfying the extrinsic constraints of the classroom described above, help overcome this gap. This can be achieved by putting together what is learned at the workplace and at school and making a link between the two.

Carpenters

In 2010, with 1060 apprentices starting an apprenticeship, carpentry was the 19th most popular profession by the number of apprentices starting an apprenticeship (Gaillard 2012). In 2011, 890 carpenter federal degrees were delivered (869 men and 21 women) and there were 2887 people enrolled in a carpenter apprenticeship (29 women), representing about 10 % of all apprentices in the building trades.

The job of a carpenter is to prepare, cut, and assemble wood pieces to create the frames and roofs of buildings (see Fig. 1). Carpenters work on new buildings as well as on older ones that they renovate. Their job mainly consists in five steps:

1. Read and make sense of the plans produced by the architect (or the engineer).
2. Generate a working drawing based on the plans, at a 1:1 scale.
3. Determine how much wood will be needed, select and prepare it.
4. Cut the various pieces of wood as indicated on the plan.
5. Go on the construction site and assemble all the pieces together.

Companies and professional schools both put forward spatial skills as a crucial capability asked from carpenters. Tapacarp was designed to support carpenter apprentices particularly regarding the acquisition of spatial skills.

Technological context: Tangible User Interfaces

Apprentices in VET are often interested in practical tasks, especially in manual professions. We therefore believe that they should be provided with learning environments that begin with concrete tasks and concrete objects before introducing more theoretical and

Fig. 1 Typical structures built by carpenters.

abstract concepts. Tangible user interfaces (TUIs) seem particularly adequate in the context of vocational training as they offer concrete manipulation of physical objects.

According to Shaer and Hornecker (2009, p. 4), a tangible user interface (TUI) is an "interface that is concerned with providing tangible representations to digital information and controls, allowing users to quite literally grasp data with their hands". The original motivation behind TUIs was indeed to connect the physical world with the digital one by using physical artifacts, therefore keeping the richness of physical interactions. This was novel and contrary to the main trend that focused on forcing the user into a virtual world. An exhaustive panorama of TUIs is beyond the scope of this article. Instead, we review the work on TUIs for education.

TUIs for education

Education has been one of the main domains of application for TUIs, resulting in many TUI prototypes for various learning domains. An excellent review of tangible interfaces in the domain of learning is available in the survey of TUIs provided by Shaer and Hornecker (2009, chap. 4.1). It references many recent TUI applications for collaborative problem solving, tangible programming, and storytelling. The focus here is on TUIs that are closer to the one mentioned in this article, i.e. TUIs that concern the creation and manipulation of three-dimensional (3D) shapes, the development of geometrical and spatial related concepts, and TUIs that are used on interactive surfaces.

Several systems have been developed that allow users to build 3D models through the manipulation of physical blocks (Girouard et al. 2007; Anderson et al. 2000). Song et al. (2006) developed an interface that allows users to annotate physical 3D models and integrate the annotations back into a CAD software. Several works have focused on exploring TUIs and spatial skills. Kim and Maher (2008) showed that designers using the TUI (as opposed to a traditional GUI) recognized more spatial relationships, were more immersed in the task, and discovered new visuo-spatial features when revisiting their design. Quarles et al. (2008) compared the use of a TUI, a GUI, and a physical interface to learn how to operate an anesthesia machine. They found that the TUI significantly helped users with low spatial cognition ability, whereas the physical interface and the GUI did not. They noted that by merging physical objects and computational media, the TUI made it easier for users to perform spatial cognition tasks.

The TUI presented in this article is a top-down interactive surface. An interactive tabletop surface is a horizontal surface that serves both as an input and output space. It is in general composed of a camera and a projector which can be placed either above the surface or under it. As detailed in Dillenbourg and Evans (2011), there are advantages and drawbacks

to both approaches. One of the most fully developed top-down tabletop systems is TinkerLogistics, created by Zufferey et al. (2009) and later improved by Do-Lenh (2012). The aim of the system is to teach logistics to teenaged apprentices by allowing them to build warehouses and evaluate them. The system is composed of small-scale shelves that the users can freely manipulate. Each shelf is equipped with a fiducial marker which allows the system to track its position accurately, build the model of a warehouse, and project feedback on the shelves. The users control the system through a paper interface. They can then run simulations and observe various metrics reflecting the performance of the warehouse. Contrary to constructive assembly systems such as Smart Blocks or Digital Manipulatives, the shelves are simple objects that do not contain any electronic equipment.

Another top-down system, without augmentation, is the one created by Horn et al. (2009) for practicing tangible programming. The system is aimed at kids who, through tangible puzzle pieces, can put together several programming instructions to direct a robot. The system identifies the puzzle pieces through a camera, interprets them into program and command a robot to execute the program.

Many other tabletop TUIs for learning have been developed over the years. Rather than listing them all, the next section summarizes the potential benefits of such systems for learning.

Potential benefits for learning

There are many reasons why tabletop TUIs could benefit learning. The most relevant ones in the context of this study are increased usability, physicality and its link with cognition, multiple external representations, and the impact of TUIs on collaboration, co-location, and simultaneous interactions.

Increased usability The learners' familiarity with the manipulation of physical objects can increase usability and limit the cognitive effort dedicated to interact with the system, and subsequently allow learners to focus on the core of their task (Manches and O'Malley 2012; Fitzmaurice and Buxton 1997). Tuddenham et al. (2010) found that tangible objects were easier for the user to acquire than multi-touch or mouse, and that once acquired, they led to an easier and more accurate manipulation. However, the ease of manipulation can sometimes be counterproductive for learning if it leads to too much manipulation (Cuendet et al. 2012c; Do-Lenh et al. 2010; Price et al. 2009). In the context of a classroom, where time is a scarce resource, increased usability and intuitiveness of manipulating objects can be decisive.

Link with cognition and spatial tasks Assuming that perception and cognition are linked (e.g. Pecher and Zwaan 2005), gestures and physical movement can benefit learning (Goldin-Meadow 2003; O'Malley and Stanton Fraser 2004). Physicality can also be beneficial for specific tasks, and especially for spatial ones, thanks to their haptic and proprioceptive properties (Sharlin et al. 2004; Marshall 2007).

Multiple external representations Zhang and Norman (1994) showed that external representations can lead to higher problem-solving performance, because they can reduce the load on working memory, anchor and structure cognitive behavior, and provide information that can be directly perceived and used without further processing. Tangibles can serve as external representations, and especially so in the case of spatial reasoning tasks, where they can serve to anchor the spatial reasoning (Cuendet et al. 2012b). Price et al. (2009) also

found that different geographical locations (i.e. co-located vs. discrete) of a representation led to an understanding of concepts at different levels of abstraction.

The potential benefits of multiple external representations (MERs) for learning are summarized in Ainsworth (1999) who argues that MERs can help learning by providing complementary processes and information, by constraining interpretation, and by helping to construct deeper understanding. However, the dosage of cognitive load of MERs upon the learner has to be well chosen. Using an additional representation leads to increased learning if it is used to constrain interpretation, but results in lower learning when the additional representation merely complements other representations.

Collaboration, co-location, and simultaneous interactions Tangible tabletops provide a shared workspace for their users. Having a shared workspace helps learners be aware of each other's actions, which in turn can lead to fluid collaboration and interaction (Hornecker et al. 2008; Ha et al. 2006; Rogers and Lindley 2004; Hornecker and Buur 2006). Rogers and Lindley (2004) found that when using a horizontal display learners generated more ideas than when using a vertical display. Hornecker et al. (2008) compared multi-touch and multiple mice and found that the multi-touch condition led to more actions that interfere with each other, but that interference was quickly resolved, leading to higher levels of awareness between participants. Ha et al. (2006) reported a similar result comparing stylus, mouse, and touch input: the touch input led to a higher awareness of each other's actions, but some participants expressed concern about physical interference and collisions. Tangibles are however not always promoting collaboration. Marshall et al. (2009) found that children resolved conflicts by moving physical objects out of reach and conclude that the tangibility of an interface can break collaboration rather than foster it. Finally, Speelpenning et al. (2011) observed that TUIs compared to multi-touch can increase the negotiation of objects but that usage of the interface depended greatly on group dynamics and not only on the type of interface

TUIs and classroom orchestration

The notion of classroom orchestration with a TUI in a VET context was first addressed by Zufferey et al. (2009) and further developed by Do-Lenh (2012), who created the TinkerLogistics described above. Do-Lenh (2012) showed that the system positively impacted learning in the classroom mainly through the development of appropriate orchestration tools. Those tools included a card with which the teacher could instantly pause the system of all the groups and a large display on the wall showing the progress of each of the groups. The pause card allowed the teacher to quickly get the attention of all the students, whereas the board allowed him to monitor all the groups from a distance. Topobo (Parkes et al. 2008), a construction kit with kinetic memory, has also been widely deployed into classroom. Results of a longitudinal study showed that Topobo was especially useful in promoting collaboration and cooperation between children. Educators experiencing the use of this tool concluded that they needed prior training and some set of exercises prepared for them that matched the curriculum. A similar constructionist approach is used by Blikstein (2008). The basic idea behind his work is to allow students to explore electronic concepts in the context of the classroom with material that match their every day life and socioeconomic background. This approach was used successfully in several classrooms in Brazil.

The TUI presented in this paper, TapaCarp, was previously tested in the classroom as well (Cuendet and Dillenbourg 2013). TapaCarp was designed in a participatory way with carpentry teachers and apprentices, with the goal that it should be integrated in existing

pedagogical practices at the school. The system was tested in the classroom regarding the usability and the learning performance. Results on usability were globally positive and showed that there was potential for such a system in a classroom. As a result of this study and other similar ones, Cuendet et al. introduced 5 design principles to reduce the orchestration load (Cuendet et al. 2012a): integration, flexibility, empowerment, awareness, minimalism. Their definition is shown in Table 1.

The integration of TapaCarp into existing practices and activities followed a request from the teacher and allowed TapaCarp to be accepted and deployed in the classroom. However, the activities done with TapaCarp were so close to the ones done in the regular curriculum that one can wonder if TapaCarp was really a valuable contribution. If the activities are so similar, why could they not be done without TapaCarp? Could it not be more valuable to use TapaCarp and its unique features to design new pedagogical scenarios that could not be achieved without it? These questions led to the design of a new learning activity with the aims to fit into the particular context of carpenters' vocational training while keeping the orchestration load low. The next section describes the activity in more detail, as well as the study in which it was evaluated.

Experiment and method

Activity

Context and goal

A key element in carpentry is the competency to read and write plans. Plans tell carpenters how to cut the beams and how to put them together. On a construction site, the plan is the only reference and communication means between all stakeholders. Currently, carpenter apprentices develop reading and writing plans skills during drawing classes at school. The plans used in these classes serve mainly as a technical exercise and a means of assessment for the teacher. The objects they feature do often not have any carpentry meaning, but are instead abstract shapes (such as a truncated pyramid) that are used because they allow to highlight a particular geometrical difficulty. Moreover, these plans are (usually) not used to build an object. This is justified by the limited time available in the classroom and by the absence of a workshop in it, but is nonetheless different from the workplace, where the plan is a means to the creation of an object rather than an end in itself.

Table 1 The 5 design principles from Cuendet et al. (2012a)

Integration	Orchestration load decreases if the learning environment is integrated in the workflow.
Empowerment	Orchestration load decreases if the learning environment allows the teacher to keep a central point in the classroom interactions when it is necessary.
Awareness	Orchestration load decreases if the learning environment provides the teacher with permanent awareness of the state of all students in the class.
Flexibility	Orchestration load decreases if the learning environment is flexible enough to adapt the activities to the evolution of the scenario (e.g. the state of students, the time remaining) and to accommodate unexpected events.
Minimalism	Orchestration load decreases if the learning environment does not provide more information (and functionalities) than what is required at a given time.

Making a connection with the workplace is a key constraint of the VET dual carpenter apprenticeship system. The main goal of the activity presented in this article is to match this constraint while also respecting the usual extrinsic constraints linked to the classroom (time, space, noise, etc.). For example, bringing a band-saw in the classroom is not an option (noise constraint). Neither is asking carpenters to work on real-world objects over several weeks and cut them in their company workshop (time constraint).

The activity creates a connection to the workplace by giving the plan the same role it has at the workplace: a means of communication and action. It includes the full workflow of carpentry, from the design of the object to the production of its final physical artifact, but does not require a full-fledged workshop. Importantly, although its duration varies depending on the apprentices' level and on the difficulty of the objects at hand, it should be feasible in a lesson's time (45 minutes). It therefore meets the "workplace" constraint while not breaking the time constraint.

Besides meeting constraints, the activity should (obviously) also help carpenters learn. Here are some reasons why this should be the case. First, the activity includes the entire workflow of carpentry in a short period of time, which should allow apprentices to make a link between the various steps of their work (drawing, marking out the wood, cutting). Second, performing both a school activity (drawing) and creating a final object as done at the workplace helps them link the workplace and the school. It allows them to make sense of what they learn, which is commonly referred to as meaning-making (e.g. Suthers 2006). Third, the confrontation created by the SWISH scenario (described below) should benefit learning (e.g. Butterworth and Light 1982; Vygotsky and Cole 1978), by forcing them to find the source and cause of a mistake and to negotiate until they agree. Finally, the class debriefing allows them to learn from their own and their peers' mistakes and to learn by inducing rules from a selected set of objects (contrasting cases, Schwartz et al. (2011)).

The collaboration between students is designed according to the "Split Where Interaction Should Happen" pedagogical script (SWISH, Dillenbourg and Hong (2008)). The idea of SWISH is to split a task between group members, thereby forcing students to interact with each other to complete the task (Fig. 2). Learning results from the effort to build a shared solution and understanding.

In a carpentry company, the plan is made by a specialized carpenter. The other carpenters must then interpret the plan and mark out the wood to indicate how it must be cut. This is therefore a natural point for the SWISH, and we chose to put the split on the transition from the plan to the marked out object.

Description

The activity is composed of two phases. The first phase involves a pair of students at a time and is composed of 5 steps, described in more detail below and summarized in Table 2;
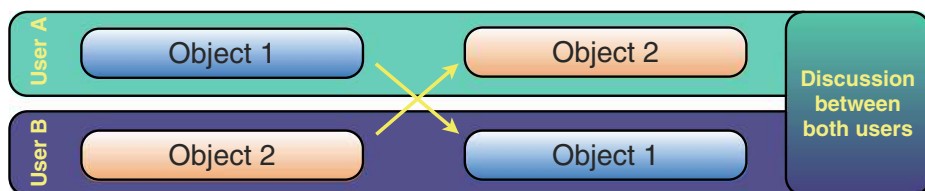

Fig. 2 A graphical representation of the SWISH script

Table 2 Description of the first phase of the activity. The bold line marks the split of SWISH

Step	Name	Description	Timing
1	Cut object virtually	Using TapaCarp, define cutting planes.	5–15'
	Print plans	The three orthographic projections of the object are printed on an A3 sheet of paper. The views include the cutting planes done in step 1. All printed lines are stipple and very light.	2'
2	Improve plans	The apprentice must improve his own plan so that it is understandable by his colleague.	5–15'
	⟷ **Exchange plans**	**The plans of the students are exchanged.**	–
3	Mark out block	Based on the plan received from his colleague, the apprentice marks out a block.	5–10'
4	Cut block	Based on the marks he made, the apprentice cuts the block.	5'
5	Within group peer assessment	The two apprentices are brought together to compare the objects.	10'

most of them are also illustrated in Figs. 3 and 4, and in Video 1[1]. The second phase is a class debriefing (Fig. 4).

1. Cutting the object with TapaCarp The apprentice is prompted with a perspective view of the object he has to cut (one of the objects shown in Fig. 5). This object is different from the one given to his partner, and an apprentice has no knowledge of the object given to his partner. The perspective view given to them does not include any dimensions. The apprentice must therefore place the position and orientation of cutting planes so that the general shape of his object will resemble the object shown to him, but cannot be blamed for small inaccuracies. At the end of this step, the plan is printed.

2. Improving the plan The printed plan displays the cuts made on TapaCarp, but in a "dumb" way: all the cuts are transverse to the object and are displayed with light stipple lines. That is, the machine does the tedious job of creating a draft of the three orthographic projections with the right dimensions, and it is the student's job to decide and indicate what line should be solid, stipple, or simply non existing. These line properties provide necessary information to make out the shape of the object. An example of a raw plan is given in Fig. 6a with the corresponding plan improved by a student (Fig. 6b).

3. Marking out the object After they have been improved, the plans are exchanged. Based on the plan received, the student marks out a polystyrene block so that he can cut it (Fig. 3c). A key aspect of the activity is that the two apprentices paired together are working on two different objects and have no knowledge of their partner's object. They do not know what the shape of their partner's object is, and the only way for them to determine this is through their partner's plan.

4. Cutting the block Each student cuts out the object he has marked out, using a hot wire cutter and/or a utility knife (Figs. 3e and 3d).

---

[1] http://www.youtube.com/watch?v=1S3gURKUjSA
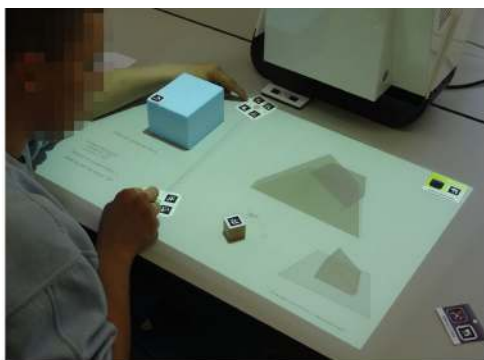
(a) Cut object virtually with TapaCarp.
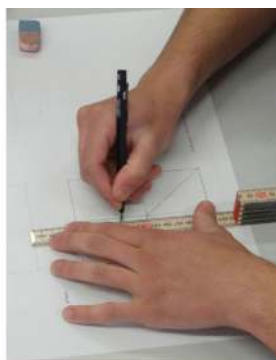
(b) Plan improvement.

(c) Marking out the block.

(d) Knife cutting.

(e) Hot wire cutting.

(f) Group discussion.

(g) Debriefing preparation.

Fig. 3 Steps of the first part of the activity

5. Within group peer assessment Pairs meet with their objects and plans and compare the objects with the printed perspective view that served as a prompt (Fig. 3f). A double-sided A4 paper sheet that shows a perspective view of the desired object as well as the correct three orthographic views supports the discussion. The evaluation is done mostly

Fig. 4 The second part of the activity: class debriefing

graphically by circling the differences between the produced object and the desired one. For each mistake they find, both students must agree on where the mistake stemmed from (the plan or the cutting), and must explain the reason of the mistake and how it could be avoided.

After all groups have completed these 5 steps in the first phase, the teacher collects all plans and objects and evaluates their accuracy, i.e. their ressemblance with the initial object. Then, all apprentices come together for a class debriefing where the teacher discusses typical errors on the plans and on the objects with the apprentices and together they identify solutions for avoiding those mistakes.
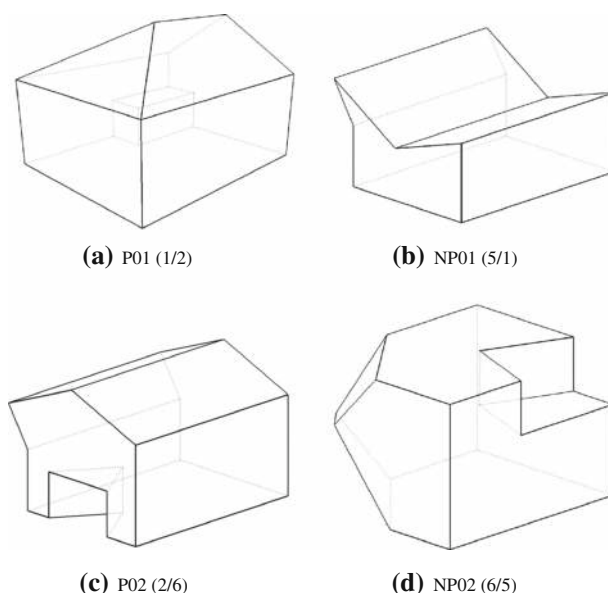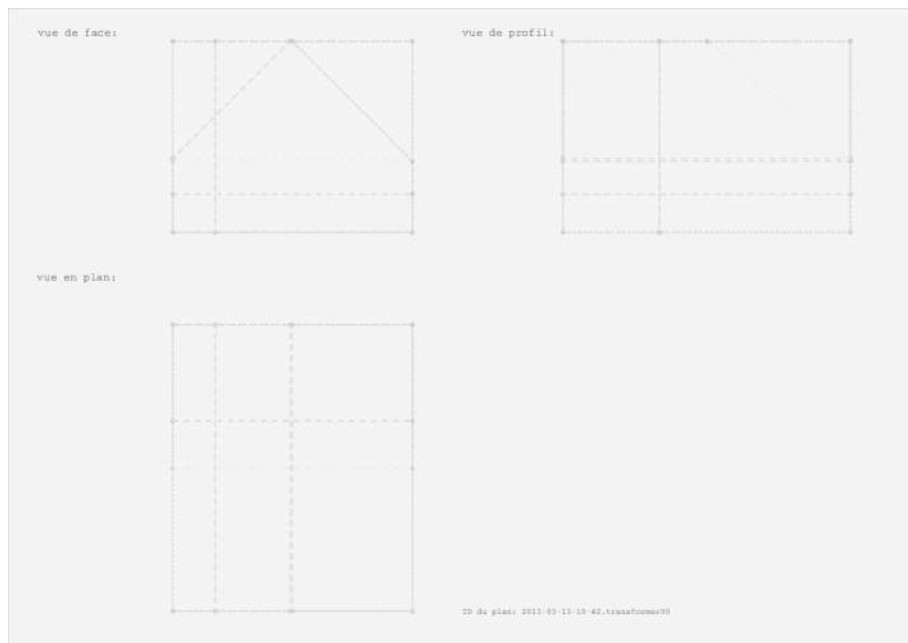

(a) P01 (1/2)

(b) NP01 (5/1)

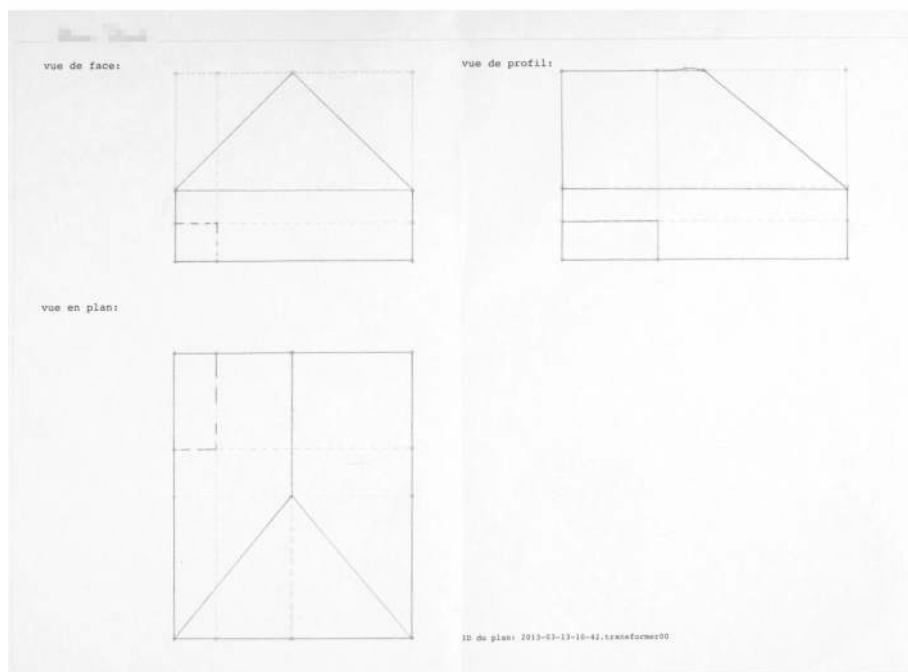**(c)** P02 (2/6)

(d) NP02 (6/5)

Fig. 5 The four objects that the participants were asked to reproduce. The name of the object indicates whether it was plausible (P) or non-plausible (NP), and whether it was used on the first (01) or second day (02). The numbers in parentheses are the plausibility and difficulty scores, respectively

(a) The plan saved.

(b) The plan once printed and improved by the apprentice.

Fig. 6 The plan (a) saved by TapaCarp; (b) the same plan improved by an apprentice

Class debriefings

Class debriefings were prepared by the teacher, with the help of a researcher, by examining the objects and the plans produced by the apprentices (Fig. 3g). All the objects – each object together with its plan and the group discussion sheet – were placed next to each other. They were sorted by object type, and then by correctness. The teacher made a list of points to discuss based on the objects produced by the apprentices with the help of the researchers. This is important: although some general points came up in each of the debriefing sessions, each debriefing session was unique because it was based on the objects produced by the apprentices, and the mistakes that they made. The teacher led the debriefing session and engaged apprentices by asking them questions.

The class debriefing is a key part of the activity. The students receive a feedback on their performance and get the explanations on why they failed, if they did. Even if they do not, they do reflect on the mistakes made by others. However, it is also a difficult activity for the teacher, because he has little time to prepare it. He also needs to adapt his message as much as possible to what the students have done because a tailored feedback is what helps the students understand their mistakes and build their knowledge.

Research questions

The main purpose in designing the new activity was to bridge the gap between the school and the workplace while keeping it usable in the classroom. This can be translated into the following research questions:

- Does the activity help create a connection with the workplace context?
- Is the activity usable in a classroom environment?

Measuring learning was not the foremost goal of this study: the environment is not controlled, the number of concepts to learn is large given the short activity, and time to assess learning is limited. Nevertheless, this is a learning activity and we must try to assess learning. Additionally, we also study the impact of plausibility of objects on the learning gain. Working with plausible objects may be more motivating and seem more relevant, but might also be easier, thereby reducing the learning gain. This leads to the two following secondary research questions:

- What is the impact of the activity on the learning performance?
- Are plausible objects more efficient for learning than non-plausible ones?

Tapacarp

TapaCarp, is a top-down camera-projector tabletop system. The projection area is about 50 cm by 70 cm. Users interact with the system by manipulating objects equipped with fiducial markers. These objects are tracked by the system thanks to a tag tracking library (Bonnard et al. 2013). and are the tangible interface through which users interact with the system. More information on TapaCarp and its design can be found in Cuendet (2013).

For this activity, the interface given to the apprentices was composed of a polystyrene block, two "selectors", one token, and five cards (Fig. 7a). Using this interface, the users had to create virtual cuts in the polystyrene block. To do so, they positioned the cutting plane around the polystyrene block using the two selectors (Fig. 7b), and adjusting its orientation
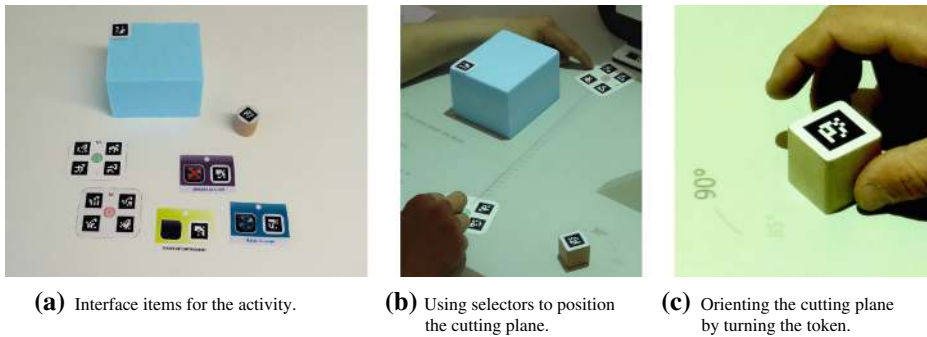
(a) Interface items for the activity.   **(b)** Using selectors to position the cutting plane.   **(c)** Orienting the cutting plane by turning the token.

Fig. 7 The components of the interface, and the manipulations of the cutting plane

by turning the wooden token (Fig. 7c). The cards are used to utter actions, such as validating or removing a cut.

One of the researchers kept some cards: the cards to quit the activity and launch the activity, and the card to print the plans (shown on the right in Fig. 7a). The card to print the plans terminates the activity and creates a file containing the orthographic projections of the cuts done by the student (Fig. 6a). The typical display of the interface for this activity is shown in Fig. 3a. The block is placed on the left side slightly above the horizontal median of the workspace. Below the block, a summary of the cuts made so far, as well as information about the current cut, are displayed. On the right of the block, a preview of the block is shown that includes the cuts already made and the one being made. On the bottom right, the object as it would be cut by the current cut is shown. The lower central area is left blank so that users can manipulate the cards and the token there.

Participants

Forty second-year carpenter apprentices, spread over three classes, participated in the study which took place at the practical school. Apprentices typically work in pairs at school and these pairs were kept for the study. The apprentices were rewarded for their participation with a chocolate bunny. Because the testing material was designed specifically for this experiment, a fourth class composed of fourteen second-year apprentices took the pre-test and post-test to control for potential levels of difficulty between the two tests. These participants simply took the pre-test and post-tests, and were given a one hour drawing task in between the two tests. We refer to them as the control group.

In the control condition, because the apprentices simply took the test, the teacher did not play any significant role. In the experimental condition, students work on their own under the technical supervision of two researchers. The teacher, who was also taking care of the rest of the class, was present about half of the time. He prepared the classroom debriefings with the two researchers and conducted them on his own.

Flow of the experiment

The study took two days for each of the three classes. The first day, the two experimenters explained the global flow of the study (1 minute), demonstrated the use of TapaCarp (3 minutes), and administered a pre-test (15 minutes). Pairs of apprentices went through the 5 steps of the group activity phase consecutively (Fig. 8). When all pairs had finished, the

teacher and the two experimenters prepared the class debriefing session (10 to 15 minutes). The apprentices were then brought back in for the debriefing session (no time constraint imposed, but ended up lasting between 10 and 20 minutes). The procedure during the second day was similar to the first day, except that the objects were more complex, the debriefing session was followed by the post-test, and there was no pre-test in the morning.

Tests

The questions in both tests were of the same kind but with a different object. All questions were built around an object, and there were 4 questions per object. The first and second question tested the ability to match the three orthographic projections with a perspective view and to correct them if a view was wrong. The third question asked whether all three views were necessary to be able to cut the object. The last question showed how the object had been cut based on the plan given in question 1, and asked whether the object was correct. If not, participants were to tell whether the mistake came from the plans or from the cuts.

Material

The objects that the participants had to reproduce are shown in Fig. 5. There were four objects in total: (a) and (b) were used on the first day of the experiment, and (c) and (d) on the second day. The objects were classified into one of two categories according to their degree of plausibility: (a) and (c) are *plausible*, whereas (b) and (d) are *non-plausible*. Plausible objects are those that carpenters could expect to find in the real world, typically a house-like shape. Non-plausible objects, are objects that are not necessarily more difficult to create, but whose shape would seem less natural in the construction world. These objects will be referred to as <plausibility><day of use>, e.g. P01 for the plausible object of the first day, or NP02 for the non-plausible object of the second day.

To determine the plausibility of objects, 13 objects were rated by 10 persons from our laboratory on a plausibility scale from 1 (plausible) to 7 (non-plausible at all). The plausibility was then defined as the median of all the ratings. Besides, a difficulty score was computed which considered the following criteria: the number of cutting planes; the number of axes of the cutting planes (i.e. the number of cutting planes not parallel to each other or to an edge of the object); the number of non-transverse cut; the number of new edges (created by a cut) not parallel to any other edges of the objects. The final difficulty score was normalized to range on a scale from 1 to 7.

The four objects chosen for the experiment were two objects with low difficulty for the first day, and two objects with a higher difficulty for the second day. On each day, one object had to be plausible while the other had to be non plausible (see Fig. 5; plausibility and difficulty scores are shown in parentheses).

The participants were randomly assigned to the objects on each day and we tried to balance them between the 4 possible combinations (P01/NP02;NP01/P02;P01/P02;NP01/NP02).

Data collection

Data were collected in various ways. The plans, final objects and group self-assessment sheets were kept for later analysis. All the manipulations done on TapaCarp were saved into log files. TapaCarp's camera was also used to save one snapshot of the student's workspace every second. The group peer assessments and the classroom debriefings were audio and

video recorded using a Sony Bloggie video camera equipped with a fisheye lens. Semi-structured interviews were conducted with the teacher at the end of each day and students were asked for a global feedback on the activity after the class debriefing. Observations were also made by the two researchers that were in the classroom.

For the statistical analysis, we checked that the variances of the various groups were homogeneous (homoscedasticity) and that the data were normally distributed, by using Bartett's and Shapiro's tests, respectively, and plotting the quantile-quantile (Q-Q) plot.

Results

Usability of the activity

The activity "worked well". All apprentices were able to complete the activity within 45 minutes on the first day and less than 40 minutes on the second day. In general, two or three groups worked simultaneously (each at a different stage of the activity). This allowed the entire activity, including the class debriefing, to be done in half a day.

There was great variability among apprentices in the usage of TapaCarp on the first day (they were using it for the first time). Most of them had no problem using it but two of them had to start over once. The apprentices had no problem understanding the rest of the activity (drawing, cutting, discussing). The apprentices worked independently and quietly. The idle time was in general short except in a few cases where one apprentice had to wait a few minutes for his partner to finish drawing the plan before being able to continue. The difference in time completion was due to difficulty with TapaCarp or to a level difference between students.

The main difficulty was with the group discussion: some pairs started heated discussions right away, inspecting the objects and noticing every minor mismatch with the original prompts, while others did not see any problems (even when the manufactured object had major flaws) and only started talking after a researcher prompted them to do so.

The informal feedback gathered from students was largely positive. They reported liking working in pairs and having to comment on each other's work and using this "cool technology". Only two apprentices said they did not understand why this technology was used and that they would rather draw the plan directly instead of first designing it with TapaCarp.

Fig. 8 The physical organization of the classroom in stations increases awareness.

Artifact analysis

The correctness of the objects produced serves as an indicator of the performance. The correctness of the object was scored according to four levels: the object did not resemble the target object at all (score: 1); the object had one major mistake but the object could be recognized (2); the object was mostly correct, but contained some inaccuracies (3); the object was correct. The scoring was done by two researchers (4).

The details of the correctness score are shown in Table 3. The global correctness score was 2.76 and was similar on both days (2.8 on the first day and 2.73 on the second one). The global score between the plausible and non-plausible objects was also similar overall (2.73 for plausible versus 2.8 for non-plausible). However, the average score differed depending on the objects: NP01 and P02 had a score higher than 3 on average, whereas the two other objects had an average score of 2.35. While the average correctness hints that there was not a big difference between the two days, the distribution of the results tells a different story. As can be seen in Fig. 9, which shows the distribution of the correctness for each of the four objects, objects that had a similar average correctness score did not necessarily show a similar correctness pattern. NP02 was the only object that had scores ranging from completely wrong to completely correct. None of the groups managed to build P01 correctly, while P02 was never completely wrong, and NP01 was either correct or contained only minor mistakes.

The objects for the second day were more difficult, and the similarity of the performance on both days suggests that the apprentices performed globally better on the second day. Looking at the number of correct objects produced (those with a score of 4) corroborates this observation: 13 correct objects were built on the second day versus only 5 on the first day, and this despite the higher level of difficulty. This is a significant difference ($F[1,78]=4.74$, $p=.03$).

Class debriefings

Each class had two debriefing sessions, one at the end of each day. In total, for the three classes, 6 debriefing sessions took place. Each of them lasted for 12 to 20 minutes (no time

Table 3 Details of the correctness of the objects and the plans

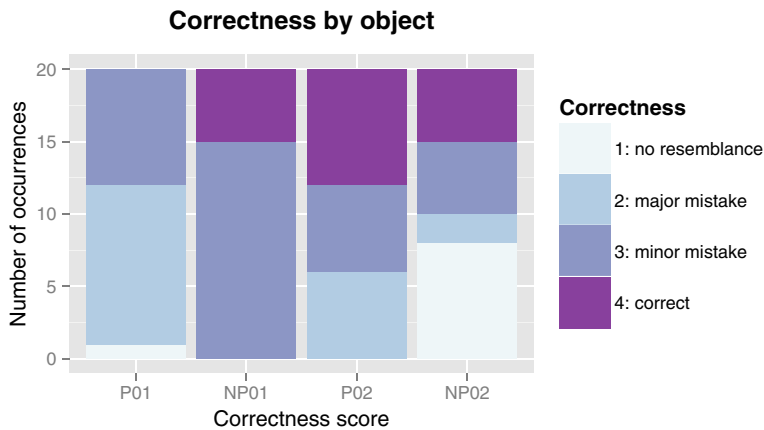|	Object correctness score					Plan
	1	2	3	4	average	# correct
P01	1	11	8	0	2.35	1
NP01	0	0	15	5	3.25	5
P02	0	6	6	8	3.1	6
NP02	8	2	5	5	2.35	4
plausible	1	17	14	8	2.73	7
non-plausible	8	2	20	10	2.8	9
day 01	1	11	23	5	2.8	6
day 02	8	8	11	13	2.73	11
overall	9	19	34	18	2.76	16

Fig. 9 Correctness score by object.

limit was imposed) and involved all students of a class simultaneously. The researchers did not intervene, unless asked a direct question by the student or the teacher (this happened in two sessions). While all the sessions were different from one another, the same subjects came up in several sessions, as shown in Table 4.

Lack of observation

One subject that was mentioned on each of the first day sessions was the lack of observation. Only 5 out of 40 objects were completely correct, and the cause for most of the mistakes was in the interpretation of the prompt object. The teacher pointed out that, despite the absence of dimensions – the reason given by students to justify this kind of mistakes – they should be able to observe and reproduce an object accurately.

Link with workplace and plan design

The teacher made links with the workplace very frequently in each session. For instance, about the lack of dimensions, he said: "When the client comes to your company, he comes with an idea. Even if he puts dimensions, they might be vague. We have to interpret so that the final object satisfies the client's requirements. So in this activity you have to analyze and observe what you see." He even did so with drawing practices. When he asked why they had drawn all three orthographic projections of an object (two only were necessary), students said "this is how we have learned to do it in the drawing class". The teacher then pointed

Table 4 Summary of the subjects discussed in each of the debriefing session. The name of the session is the class identifier followed by the day of the session

A-01	A-02	B-01	B-02	C-01	C-02	
lack of observation	✓		✓		✓	
plan design	✓		✓	✓	✓	✓
link with workplace	✓	✓	✓	✓	✓	✓
plan reading	✓	✓	✓	✓	✓	

out that when they draw a hip rafter at work, they only draw two views. Many of the links to the workplace came up when discussing the importance of the plan in the real world. Four main areas were pointed out: to include only the necessary information, to respect drawing conventions, to be consistent, and the importance of the plan as a means of communication.

Plan reading

Although most of the mistakes came from the plan, some correct plans also led to wrong objects. The teacher reminded his students that "you do not interpret a plan, you read it". He made the link with the drawing classes that they attended in school and recommended that they use this knowledge to apply a structured step-by-step approach to mark out the object and cut it.

Tests performance

There was no significant learning gain difference between the control and the experimental group (t[25]=.13, p=.89). The control group performed slightly worse at the post-test (66.5 % versus 73.1 %), but this lower performance could already be observed in the pre-test (52.4 % versus 59.3 %), and the relative learning gain (RLG) was comparable (28.7 % versus 29.9 %). The global positive learning gain of the control group could indicate that the post-test was easier than the pre-test, or that there was a learning effect due to the pre-test. In any case, students were faster at solving the post-test (10 minutes on average versus more than 14 minutes for the pre-test). The lack of a learning difference between the two groups is obviously disappointing, because it means that apprentices who took part in our activity during two days did not improve more than students who did not follow any dedicated training. Several possible explanations for this will be discussed below in the discussion section.

The effect of plausibility

The better performance on the second day is really only noticeable for the plausible objects: 8 were completely correct on the second day, versus none on the first day. This improvement is quite compelling given that P02 was more difficult than P01. Its cause could be that apprentices learned the task or that they learned a skill, or both. Whatever the reason, it is interesting to identify whether it came from the plans or from the cutting. The source of the error can be on the plan, on the cut, or on both. Table 5 shows the source of the error for each object. For plausible objects, the reduction of mistakes between day 1 and day 2 came mostly from better plans (17 plan mistakes on day 1 versus 8 on day 2).

Table 5 Source of errors for each type of object

	plan	cut	shared	total
P01	17	0	3	20
P02	8	3	1	12
NP01	14	0	1	15
NP02	7	3	4	14

It is harder to assess the performance evolution for non-plausible objects. On the one hand, the total number of mistakes remained similar and the number of fully correct objects was equal on both days. On the other hand, the average correctness score went down by almost a point between day 1 and 2. More specifically, all NP01 objects were correct or contained only a small mistake, whereas on the second day half of the objects were significantly wrong (8 objects were completely wrong and 2 contained a major mistake). This hints that the gap in the task difficulty between two objects was bigger for the non-plausible objects than for the plausible ones.

One explanation is that the criteria used to assess the difficulty level of an object do not capture everything related to the difficulty level. While this is possible, it is hard to find additional criteria that could reflect better this gap of difficulty. A more likely explanation is that there is an interaction effect between the plausibility and the difficulty of an object: once a certain level of difficulty is crossed, the plausibility becomes a crucial factor. This can be seen with the plans of the two difficult objects (Fig. 10). Although there were fewer cuts on NP02 than on P02, the P02 plan (Fig. 10a) is easier to interpret for a human brain that has been trained to read plans of houses than the plan of NP02 (Fig. 10b).

When asked about their strategies to create a correct object, successful groups said that they had to imagine how the three views would look like based on the perspective view. During a class debriefing, two apprentices explained that they made sense of plans that had many cuts by creating a mental model of the object; they added that making a mental model of a house was easier than making one of an object that is unfamiliar. In other words, Fig. 10b by itself is hard to decipher, but once one knows what NP02 looks like, it is not more difficult to create than P02. This could explain why the objective criteria did not capture accurately the difficulty level: only criteria related to the number and nature of cuts were taken into account. These might be valid criteria for computers, but for humans the plausibility of the object should be factored into the difficulty score, because the first step made by humans when solving such a problem is to try to imagine the 3D representation of the object.
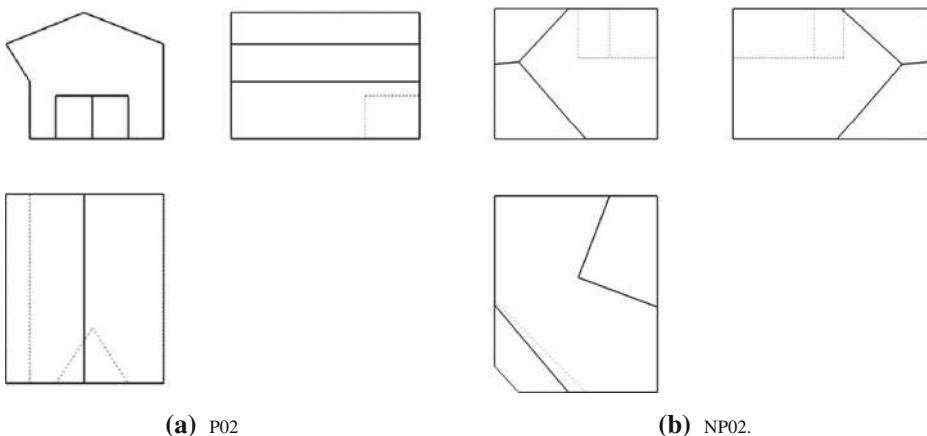

(a) P02   **(b)** NP02.

Fig. 10 Orthographic projections for the objects of the second day

Discussion

The goal of this study was to assess the new learning activity designed with TapaCarp and to do it in a classroom environment. We aimed to test the usability of the activity and whether it helped create a connection to the workplace. Additionally, we also sought to examine the learning performance of students and whether varying the plausibility of the objects influenced learning. We now discuss these four aspects in light of the results reported above.

Usability of the activity

As explained in the introduction, one worry when introducing technology in the classroom is the potential increase of the orchestration load. We use Cuendet et al.'s 5 design principles to analyze the impact of the activity on the orchestration load.

The **integration** principle was, by design, only partially respected. The activity was meant to be of a new sort, using technology for part of it and integrating with existing practices (e.g. drawing plans on paper) for the rest. It nevertheless succeeded in integrating into the larger context of dual training by being executable within a day and connecting a school activity with the workplace (see next section).

Having to pair students and the SWISH script limited **flexibility**. This manifested when some students were faster than their partner and had to wait several minutes without work to do. On the other hand, there was a lot of flexibility for the teacher in analyzing and organizing the artifacts for the debriefing session. Objects were easily compared and when mistakes were made, the teacher could easily identify their source by having both the plan and the group debriefing session close to it.

Empowerment was obvious in the class debriefings. The teacher had the full attention of the students who were eager to understand their mistakes. During the activity, the print card was not given to the students, forcing students to call the teacher before they could print their plan and therefore empowering him. In contrast, the need for a mediator to foster comparison during the group peer assessment indicates that in the absence of a mediator, an empowerment mechanism would probably be required for this part of the activity.

The physical organization of the classroom ensured great **awareness** for the teacher: he could see where students were in the workflow of the activity. Awareness was also high during the classroom debriefing with all the objects readily visible and reachable.

Finally, a special effort was made to keep the number of features and objects minimal in the TapaCarp interface. This enforcement of the **minimalism** principle paid off, as only two students had trouble with the interface. Globally, TapaCarp was also used in a minimalist way, much of the activity being done without it. This ensured smooth transitions between the various stages of the activity. It also helped introduce TapaCarp to the students and avoided a startling effect for some students afraid of technology.

Overall, the 5 design principles were well respected and the classroom orchestration load was kept low, indicating that the usability of the activity was high.

Connections with the workplace

One of the goals of this activity was to connect the school activity with the workplace. Apprentices did not make connections with the workplace during the execution phase of the activity (at least not explicitly). However, during the reflective phases of the activity, and

especially during the class debriefings, many connections to the workplace were triggered by the teacher and led to meaningful discussions about the workplace.

As mentioned in the results section, the links that the teacher made revolved around four subjects: to include only the necessary information, to respect drawing conventions, to be consistent, and the importance of the plan as a means of communication. Having the objects readily at hand allowed him to give concrete examples of what he was explaining.

Although this is hard to measure, there are signs that this class debriefing participated in an important meaning-making process for the apprentices. One indication that this happened is that apprentices paid attention to the points made by the teacher and did not repeat (or less often) the same mistakes on the second day. For example, the number of plans with inconsistencies decreased between the first and the second day. Apprentices did not question the validity of the activity, whereas the teacher had previously told us that they had repeatedly questioned the usefulness of the drawing activities.

One thing that could be improved in the future is to have apprentices come up with connections to the workplace on their own. For example, during the group peer assessment phase, if the object produced has mistakes they could be asked to very briefly describe a situation at work in which they faced a mistake that was similar to the one they spotted. Alternatively, one could also wrap the activity in a scenario in which one apprentice is the client and the other one the carpenter, with the hope that they would naturally replay situations they experienced at work.

Learning gain

Learning was measured in two ways: with the pre-test and post-test, which showed no significant difference between the experimental group and the control one, and by a task performance measure (assessing the evolution of the quality of the artifacts produced on both days), which did.

The activity was rich and made apprentices work on several competencies. Yet, because of a limited amount of time, the tests could only assess some precise competencies. The richness of the class debriefings indicates that many aspects in relation to understanding a plan were addressed, and it may therefore be that the apprentices developed other skills that were not captured in the post-test. It may also be that some competencies, even if addressed and tested, need more time to develop, and that an effect could only be observed over a longer treatment. A third explanation is that the test questions might not have been discriminative enough to distinguish finer nuances in learning.

These explanations are corroborated by the fact that a more performance-based measure of learning indicated that learners progressed: substantially more artifacts were built correctly on day 2 compared to day 1, despite more difficult tasks on day 2. A longer intervention may have shown further improvements because of more exposure to the learning material but also because the activity may have been more rehearsed by then. This hints to a sixth design principle (in addition to integration, flexibility, empowerment, awareness, and minimalism): rehearsal. In order to be effective, an activity should be rehearsed several times, by both the teacher and the learners. This will need to be confirmed in the future.

The effect of plausibility

There was progress between the first and second day in the quality of the artifacts produced. In particular, fewer plans with inconsistencies were drawn and more correct objects were created on the second day, despite the higher complexity of objects on the second day.

The improvement was more evident on the plausible objects, which indicates that for complicated objects, plausibility plays a key role: the more plausible an object, the easier it is to imagine, even though objective criteria might deem the two objects of equal difficulty. However, the limited number of objects does not allow us to draw strong conclusions, and further testing with a greater variety and number of objects would be needed to confirm the effect of plausibility.

Limitations

Although measuring learning was not the main goal of this study, the lack of a learning effect on the pre-test and post-test measures is disappointing. There are three main challenges in better capturing learning impact of the activity presented in this article. First, learning is extremely rich and hard to measure in short tests, and longer tests are not an option due to the limited classroom time available. Second, a longer treatment would be required to see an effect. Third, test performance depends on the motivation and state of mind of students, which vary substantially and are hard to monitor. A longer use of TapaCarp and a longer and more varied evaluation phase would be needed to assess the exact impact of TapaCarp on learning.

The single deployment also makes it difficult to judge how the activity would work in different contexts. We only worked in one school, with one teacher, and 3 classes.

In the study reported here, researchers helped the teacher run the activity. This is not required by the activity, which could be run by a single teacher, but was necessary because the teacher was not familiar with the activity itself and with the TapaCarp system. Having the activity be run solely by the teacher would require a deeper training on the usage of TapaCarp.

Conclusions

A novel classroom learning activity was presented that aimed to bridge the gap between the classroom and workplace environment in a dual carpentry apprenticeship while keeping the orchestration load low. A tangible user interface, TapaCarp, served as a necessary launchpad for the activity, but was only used for the first part of it.

The activity respected 5 design principles known to reduce orchestration load and its usability was high. The class debriefings, based on the artifacts created by the apprentices, allowed the teacher to touch upon many practical points and make frequent connections to the workplace, thus participating in a meaning-making process for the apprentices. Although the test measures did not show any learning gain, the activity was rather successful: the performance improved from the first day to the second day of the activity, the teacher was enthusiastic, and both the apprentices and the teacher showed a high engagement. Varying the plausibility and difficulty of the objects indicated that once a certain level of difficulty is crossed, a higher plausibility leads to a higher rate of success. This is of interest for the design of future learning material.

This study extends our first deployment of TapaCarp in a classroom (Cuendet and Dillenbourg 2013). The main difference in this second study is that we designed a new pedagogical scenario instead of integrating TapaCarp in existing pedagogical practices. We focused on designing a meaningful activity for carpenter apprentices and sought to use TapaCarp only when it had an added value. As a result, only a small part of the activity was done with TapaCarp, while the rest of the activity used techniques and processes that apprentices already knew. This mix of technical novelty and usage of existing techniques allowed

for a high usability of the activity and both the teachers and the apprentices showed high engagement and gave positive feedback.

The design of this activity has convinced us that the way to integrate computer technology in classrooms should be to mix it with existing practices instead of replacing them. We found the concept of classroom orchestration and thinking about extrinsic constraints an excellent way to achieve this, and would like to encourage other researchers to develop learning activities along these lines.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2-3), 131–152.

Alcoholado, C., Nussbaum, M., Tagle, A., Gomez, F., Denardin, F., Susaeta, H., Villalta, M., & Toyama, K. (2011). One mouse per child. *Journal of Computer Assisted Learning*.

Anderson, D., Frankel, J.L., Marks, J., Agarwala, A., Beardsley, P., Hodgins, J., Leigh, D., Ryall, K., Sullivan, E., & Yedidia, J.S. (2000). Tangible interaction + graphical interpretation: a new approach to 3d modeling. In *Proc. of SIGGRAPH*.

Blikstein, P. (2008). Travels in Troy with Freire, chapter 14, Sense Publishers.

Bonnard, Q., Zufferey, G., Mazzei, A., Cuendet, S., Li, N., & Dillenbourg, P. (2013). Chilitags, robust fiducial markers for augmented reality.

Brousseau, G. (1997). *Theory of Didactical Situations in Mathematics: Didactique des mathmatiques, 1970-1990*: Springer.

Butterworth, G., & Light, P. (1982). *Univ of Chicago Pr (Tx)*: first edition edition.

Cuendet, S. (2013). Tangible Interfaces for Learning. PhD thesis, IC, Lausanne.

Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2012a). *Designing Augmented Reality for the Classroom*: Computers and Education. Special Issue on Augmented Reality for Learning.

Cuendet, S., Bumbacher, E.W., & Dillenbourg, P. (2012b). Tangible vs. virtual Representations: when Tangibles Benefit the Training of Spatial Skills. In *Proc NordiCHI* (p. 2012).

Cuendet, S., & Dillenbourg, P. (2013). The Benefits and Limitations of Distributing a Tangible Interface in a Classroom. In *CSCL 2013*. Best paper award.

Cuendet, S., Jermann, P., & Dillenbourg, P. (2012c). Tangible interfaces: when physical-virtual coupling may be detrimental to learning. In *Proceedings of the 2012 British Computer Society Conference on Human-Computer Interaction*.

Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal Of Computer-Supported Collaborative Learning*, *491514*, 6.

Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning*, *3*(1), 523.

Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In *In New Science of Learning*. Springer Science+Business Media.

Dillenbourg, P., Jrvel, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In Balacheff, N., Ludvigsen, S., Jong, T., Lazonder, A., & Barnes, S. (Eds.) *Technology-Enhanced Learning*, (pp. 3–19). Dordrecht: Springer Netherlands.

Dillenbourg, P., Nussbaum, M., Dimitriadis, Y., & Roschelle, J. (2012). Design for classroom orchestration. *Computers & Education*.

Dimitriadis, Y., Asensio-prez, J.I., Hernndez-leo, D., Roschelle, J., Brecht, J., Chaudhury, S.R., Digiano, C., & Patton, C.M. (2007). From socially-mediated to technology-mediated coordination: A study of design tensions using group scribbles.

Do-Lenh, H.S. (2012). Supporting Reflection and Classroom Orchestration with Tangible Tabletops. PhD thesis, EPFL, Lausanne.

Do-Lenh, S., Jermann, P., Cuendet, S., Zufferey, G., & Dillenbourg, P. (2010). Task performance vs. learning outcomes: a study of a tangible user interface in the classroom. In *Proceedings of the 5th European conference on Technology enhanced learning conference on Sustaining TEL: from innovation to learning and practice*. EC-TEL'10, p. 7892, Berlin, Heidelberg. Springer-Verlag.

Fitzmaurice, G.W., & Buxton, W. (1997). An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*. p. 4350, New York, NY, USA. ACM Press.

Gaillard, L. (2012). Perspectives de la formation - scnarios 20112020 pour le degr secondaire ii. Technical report, Office fdral de la formation professionnelle et de la technologie OFFT.

Girouard, A., Solovey, E.T., Hirshfield, L.M., Ecott, S., Shaer, O., & Jacob, R.J.K. (2007). Smart blocks: a tangible mathematical manipulative. In *Proc. of TEI*.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*: Harvard University Press.

Ha, V., Inkpen, K., Mandryk, R., & Whalen, T. (2006). Direct intentions: the effects of input devices on collaboration around a tabletop display. In *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, 2006. TableTop 2006,* page 8 pp.

Horn, M.S., Solovey, E.T., Crouser, R.J., & Jacob, R.J. (2009). Comparing the use of tangible and graphical programming languages for informal science education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09, p. 975984, New York, NY, USA. ACM.

Hornecker, E., & Buur, J. (2006). Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06, p. 437446, New York, NY, USA. ACM.

Hornecker, E., Marshall, P., Dalton, N.S., & Rogers, Y. (2008). Collaboration and interference: awareness with mice or touch input. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. CSCW '08, pp. 167–176, New York, NY, USA. ACM.

Kim, M.J., & Maher, M.L. (2008). The impact of tangible user interfaces on designers' spatial cognition. *HumanComputer Interaction*, *23*(2), 101–137.

Manches, A., & O'Malley, C. (2012). others Tangibles for learning: a representational analysis of physical manipulation. *Personal Ubiquitous Comput*, *16*(4), 405419.

Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proceedings of the 1st international conference on Tangible and embedded interaction*. pp. 163–170. ACM.

Marshall, P., Fleck, R., Harris, A., Rick, J., Hornecker, E., Rogers, Y., Yuill, N., & Dalton, N.S. (2009). Fighting for control: children's embodied interactions when using physical and digital representations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09, p. 21492152, New York, NY, USA. ACM.

Moraveji, N., Ringel Morris, M., Morris, D., Mary, C., & Riche, N. (2011). Classsearch: Facilitating the development of web search skills through social learning. In *Proc. of CHI*.

O'Malley, C., & Stanton Fraser, D. (2004). *Literature review in learning with tangible technologies*. Technical report, FutureLab.

Parkes, A.J., Raffle, H.S., & Ishii, H. (2008). Topobo in the wild: longitudinal evaluations of educators appropriating a tangible interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08, p. 11291138, New York, NY, USA. ACM.

Pecher, D., & Zwaan, R.A. (Eds.) (2005). *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*: Cambridge University Press.

Price, S., Falcão, T.P., Sheridan, J.G., & Roussos, G. (2009). The effect of representation location on interaction in a tangible learning environment. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*. TEI '09, pp. 85–92, New York, NY, USA. ACM.

Quarles, J., Lampotang, S., Fischler, I., Fishwick, P., & Lok, B. (2008). Tangible user interfaces compensate for low spatial cognition. In *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on* (pp. 11–18).

Rogers, Y., & Lindley, S. (2004). Collaborating around vertical and horizontal large interactive displays: which way is best?. *Interacting with Computers*, *16*(6), 1133–1152.

Roschelle, J., Rafanan, K., Estrella, G., Nussbaum, M., & Claro, S. (2009). From handheld collaborative tool to effective classroom module: embedding cscl in a broader design framework. In *Proc. of CSCL*.

Roschelle, J.M., Pea, R.D., Hoadley, C.M., Gordin, D.N., & Means, B. (2000). Changing how and what children learn in school with computer-based technologies. *Future of Children*.

Schwartz, D.L., Chase, C.C., Oppezzo, M.A., & Chin, D.B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759–775.

Shaer, O., & Hornecker, E. (2009). Tangible user interfaces: Past, present and future directions (preprint). *Found. Trends Hum.-Comput. Interact*, *3*(1-2), 23–27.

Sharlin, E., Watson, B., Kitamura, Y., Kishino, F., & Itoh, Y. (2004). On tangible user interfaces, humans and spatiality. *Personal Ubiquitous Comput*, *8*(5), 338346.

Song, H., Guimbretire, F., Hu, C., & Lipson, H. (2006). ModelCraft: capturing freehand annotations and edits on physical 3D models. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*. UIST '06, p. 1322, New York, NY, USA. ACM.

Speelpenning, T., Antle, A.N., Doering, T., & van den Hoven, E. (2011). Exploring how tangible tools enable collaboration in a multi-touch tabletop game. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part II*. INTERACT'11, p. 605621, Berlin, Heidelberg. Springer-Verlag.

Suthers, D. (2006). Technology affordances for intersubjective meaning-making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, *315337*, 1.

Tuddenham, P., Kirk, D., & Izadi, S. (2010). Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10, p. 22232232, New York, NY, USA. ACM.

Vygotsky, L.S., & Cole, M. (1978). *Mind in society: the development of higher psychological processes*. Cambridge: Harvard University Press.

Zhang, J., & Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122.

Zufferey, G., Jermann, P., Lucchi, A., & Dillenbourg, P. (2009). Tinker Sheets: Using Paper Forms to Control and Visualize Tangible Simulations. In *Proc. of TEI*.