Tangible Interfaces for Learning: Training Spatial Skills in Vocational Classrooms

THÈSE N^{O} 5931 (2013)

PRÉSENTÉE LE 8 NOVEMBRE 2013

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE D'ERGONOMIE ÉDUCATIVE
PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Sébastien CUENDET

acceptée sur proposition du jury:

Prof. P. Fua, président du jury
Prof. P. Dillenbourg, directeur de thèse
Prof. P. Blikstein, rapporteur
Dr R. Boulic, rapporteur
Prof. E. Hornecker, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2013

Acknowledgements

Contributions to this thesis

I would like to start by thanking my advisor, **Pierre Dillenbourg**, for his precious help and support throughout my PhD. Pierre was always available when required and was always supportive, encouraging, and understanding. Moreover, his motivation and personal involvement in the visit of companies (Chapter 3) and the first classroom study (Chapter 7) were integral in bootstrapping the project.

Even if he is not officially my co-advisor, **Patrick Jermann** played this role for most of my PhD. Our discussions on the strengths and weaknesses of the Swiss dual apprenticeship system helped me gain a better understanding of this system. I also appreciated his patience in discussing statistics, was inspired by his passion for data analysis, and learned a lot from designing with him the feedback experiment reported in Chapter 6.

The quality of this thesis would have been much lower without the careful proof-reading and many suggestions of **Jessica Dehler**. Jessica significantly contributed to the second classroom study presented in Chapter 7. The many passionate discussions that ensued from this study helped refine the final message and contributions of this thesis, and for that I am deeply thankful to her.

The study on spatial skills reported in Chapter 4 was done with the help **Christoph Arn** and **Engin Bumbacher**. I am grateful to them, to the schools that welcomed us (the professional schools in Bulle, Thun, and Zollikofen, and the gymnase de Burier), and to their students who agreed to freely participate in it, for making this study possible.

During the design implementation of TapaCarp, I received help from many people. I benefited from the pioneering work done by **Guillaume Zufferey** in his own PhD thesis on the Tinkerlamp (Zufferey, 2010). I also found encouragement from collaborating closely on the development of the Tinkerlamp with Guillaume, but also **Quentin Bonnard**, **Andrea Mazzei**, **Son Do-Lenh**, and several EPFL Master's students. **Stéphane Testuz**, **Yongsung Kim**, and **Sara Alsudairy** helped implement the activities and build the physical interface used in the two studies presented in Chapter 6 and the first one of Chapter 7. For this last experiment, Stéphane also helped run the study itself. The main effort on e-TapaCarp (Chapter 8) was made by **Claude Bossy** during his Masters Thesis (Bossy, 2013).

This thesis is part of a project financed by the State Secretariat for Education, Research and

Innovation (**SERI**). The exchanges with the other researchers involved in the project were rich and contributed to this thesis. Special thanks to **Giulia Ortoleva** for her help in running the second study of Chapter 7. I would also like to thank the members of the advisory board of this project for their interest, constructive criticisms, suggestions, and many encouragements.

One of the fundamental aspects of this work is that it is anchored in the practice. This would not have been possible without the agreement of the management of the **Centre d'Enseignement Professionel de Morges** (CEPM) and that of l'**École de la Construcion à Tolochenaz**, who were our partners all along this project. I am also thankful to the companies that we visited (visits reported in Chapter 3): Volery Frères SA, in Aumont; Diserens J.F. SA, in Dompierre; Menuiserie-Charpente Rottet Sàrl, in Essertines-sur-Yverdon; André SA, in Yens; and Hubert Constructions Bois SA, in Bex.

Discovering the world of vocational education and training in carpentry was a fascinating experience, and the credit goes to my guides in this discovery: the carpentry teachers. First of all, thanks to **Philippe Ogay** for the numerous hours spent explaining the intricacies of descriptive geometry applied to carpentry, and the habits and custom of the carpentry world, and for his personal investment witnessed, among other things, by his volunteering for cutting the wooden blocks used in the TapaCarp interface. Thanks to him and his colleagues, **Pascal Wulliamoz** and **Laurent Rüfenacht** for putting up with my disturbing presence in their classrooms when running the studies of Chapter 6 and the first one of Chapter 7. **Sandro Melchior** let me enter his workshop and run the second study of Chapter 7. I am grateful for his infectious enthusiasm and for taking the time to share his passion for teaching.

The **CHILI lab** is fortunate enough to have many great **visitors** and I have benefited from their questions, criticism, and suggestions. A special mention goes to **Dan Schwartz**, who made many valuable suggestions on the interpretation of the results, especially for the first study reported in Chapter 6.

Last but not least, I would like to thank all the members of my jury for providing insightful suggestions on how to improve my thesis. They not only assessed my work as their duties required, but also made many great suggestions that improved the overall quality of this thesis.

Parts of this thesis have been published in peer-review conferences (Cuendet et al., 2012c,b; Cuendet and Dillenbourg, 2013) and a journal ((Cuendet et al., 2012a)), and have benefited from the insightful comments and suggestions of anonymous reviewers. I would like to thank them, as well as my co-authors for helping me refine my research.

Note on "I" versus "we". As made obvious by the previous paragraphs, the work presented in this thesis resulted from a collaboration with several other researchers. As a consequence, I wrote most of my thesis in the first person plural form ("we"), and switched to the first person singular form ("I") when referring to actions or statements for which I am the only one accountable.

The rest of the story

Throughout my life as a researcher, I have been fortunate to meet exceptional people who took time to mentor me. I want to particularly thank **Dilek Hakkani-Tür** for helping me during my first steps in the academic world, and **Elizabeth Shriberg** for showing me by example what academic excellence is.

During my PhD, in the summer 2012, I had the privilege to intern for 3 months at Microsoft Research India. This was an extraordinary experience, from which I learned a lot in many ways, and I am thankful to **Ed Cutrell** for offering me this opportunity, and to **Indrani Mehdi** and **Andrew Cross** for their collaboration and friendship.

The path that leads to a thesis is emotionally bumpy. This is even more true when pursuing interdisciplinary research, as we do in CHILI lab. It is therefore primordial to have a team of close colleagues to be able to share the ups and downs of the PhD life. I am grateful that I could find such a **team in CHILI lab**, and share both successes and failures with **Andrea**, journeyman from the start to end, **Florence**, always there to help, **Guillaume**, **Hamed**, **Himanshu**, **Jessica**, **Julia**, **Khaled**, **Kshitij**, **Patrick**, **Quentin**, and all the others! I will treasure the memory of late nights on days preceding the experiments and paper deadlines; the beers at Sat to celebrate successes, or forget temporary failures; the many birthday bombings; and the outings, every year, in the snow.

The time of this PhD also coincided with the formidable adventure of Wizbee's creation. I really enjoyed having this parallel adventure, and I would like to thank my **colleagues and friends at Wizzy**, Jean-Philippe, Thomas, Javier, Kepa, and Clément for making it possible.

I am also indebted to all my "non research" friends for helping me step back and reconnect with the real world from time to time by sharing a beer or a meal, doing sports together, or just hanging out. Thank you friends!

I am grateful to my family for their strong support during these 4 years: **Gwendeline**, **Ludovic**, and **Matis** and **Juliette** for their numerous visits who have often lit up my days; **Matthieu** and **Coralie** for the lunches shared on campus; and of course, both my incredibly loving and supportive **parents** for their incessant interest and effort to understand what exactly I was doing.

Finally, I wish to thank **Lorin**, my wife, for her inexhaustible patience, and her unfailing love and support throughout this adventure. I dedicate this thesis to her.

Lausanne, October 2013 S.C.

Résumé

De nombreuses affirmations ont été faites au sujet des avantages des interfaces tangibles pour l'apprentissage. Ainsi, les interfaces tangibles amélioreraient la collaboration entre apprenants, offriraient une plus grande facilité d'utilisation, et permettraient aux apprenants d'avoir une meilleure perception de la tâche d'apprentissage, surtout dans les tâches avec une composante spatiale. Il existe cependant peu de données empiriques qui soutiennent ces affirmations, et malgré tous ces avantages supposés, les interfaces tangibles sont peu utilisées dans les écoles.

Cette thèse explore ces deux problématiques dans le cadre de la formation professionnelle des apprentis charpentiers. Le domaine d'apprentissage concerne les capacités spatiales, et en particulier le passage entre les représentations 2D et 3D d'un objet. Nous étudions (1) si les interfaces tangibles aident au développement des capacités spatiales, et (2) quels types de scénarios pédagogiques elles peuvent soutenir dans la classe. Nous suivons une approche empirique au travers d'études menées dans les classes.

Cette thèse contribue aux trois domaines de recherche suivants:

1. **Développement de capacités spatiales.** Nos résultats montrent que les interfaces tangibles peuvent aider au développement des capacités spatiales des apprentis charpentiers. La nature tangible de l'interface aide l'apprenant à mettre en relation les représentations multiples d'un objet, surtout pour des problèmes difficiles. Ceci est en particulier bénéfique pour les débutants.
2. **Apprentissage avec les interfaces tangibles.** Selon nos résultats, le seul fait d'utiliser des interfaces tangibles ne garantit pas un effet d'apprentissage, et les choix de conception sont primordiaux. En effet, de petites différences, telles que la correspondence entre la forme de l'objet tangible et sa représentation digitale, ou le type de retour donné à l'utilisateur par le système, ont un impact direct sur l'apprentissage.
3. **Technologie dans la classe et orchestration.** Nous explorons plusieurs types de scénarios pédagogiques dans la classe, que les interfaces tangibles soutiennent. Le scénario le plus prometteur est celui dans lequel l'interface tangible s'insère dans une activité comprenant également une partie sans technologie. En sus, nous proposons 5 principes de conception pour améliorer l'intégration des interfaces tangibles dans la classe, et montrons comment cette intégration peut être faite à moindre coût.

Mots-clé: interface tangible, orchestration de classe, développement de capacités spatiales, apprentissage collaboratif supporté par ordinateur, formation professionnelle, réalité augmentée.

Abstract

There have been many claims that Tangible User Interfaces (TUIs) can have a positive impact on learning. Alleged benefits include increasing usability, improving engagement and collaboration of students, and providing a better perception of the task, especially spatial ones. However, there exists little empirical data to back up these claims. Moreover, for all their potential benefits for learning, TUIs are still scarcely used in schools.

This thesis explores these two issues in the specific context of vocational education and training of carpenter apprentices. The learning objectives concern spatial skills and in particular, the mapping between 2D and 3D representations. We study (1) whether TUIs can support the training of spatial skills, and if so what features allow them to do so, and (2) what kinds of classroom pedagogical scenarios TUIs can support. We follow a design-based research approach and run empirical studies, mostly in classrooms.

The contributions of this thesis touch on three research domains:

1. **Spatial skills.** Our results show that TUIs can help teach spatial skills to carpenter apprentices. The tangible nature of TUIs can help the learner relate multiple representations of an object, especially for difficult problems. It can also lower the barrier to entry into a learning domain for beginners.
2. **Learning with TUIs.** According to our results, TUIs can benefit learning, but the mere fact of using TUIs does not guarantee learning. Instead, special attention needs to be given to the design of the TUI. Small design variations, such as the physical correspondence between the tangible object and its virtual representations, or the type and timing of feedback given to the user, can have a significant impact on learning.
3. **Classroom technologies and orchestration.** We explore several classroom pedagogical scenarios that TUIs can support. The most promising one is to use a TUI as part of a hybrid classroom learning activity that includes both TUI and non TUI steps. Additionally, we devise two ways to promote the integration of TUIs in classroom. First, we introduce 5 design principles that reduce the classroom orchestration load. Second, we show how new web technologies can be used to deploy TUIs in schools at a lower cost.

Keywords: tangible user interfaces, classroom orchestration, spatial skills training, computer-supported collaborative learning, vocational training, augmented reality.

Contents

1 Introduction

1.1 Motivation

It is 7.45 a.m. on a Thursday morning. Joël, 16, is about to enter the classroom where he goes every Thursday to learn about the fundamentals of carpentry, as well as some general knowledge. Joël is a carpenter apprentice, and this is the only day he spends at school during the week. The four remaining days of the week, he works in a carpentry company, where he works as a carpenter under the supervision of trained carpenters.

As he is about to enter the classroom, Joël turns off his phone, as it is not allowed to use it in the classroom. Actually, any technology does not seem welcome in this classroom. Sure, there are two or three dusty computers in a corner of the classroom, and an overhead projector that the teacher uses episodically. But he never uses technology in this classroom, and that is very different from the rest of his life, where he uses technology to keep in touch with his friends, plan his vacation, take notes and pictures at work, and follow the news about his favorite ice-hockey team. Even at work, technology is used heavily, be it to create the plans of wood structures with a computer-aided design (CAD) software, or sometimes, to cut wood by using a computer numerical control (CNC) wood router.

These observations are not an idiosyncrasy of Joël's school. Although computers are present in many schools, they are still scarcely used for learning and teaching. Developing technology to improve learning has been a quest for educators since the early 1980s and the advent of personal computers. But despite much effort, technology did not change education in the way many predicted it would. The general reluctance to change from teacher unions, parents' fear that their children might get a poorer education, political motivation (and absence thereof), slowness in adapting curricula and learning material, the lack of a proven added value of computer technology for learning, or lack of money, are all reasons cited to explain why this is the case.

However, if technology can help Joël in so many domains of his life, why could it not help him learn his job, too? Investors seem to think that it could. According to GSV Advisors (Quazzo et al., 2012), a consultancy, total investments in educational technology in the United States reached $1.1. billion in 2012, twice what they were in 2010. Some education companies are gaining popularity, such as the Khan Academy, a non-profit educational website that features over 4000 videos on various school subjects, which has about 6 millions visitors a month, and is now being deployed in some schools in California.

This thesis investigates whether an alternative technology, tangible user interfaces (TUIs), could meet a greater success in classrooms than traditional, mouse and keyboard, computers. TUIs were introduced by Ishii and Ullmer (1997) and allow users to interact with digital information through physical objects. There are several reasons why TUIs can foster learning. For example, their physicality can increase usability, lower the barrier to entry for novice users, and improve collaboration between students. Compared to traditional computers, which provide information mainly through visual stimuli, TUIs can provide additional stimuli through touch, or spatial information. This can be useful in some particular domains, especially spatial ones (e.g. it is easier to see relief on a 3D map than interpret it from contour lines). In the classroom, tabletop TUIs may ease the transition between the digital and physical worlds through the use of physical objects on a tabletop. This type of usage is closer to the practice of the classroom than manipulating a mouse and keyboard, and may help teachers in orchestrating their classroom. Finally, in the specific case of carpenter training, TUIs allow technology to be close to the "hands-on" culture of carpentry.

For all the potential benefits of TUIs for learning, TUIs are still used scarcely in formal learning environments. One reason for this is that there exists little empirical data to confirm these benefits. Many creative TUIs for learning have been designed, addressing various topics, and researchers have often been able to show that they lead to higher engagement. But comparative studies showing that TUIs can actually increase the learning performance are rare. Without proper evidence of better learning, it is difficult for policy makers to encourage this type of learning technologies. But even with solid evidence of better learning with TUIs, it could still be hard to introduce TUIs in classrooms. The classroom is a complex ecosystem, with many constraints and a fragile balance that the teacher strives to maintain. A new technology would surely disrupt the classroom. One solution to get teachers on board is to design the technology with the classroom in mind and involve them in the design process. However, it is tempting for designers of learning technologies to create these technologies outside of the classroom and deliver them to the teacher as a finished good.

The work presented in this thesis studies TUIs and learning in a specific context: the training of spatial skills in the initial vocational training of carpenters in Switzerland. Vocational training is a key component of the Swiss education system, and the most popular one: more than two-thirds of students do an apprenticeship (about one-third go to high school). The apprenticeship is a mix of school teaching and work practice. At school, one of the main skills learned by carpenters is linked to interpreting 2D plans into 3D building, and vice versa, which is linked to spatial skills. Spatial skills are hard to acquire. They are also gaining interest among educators, as some studies have shown that having good spatial skills can help in science, technology, engineering, and mathematics, also known as STEM fields (Wai et al., 2009; Newcombe, 2010; Sorby et al., 2013). Moreover, a recent meta-analysis by Uttal et al. (2012) showed that they are trainable, something that had long been debated.

This thesis touches on three research domains: the training of spatial skills, learning with TUIs, and classroom technologies. I will first try to show how TUIs can help develop spatial skills, which would contribute to both the training of spatial skills and learning with TUIs. I will then present ways to make it easier to integrate a TUI in a classroom. This is formalized by the following research objectives.

1.2 Research Objectives

The lack of technology mentioned for education in general also exists in vocational education and training. Aware that technology could help apprentices, the Swiss State Secretariat for Education, Research and Innovation (SERI) is funding research to study how technology can help in the initial vocational training. The work presented in this thesis is part of this effort, and the research objectives, presented below, are all grounded in the context of initial vocational training of carpenters in Switzerland.

1. **Investigate whether TUIs can help train spatial skills for carpenters.** We make the hypothesis that some features of tangibles can help train spatial skills, but that minor design changes can greatly affect learning. In particular, we study the influence on the learning experience of (1) varying the shape of the tangible object, and (2) varying the type of feedback that is given to the learner by the system.
2. **Investigate how TUIs can be used in a classroom to train spatial skills.** We make the hypothesis that one reason why TUIs are not used much in classrooms is that they are not adapted to classroom scenarios. Taking the classroom constraints into account in our design, we look at what pedagogical classroom scenarios TUIs can best support.
3. **Show how low-cost TUIs can be built.** TUIs often require dedicated, expensive, hardware, which may also explain their lack of deployment in schools. New web technologies make it possible to build cheaper, potentially highly scalable, TUIs. We show through an example how this can be done, and indicate what some of the pitfalls of such an approach might be.

1.3 Organization of This Thesis

The remainder of this thesis is organized as follows. The next chapter gives an overview on previous research on TUIs for learning, and on the most recent theories on integrating technologies in the classroom. Chapter 3 describes in more details the context in which this work was done, especially the vocational training of carpenters. Chapter 4 first briefly describes previous research on spatial skills, and then reports on a study that we conducted to gain a deeper knowledge on carpenter apprentices and their spatial skills.

The research methodology is described in Chapter 5, together with the TUI that we created and the design process that led to it.

The empirical studies that brought the main results of this thesis are presented in Chapters 6, 7, and 8. Chapter 6 reports on two semi-controlled studies that aimed at identifying what features of TUIs can help train spatial skills (research objective (RO) 1). The studies reported in Chapter 7 focus on the integration of our TUI in the classroom (RO 2). In Chapter 8, I describe how we designed a low-cost TUI and report on the results of a first study (RO 3).

Chapter 9 summarizes the main findings of this thesis, mentions its limitations, and highlights potential future research directions.

2 Related Work

This thesis is the result of an interdisciplinary work, as illustrated by Figure 2.1. It describes the design and implementation of a tangible user interface (TUI) that aims at helping students develop a specific skill. It is meant to be used in the classroom, either in an individual or collaborative fashion.

In this chapter, I review the related work on TUIs and on Computer-Supported Collaborative Learning (CSCL). The next chapter will detail the learning context and skill targeted by the learning technology.

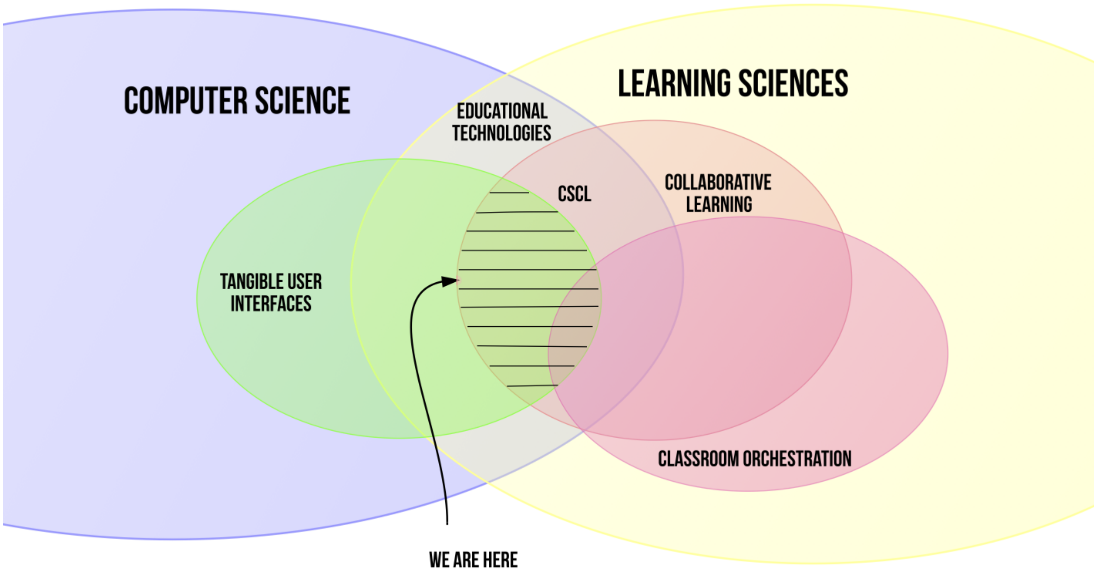


Figure 2.1: A graphical overview to situate this work.

2.1 Tangible User Interfaces

Tangible User Interfaces (TUIs) were first introduced as Graspable Interfaces (Fitzmaurice, 1996; Fitzmaurice and Buxton, 1997), which aimed at increasing the manipulability of digital objects through the use of wooden blocks. The main advantage of a Graspable Interface over a traditional GUI is that it provides concurrent access to several physical input devices, as opposed to one (the mouse). Ishii and Ullmer (1997) presented the concept of Tangible Bits, which then led to the proposition of Tangible User Interfaces. Tangible Bits built upon the concept of Graspable Interfaces and added two more components: interactive surfaces and ambient media for background awareness.

According to Shaer and Hornecker (2009), 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 (see Figure 2.2). This was novel and contrary to the main trend that focused on forcing the user into a virtual world.

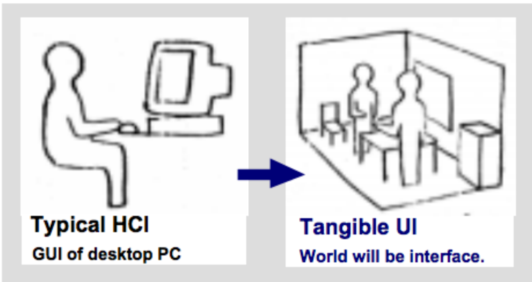


Figure 2.2: The original motivation for developing TUIs: give the user a physical grasp on the digital world instead of forcing the user into a virtual world (Ishii and Ullmer, 1997).

While Ishii and Ullmer (1997) coined the term "Tangible User Interfaces", Wellner (1991) had actually already implemented the ideas of augmented reality workspace. The DigitalDESK that Wellner created was a traditional desk augmented with digital capabilities thanks to a projector and a camera located above the tabletop. The difference with the work on Tangible Bits was the emphasis that the latter put on graspable objects and on ambient media.

The seminal work by Ishii and Ullmer (1997) was followed by a great variety of research spanning different domains through the development of many proof-of-concept systems. Conceptual frameworks and taxonomies also emerged that tried to scope the various shapes that a TUI could take and to classify them. A widely used classification is the one by Ullmer et al. (2005), who distinguished three dominant types of TUIs: interactive surfaces, token+constraint, and constructive assembly (Figure 2.3). Interactive surfaces designate systems in which objects are placed and manipulated on a planar surface. Constructive assembly are made of elements that can be connected together. The order and the shape in which they are connected can be interpreted by the system. Finally, token+constraint systems provide a mechanical structure for the manipulation of tokens. The physical constraints imposed on the manipulation of the tokens can guide the users in their interaction with the system.

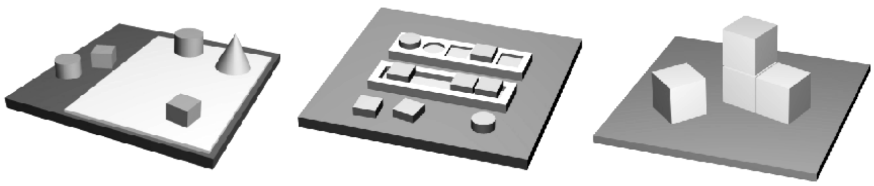


Figure 2.3: The three dominant types of TUIs identified by Ullmer et al. (2005): interactive surfaces, token+constraints, and constructive assembly.

While these classifications can serve as a way to structure the different types of TUIs, the classification of a TUI into one of those three categories is not always evident to apply, as noted by Shaer and Hornecker (2009). In fact, researchers have dedicated much effort into identifying classification and terminology of TUIs. This was partly needed because TUIs emerged from several research fields, such as augmented reality, tangible tabletop interaction, ambient displays, and embodied user interfaces, and partly because they were applied to many different domains. Shaer and Hornecker (2009, pp. 17-21) provide a summary of the recent research on unifying perspectives between the different research fields and domain applications.

The goal of this review is not to present an exhaustive panorama of TUIs, since it was provided by (Shaer and Hornecker, 2009). Rather, I will foremost focus on the the link between TUIs and learning.

2.2 Tangibles and Learning

2.2.1 Physicality and Learning

Education has been one of the main domains of application for TUIs, resulting in many TUI prototypes for various learning domains. The idea of using physical objects for learning is not new. The Swiss educator Johann Heinrich Pestalozzi (1746-1827) was among the first advocates for hands-on learning (Pestalozzi, 1803). The world's first kindergarten, created by Friedrich Froebel, was filled with objects carefully designed to help children recognize patterns and forms. Montessori (1912) built on Froebel's work and created new activities and material to develop children's sensory capabilities .

Jean Piaget later provided a theory for these educational ideas, a theory referred to as constructivism (Piaget, 1972). Piaget's idea was that children should first construct knowledge through concrete actions before formalizing their knowledge. Seymour Papert collaborated with Piaget during many years and shared many of Piaget's ideas. However, Papert was not only an educator, but also a mathematician and computer scientist and was convinced that children could benefit from the use of computers for learning and developing creativity (Papert, 1980). Together with Sherry Turkle, he argued that concrete manipulations should not be undermined, as they are too often, when compared with abstract reasoning (Turkle and Papert, 1990).

Papert's theory, constructionism, builds upon Piaget's work and claims that the construction of knowledge is particularly efficient when students build and share objects (Harel and Papert, 1991). Along with his theory, Papert also created Logo, a programming language that children can use to explore and solve problems. The Logo programming language was used in several behavior construction kits, the most famous of them now being the well known LEGO® Mindstorms (Resnick, 1993). Research showed that using toys derived from Papert's work (Resnick et al., 1998) children were able to explore concepts that were previously "too advanced" for them. Those toys also allow the design of new learning activities in which children can conduct their own science experiments, create autonomous creatures and active environments.

Although it has been widely supported and adopted, and despite the fact that others have continued to follow this approach successfully (e.g. Buechley et al., 2008; Blikstein, 2008), the use of manipulatives and physical objects for early education has also been questioned by recent research in the learning sciences (McNeil and Jarvin, 2007; Klahr et al., 2007; Triona and Klahr, 2003; Zaman et al., 2012). McNeil and Jarvin (2007) pointed out that students might have fun at the expense of deep learning and that learning might be too difficult because of the dual representation offered by manipulatives. Triona and Klahr (2003) compared physical and virtual interfaces for learning activities and did not find any difference between the two.

Digital manipulatives do not, in the sense of Shaer and Hornecker's definition, fall into the category of tangible user interfaces. They do provide tangible manipulation of objects that are digitally augmented, but do not offer a way to manipulate digital information in the way TUIs do. However, I described their history in detail for two reasons.

The first one is that the confrontation observed in the case of digital manipulatives is characteristic of one that is also often observed for TUIs. It is a confrontation between an enthusiastic view based on common sense but little empirical evidence on the one hand, and a more skeptical one based on empirical evidence or the lack thereof on the other hand.

The second reason is that the tangible manipulation is a key element of TUIs and one that is often cited when developing TUIs for learning. There are nevertheless additional potential advantages of TUIs for learning, which I will describe in Section 2.2.3. But first, let us look at some examples of TUIs relevant to my thesis.

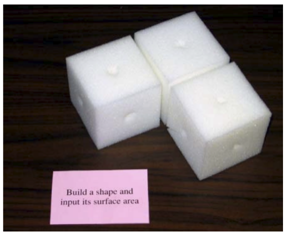
2.2.2 Relevant applications of TUIs

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. I focus here on TUIs that are closer to the one presented in this Thesis, 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.

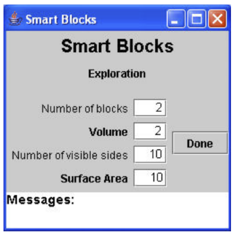
TUIs to build shapes

Smart Blocks are augmented blocks that allow users to build 3D models and to explore concepts such as the surface and the volume of a 3D model (Girouard et al., 2007). Smart Blocks use the RFID technology and smart connectors to compute the surface and volume of the 3D model. The interface of the system has both a TUI composed of cubes, connectors, and question cards, and a GUI through which users can input their answers (Figure 2.4).

(a) The blocks.



(b) The connectors.



(c) The GUI.

Figure 2.4: Smart Blocks (Girouard et al., 2007).

A more complex system is presented in Anderson et al. (2000), in which physical building blocks self-describe and interpret the structure in which they are built, making it possible to infer the geometry of the 3D model. The tangible model is combined with graphical interpretation techniques that enhance the virtual model, as can be seen in Figure 2.5.

Song et al. (2006) developed an interface to annotate physical 3D models and integrate the annotations back into a CAD software. The system, illustrated in Figure 2.6, allows users to

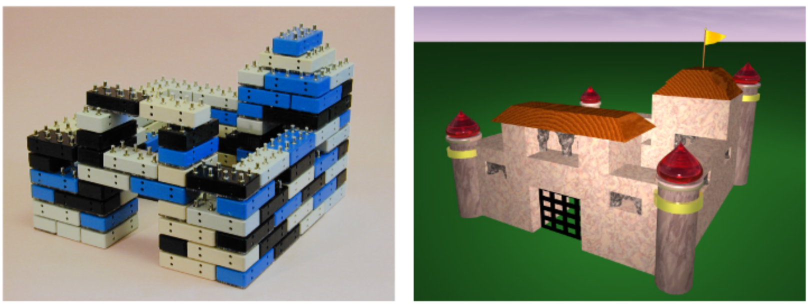


Figure 2.5: On the left, the physical model built with blocks that are able to self-describe and interpret the structure in which they are built. On the right, the corresponding virtual model automatically enhanced (Anderson et al., 2000).

comment and edit their model directly on the physical object. The changes made to the model are reported automatically into the digital model without the user having to manually report them.

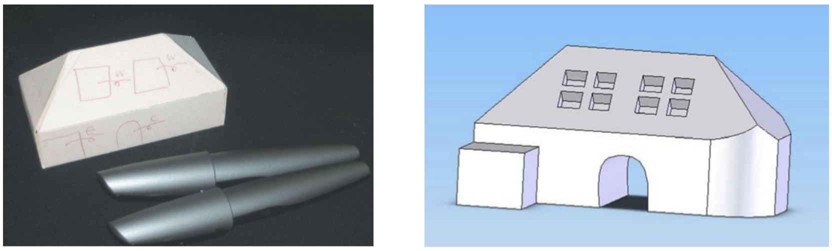


Figure 2.6: ModelCraft: annotations made with a pen on the physical model (left) can be integrated back into the CAD software (right) (Song et al., 2006).

Illuminating Clay was developed to allow the exploration of free form spatial models (Underkoffler and Ishii, 1998). A landscape is modeled using clay and its 3D geometry is captured using a laser scanner. The system then projects information such as shadow casting and land erosion back onto the clay model.

Topobo (Parkes et al., 2008) is a set of tangibles similar to Digital Manipulatives. It is a construction kit with kinetic memory, meaning that it can record and play back physical motion. The system is made of active and passive pieces. The users connect these pieces together to create an object (as shown in Figure 2.7), apply a motion to their object to "teach" it how it should move. The object can then play back this motion.

Tangibles for spatial skills

As will be explained in the next chapter, this thesis focuses on the training of spatial skills. There have been some TUIs that focused on spatial skills.

Kim and Maher (2008) studied the impact of TUIs on designers' spatial cognition. Comparing the usage of a GUI and a TUI, they showed that designers using the TUI recognized more spatial relationships, were more immersed in the task, and discovered new visuo-spatial

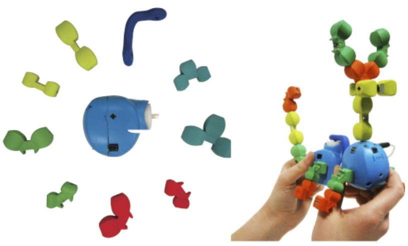


Figure 2.7: The Topobo system (Parkes et al., 2008)

features when revisiting their design. They concluded that TUIs demonstrated potential to support creative processes.

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.

Figure 2.8: ActiveCubes (Watanabe et al., 2004).

ActiveCubes (Watanabe et al., 2004) are small plastic cubes that can be connected together and that each have a small CPU. Each cube is unique and the topology of the connected shape can thus be determined by a connected computer (see Figure 2.8). ActiveCubes were used to assess spatial and constructional abilities (Sharlin et al., 2002). Users put the blocks together to match a 3D shape that is displayed on a screen. Empirical studies showed that the results of the assessment done with ActiveCubes were comparable to those done with a traditional paper-and-pencil 3D assessment.

TUIs on interactive surfaces

Of the three categories of TUIs identified by Shaer and Hornecker (interactive surfaces, constructive assembly, and token+constraint), the systems presented so far fall mostly into the constructive assembly one. This is because most of the work done on geometrical and spatial related concepts has so far been using this type of interfaces. However, the system presented in this thesis falls into the category of interactive surfaces, more specifically into the category of interactive *tabletop* surfaces, and it is thus useful to review the main findings made with these kinds of systems.

An interactive tabletop surface is a horizontal surface that serves both as an input and output space. In general, an interactive tabletop surface is 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. Projecting from the top allows to project on any surface and to augment objects placed on the surface, but it comes at the price of occlusions and the inability to detect accurately whether an object is touching the table or not. The latter can for instance make finger tracking unreliable and therefore limit the naturalness of interaction. Finger tracking is more easily done with a rear system, but such systems often come with a bulky piece of furniture and do not allow for augmenting the objects placed on the tabletop.

In both the top-down and rear configuration, interactive surfaces are well-suited for collaborative learning activities, and numerous activities have been implemented that make use of interactive surfaces.

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 teenager apprentices by allowing them to build warehouses and evaluate them. The system, shown in Figure 2.9, 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.

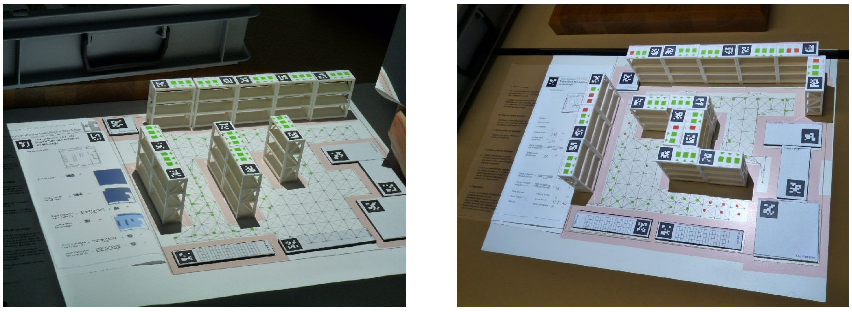


Figure 2.9: TinkerLogistics (Zufferey et al., 2009).

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 (Figure 2.10). The system reads the program, i.e. the assembly of puzzle pieces through a camera, and command a robot to execute it.

Rear systems have been more popular than top-down ones, most likely because of the availability of commercial products such as the Microsoft PixelSense or the Reactable, leading to a greater coverage of learning domains. For instance, Price et al. (2009) created a tangible tabletop interface to support children learning about the behavior of light (Figure 2.11). Using a torch and tangible elements marked with fiducial markers, children can explore concepts such as reflection, transmission, absorption and refraction of light.

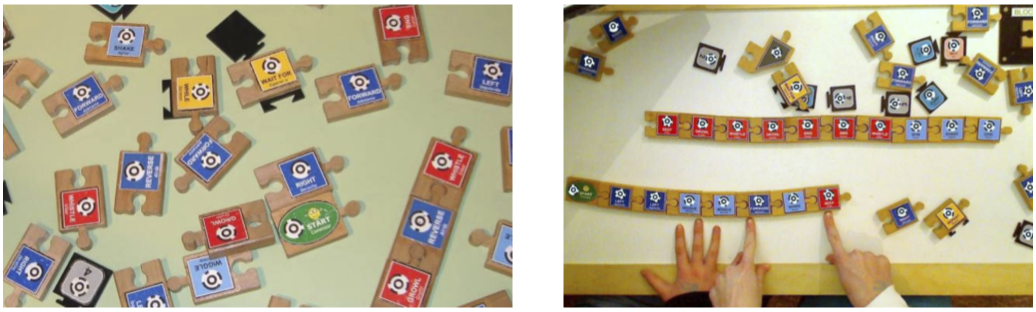


Figure 2.10: Programming a robot by assembling puzzle pieces (Horn et al., 2009).

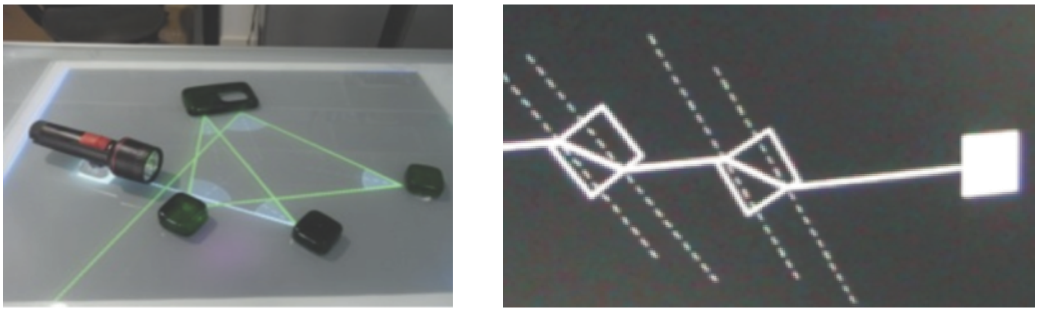


Figure 2.11: A tabletop system to learn about the behavior of light (Price et al., 2009).

The work by Antle et al. (2011) also features a rear system. Their system was designed for children to learn about sustainable land use planning. The users see a map of a region and their task is to assign land use and activities to the various regions on the map by using physical stamps (Figure 2.12). The particularity of this system is that it also includes a separate station on which children can get more information on the stamps.

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.

2.2.3 Possible benefits of tangible tabletops for learning

There are many reasons why tabletop TUIs could benefit learning. Those benefits stem from the tangible nature of the interface as well as from their tabletop component. I classified them under five categories.

Increased usability

We are used to manipulate physical objects as we do it daily in our lives. One expectation of tangibles is that the learners' familiarity with the manipulation of physical objects may increase usability and limit the cognitive effort dedicated to interact with the system, and subsequently

Figure 2.12: Learning the concepts of sustainable land use planning (Antle et al., 2011).

allow them to focus on the core of their task (Manches and O'Malley, 2012). Fitzmaurice and Buxton (1997) showed that tangible interaction could increase speed and efficiency compared to a GUI. Further research showed that TUIs could help organize information effectively because of the spatial organization that they allowed (Jacob et al., 2002; Patten and Ishii, 2000). However, Hornecker (2012) recently challenged this claim and encouraged TUI designers to reflect on how they could create "seamful" mappings (as opposed to seamless) to increase users' awareness and understanding of the systems they use.

The impact of the increased usability on learning has also been explored. Comparing a tangible interface with a multi-touch one for a problem-solving task, Schneider et al. (2011) showed that the tangible interface helped participants attain a superior learning gain through higher exploration. Similarly, 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 easiness of manipulation can sometimes be counterproductive for learning if it leads to too much manipulation (Do-Lenh et al., 2010; Do-Lenh, 2012; Price et al., 2009). This is in line with Manches et al. (2010)'s findings that the design of a TUI, and especially the manipulative properties of interfaces, can significantly impact the learning strategies or performance.

A side aspect of the increased usability is that tangibles can be more fun than traditional computer interfaces and thus lead to a higher engagement. Several studies have observed such tendency among children (e.g. Price et al., 2003; Xie et al., 2008; Brederode et al., 2005; Sylla et al., 2012). In some cases, the higher engagement can even make user look over some usability issues (Zuckerman and Gal-Oz, 2013). Doubts have however been raised about causality between engagement and learning (McNeil and Jarvin, 2007). Also, while higher engagement is perceivable with children, it has yet to be shown that this can also be the case with older users.

Physicality and link with cognition

Following the claims of constructionism, the physicality of tangibles brings potential benefits for learning. 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). This is similar to Piaget's developmental theory (Piaget, 1974), which asserts that a child builds empirical abstraction, and subsequently knowledge, from his physical actions. However, as mentioned above, those claims have recently been disputed by several researchers (McNeil and Jarvin, 2007; Klahr et al., 2007; Triona and Klahr, 2003; Zaman et al., 2012), and the relevance of applying the conclusion of Piaget's work, who worked with young children, to TUIs in general, can be questioned.

Recent research has also underlined the limits of conceptual metaphors linked with physical objects, pointing out that symbolic representations are limited to the context in which the objects are used (Manches and O'Malley, 2012). It may also require an additional cognitive effort for the user to understand to which extent an object must be considered as symbolic or as an object of interest.

Claims that physicality can be especially beneficial for spatial tasks have emerged early (Sharlin et al., 2004). As Marshall (2007) puts it: "three-dimensional forms might be perceived and understood more readily through haptic and proprioceptive perception of tangible representations than through visual representation alone". This indeed echoes some results in spatial imagery research, which showed that spatial representations provide information without intensive cognitive effort, that it can only be interpreted in a single way at a given time, and that because perception occurs before one's belief, people can build mental representations before they get locked into their beliefs (Schwartz and Heiser, 2006).

Multiple external representations

As put by Zhang and Norman (1994), "it is the interwoven processing of internal and external information that generates much of a person's intelligent behaviour". Zhang and Norman showed that using external representations can lead to a higher performance when solving a distributed cognitive task (a task that requires the processing of information distributed across the internal mind and the external environment), because they can reduce the load on the working memory, anchor and structure cognitive behaviour, and provide information that can be directly perceived and used without further processing.

The readiness of the information available in external representations was made explicit by comparing diagrams and text (Larkin and Simon, 1987; Scaife and Rogers, 1996). Larkin and Simon (1987) detailed the difference between diagrammatic representations, which are indexed by location in the plane and carry implicit information, and sentential representations, which are linear and explicit. They contrasted the computational efficiency of these two representations for solving problems in mathematics and physics and concluded that the advantages of diagrams are computational: "they can be better representations not because they contain more information, but because the indexing of this information can support extremely useful and efficient computational processes."

Tangibles can serve as external representations. Price et al. (2009) found that different geographical locations of a representation led to an understanding of concepts at different levels of abstraction. They also observed the limit of the mapping between real-world and virtual objects, when the virtual torch that was represented in the real-world by a torch could not be turned on the same way as in the real-world. As already mentioned for physicality, tangibles, through the perception and sensations that their physicality afford, make some information readily available and can complement some other kinds of representations (Schwartz and Heiser, 2006).

Having more than one external representations at a time plays a key role in problem solving and learning by helping the learner to make inferences, find invariants, and gather information in a different fashion from each of the representations (Ainsworth, 1999; Larkin and Simon, 1987; Price, 2008a). The potential benefits of multiple external representations (MERs) for learning are summarized in Ainsworth (1999), which provides a functional taxonomy of MERs. According to this taxonomy, MERs can help learn by providing complementary processes and information, by constraining interpretation, and by helping to construct deeper understanding.

However, the use of MERs does not guarantee an improved learning outcome. Indeed, it might be hard for the learner to make sense of a dual representation (McNeil and Jarvin, 2007), and depending on their use, MERs can excessively increase or decrease the cognitive load upon the learner. Ainsworth (2006) presented a framework and design guidelines that aim at avoiding the increase of cognitive load. She proposed that three fundamental aspects impact the effectiveness of MERs: the functions that MERs serve in supporting learning, the cognitive tasks given to the learner, and the design parameters that are unique to learning with MERs. Using this framework, she explains how using an additional representation led to an increase in learning when it was used to constrain interpretation, and to a lower learning performance when the additional representation complemented other representations.

Dyna-linking

One advantage of MERs is that they allow the user to see how his action on one representation impacts a second representation (Ainsworth, 2006). This dynamic link between two or more different representations is referred to as dyna-linking. Dyna-linking can be of importance in learning tasks, because it can reduce cognitive load through offloading (Ainsworth, 2006; Manches and O'Malley, 2012), therefore allowing the students to concentrate on the actions they exert on the representations and their consequences. Price et al. (2009) showed that dyna-linking can also encourage learners to try to understand concepts that are beyond their level of understanding by encouraging them explore more. One important dimension for learning is the coupling between cognition and physical experience (O'Malley and Stanton Fraser, 2004). This is important with tangibles, as they are not only external representations but also physical objects on which *actions* can be exerted.

Collaboration, co-location and simultaneous interactions

Tangible tabletops provide a shared workspace to their users. Having a shared workspace helps learners be aware of each other's actions, which in turn leads to fluid collaboration and interaction. This result has been empirically confirmed by several studies (e.g. Hornecker et al., 2008; Ha et al., 2006; Rogers and Lindley, 2004; Hornecker and Buur, 2006).

Rogers and Lindley (2004) studied the effect of the orientation of the shared display in a group setting. Comparing vertical versus horizontal display, they found that when using a horizontal display participants had a higher level of awareness of their mutual actions and that they generated more ideas.

Using in both cases a horizontal 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) found 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.

But tangibles do not always promote collaboration. Marshall et al. (2009) pointed out that sometimes the learners do not strive to collaborate, and that physical objects might help hamper collaboration. They compared the collaborative use of non-augmented physical object with the use of digital objects present on a multi-touch interactive surface. The results showed that when disputes arose over a physical object, the children using the system resolved them by moving the object out of reach or blocking access to it. Thus the tangibility of the interface served to break collaboration rather than to foster it.

Finally, Speelpenning et al. (2011) observed that "the physicality of the tangible tools facilitated individual ownership and announcement of tool use, which in turn supported group and tool awareness." Their work distinguishes four concepts, inspired from the CSCL literature, that can impact collaboration: objects of negotiation, access points, physical constraints, and awareness. Comparing TUIs and multi-touch input, they found that the awareness and negotiation for objects was greater in the TUI condition. However, they also found that the usage depended greatly on group dynamics and not only on the type of interface.

2.2.4 Conceptual frameworks

Several frameworks and taxonomies have been developed since the advent of TUIs. They have mostly focused on the integration of representations between the tangible and the digital objects (Koleva et al., 2003; Ullmer et al., 2005; Fishkin, 2004; Holmquist et al., 1999), the impact of concreteness of TUIs (Zuckerman et al., 2005), the embodied interaction tangibles lend themselves to (Dourish, 2004), and the social aspects of TUIs (Hornecker and Buur, 2006). One of the most recent frameworks, the Tangible Learning Design Framework (Antle and Wise, 2013), builds upon previous frameworks and explores the TUI elements (physical objects, digital objects, actions on object, informational relations, learning activities) from several perspectives (constructivist, cognitive, embodied, distributed, collaborative).

The framework by Price (2008b) describes the relationship between artifacts, actions, and external representations, and how these relationships affect cognition in terms of inferences and conceptual understanding. It distinguishes four parameters: location, information flow, correspondence, and modality. These parameters are particularly relevant to the work presented in this thesis and I will thus describe them in more detail.

Location is a parameter that specifies the location of the representation in the physical space. Location can be discrete, co-located, or embedded. In the discrete case, the input and the output are located separately. In the co-located case, input and output are contiguous. In the embedded case a digital effect occur within an object. Location can impact attention demands, ease of problem solving, potential for representing multiple levels of abstraction, and the kind of actions that can take place. Price et al. (2009) showed empirically that the level of abstraction that learners were able to make out from a concept was influenced by the location of the digital representation in the TUI.

Dynamics describes how information is put in or out of the system. The way in which the system outputs feedback has an impact on the learning experience, since students infer causality between their actions and the output of the system. A simple causality between action and output (a single action followed by an output) may be easy to interpret, but lead to less exploration than a complex causality. The way in which information is input in the system is measured by intentionality. An input can be either intentional when digital effects occur together with an action, generating an expected effect, or serendipitous, when digital effects are unexpected and the result of actions.

The **correspondence** parameter categorizes the mapping between objects, actions, representations, and the learning domain. There are two types of correspondence, physical and representational. The physical correspondence is the extent to which an artifact possesses the physical properties of the object that it represents. The two extreme values on the physical correspondence scale are "symbolic" and "literal". Symbolic objects have no characteristics of the object they represent, whereas literal objects are closely mapped to the domain they represent thanks to their physical properties. The representational correspondence quantifies the meaning mapping between the digital object and its physical representation. It can greatly impact the level of cognitive load needed by the learner.

Location, information flow, and correspondence are parameters that help analyze a tangible environment. However, they do not mention the type of **modality** on which the interaction between the user and the system is based. Although most information is usually represented visually in existing systems, one should not forget the power of the tactile and audio representations. Varying modalities can be used to reduce screen overload, grab the user's attention more easily, or output feedback in a different way.

2.3 Computer-Supported Collaborative Learning

As explained in the introduction of this chapter, this work lies at the intersection of learning sciences and computer sciences, with a particular emphasis on computer-supported collaborative learning (CSCL).

Although the definition of collaborative learning has been disputed (Dillenbourg, 1999), it can broadly be defined as activities in which two or more learners share and construct knowledge and understanding together (Roschelle and Teasley, 1995). Collaborative learning emerged as a research field in the 1980s. Educational researchers started to embrace social constructivism (Vygotsky and Cole, 1978) instead of individualistic view of learning. Early research focused on finding out empirical evidence on whether, and if so when, learning in collaborative fashion could be more effective than learning alone. However, researchers never managed to identify causal links between conditions and learning effects, and the global research question in the field evolved from "is collaboration positive for learning" to "under which conditions is collaborative learning effective" (Dillenbourg et al., 1995). This question was then further refined into two sub-questions: "which interactions occur under which conditions" and "what effects do these interactions have on learning".

The computer can enhance the benefits from collaborative learning situations, for instance by helping distribute roles between the learners, regulating social interaction (O'Malley, 1992), or fostering mechanisms such as mutual understanding and argumentation (see Blaye and Light, 1990, for a review). The computer can also help the teacher or tutor regulate the learning of a group in a more effective way (Dillenbourg, 2002; Bachour et al., 2010). However, the computer can also have a negative effect on learning, such as reducing discussion between learners by providing immediate feedback (Fraisse, 1987). In fact, much like the collaborative aspect of learning, the computer is not good or bad per se for learning. Its effectiveness depends on the way in which it is used as well as on the conditions in which it is used (number of students collaborating together, the type of tasks, the difference in levels between the students, etc.). Computer systems for learning need to have an integrative approach, as suggested by Rogers and Ellis (1994), who argued that to be effective, collaboration had to take into account the cognitive, social, and organizational aspects all together, instead of looking at them separately. They claimed that systems designed solely from the cognitive perspective became unusable once put in the context in which they were supposed to be used.

A good example of computer systems that sustain learning are tools that support scripting. Scripting is a way to make collaboration more productive in terms of learning. A collaboration script is defined as "a set of instructions prescribing how students should form groups, how they should interact and collaborate and how they should solve the problem" (Dillenbourg, 2002). Scripting is an effort to directly influence the interactions between co-learners with the goal of making those interactions more beneficial for learning. An example of how computers can support scripting is ArgueGraph (Jermann and Dillenbourg, 1999), a system that automatically pair students so as to maximize the distance in their belief about a topic. The idea behind ArgueGraph is to increase the chances students have to argue and negotiate a common answer, a type of interaction that has been shown to affect learning positively (e.g. Butterworth and Light, 1982)

While scripting is an effective way to foster productive collaboration, it does not consider the environment in which it happens. Recently, a movement in the CSCL community has emerged that proposes to consider organizational and social aspects, much as suggested by Rogers and Ellis (1994), but focusing on the classroom environment. This has given name to new theory called classroom orchestration (Dillenbourg et al., 2012). I will now explain the main principles of classroom orchestration and then give examples of some systems successfully deployed in classrooms.

2.3.1 Classroom orchestration

Despite all the research in technology enhanced learning, a gap remains between this research and the learning technologies available in classrooms (Roschelle et al., 2009; Dillenbourg et al., 2012). In fact, classrooms using computer technology are nowadays the exception rather than the majority. This is not for the lack of potential of computer technology for learning, since computer can support fundamental characteristics of learning, such as active engagement, participation in groups, frequent interaction and feedback, and connections to real-world contexts (Roschelle et al., 2000).

A classroom setting is a complex and peculiar one. Don Norman's assertion about everyday life applies perfectly to the classroom: "It is complex, not because any given activity is complex, but because there are so many apparently simple activities, each with its own set of idiosyncratic requirements. Take a large number of simple actions and add them up, and the overall result can be complex and confusing: the whole is greater than the sum of its parts." (Norman, 2011, p.64).

Therefore, as argued by Dillenbourg and Jermann (2010a), 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.

Those 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 technology considered by the teacher to "work well" should also be given weight (Dillenbourg et al., 2012). They put emphasis on other constraints not directly linked to learning, but to *teaching* in the classroom. For example, they detail the tension between the simultaneous need for routines and improvisation (Dillenbourg et al., 2012; Prieto et al., 2011). They analyze the dynamics of the classroom and identify three levels of interaction in the classroom: individual, group, and class levels (Dillenbourg et al., 2011). This 3-layer notation is similar to what was previously observed in the field of social psychology (see e.g. McGrath et al. (2000)).

Classroom orchestration has been a rising topic in the CSCL community in the last few years (Dillenbourg et al., 2009), and scholars are still in the process of defining it exactly (Dillenbourg et al., 2012). The core idea of classroom orchestration is that a learning technology in the classroom should address the many constraints of classroom management, which are typically the explicit constraints described in the previous section. These constraints have often been somehow neglected in CSCL either because studies where conducted in labs or because technologies were designed to be used "anywhere". Recently, Dillenbourg and Jermann (2010a) have gone further and argued that the constraints of the entire learning environment be taken into account, and have therefore added three layers to the three (individual, group, classroom): the periphery, the community, and the world.

Technology can serve classroom orchestration in two ways: by developing tools with the specific purpose of facilitating orchestration (e.g. Alavi et al., 2009), or by changing existing technology to make it more "orchestrable". The difficulty that the teacher faces when orchestrating the classroom is referred to as the "orchestration load".

2.3.2 Examples of successful classroom integration

Although the focus on classroom orchestration and taking into account the constraints of the classroom into account is relatively recent, there are learning systems that have been used successfully in classrooms.

ClassSearch (Moraveji et al., 2011), a tool for web search skills acquisition, was used in classrooms of students aged 11-14 years old. It met a discipline constraint by encouraging students to stay on-task and also minimized the teacher's effort by displaying a search history on every student's screen. However, the learning task was not part of the core curriculum.

One Mouse 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. It fulfills a financial and effort constraint by using only one computer, and offers easy monitoring for the teacher. The discipline was a problem in this case, because one student could easily disrupt the entire class. Nevertheless, the system was used successfully in classrooms with varying socioeconomics contexts in India and Chile and helped the weakest students improve their arithmetic skills.

One of the main successes of classroom integration is Group Scribbles (Dimitriadis et al., 2007), a GUI-based approach 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. With Group Scribbles, students can perform a wide range of cooperative activities and were reported to be more involved in contributing and responding to content, potentially easing the discipline of the class. The system also increases awareness of students' actions and state for the teacher.

The most relevant example of classroom integration to this research is the logistics environment created by Zufferey et al. (2009) and extended by Do-Lenh et al. (2010), presented earlier in this chapter. Do-Lenh (2012) was able to show 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 distantly. The activity developed by Do-Lenh (2012) was the first one to span over the four first planes of the classroom orchestration model by making the logistic apprentices show their supervisor, back at the workplace, what they had accomplished at school. The system is still being used in several schools.

Topobo (Parkes et al., 2008), the 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. The message from educators was 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 this 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 classroom in Brazil.

2.4 Challenges and Open Questions

While not in its infancy anymore, research on TUIs (and on learning with TUIs) is still quite recent and many questions remain open. The review of the past work made in this chapter has shown that the work on TUIs has so far mostly focused on prototypes and the production of conceptual frameworks and taxonomies. These works are often conducted with the goal of exploring the range of possibilities offered by TUIs. There is still much room for such explorations, especially in domains such as user fatigue, or malleability of the interface. However, this thesis does not go in the direction of exploring more such possibilities. Rather, it is focused on gathering empirical evidence for two aspects of TUIs: their impact on learning, and their integration in the classroom environment. Although the body of research tackling these questions is increasing, as shown by the survey presented in this chapter, this is still a new approach with many questions yet to be answered.

2.4.1 Lack of empirical knowledge on learning effectiveness

A few years ago, Marshall (2007) questioned whether TUIs really enhanced learning. He noticed that research on TUIs so far had mostly focused on exploring the possible combinations between physical and digital representations and technical novelty and that the rare empirical studies mostly focused on the motivating and engaging aspects of the interface. He reviewed the literature and showed that the claimed benefits of TUIs for learning were mostly based on assumptions and common sense and that empirical evidence supporting them was lacking. He encouraged the community to provide empirical evidence detailing "what features of a tangible interface design might be associated with particular benefits to interaction or learning and what features might be more incidental".

This was several years ago and since then some researchers have focused on theoretical aspects of TUIs for learning (Price, 2008b; Antle and Wise, 2013) as well as providing empirical evidence from field experiments (Horn et al., 2009; Zufferey et al., 2009; Do-Lenh et al., 2010). However, despite these efforts, many of the questions asked by Marshall still remain open today (Zaman et al., 2012; Antle, 2012). In a recent overview of the potential benefits of tangibles for children, Zaman et al. (2012) ask for caution in being too positive about tangibles. The authors describe three types of studies through which tangibles have mostly been evaluated (single case evaluation studies, comparative evaluation studies, and design case studies). While they acknowledge that "understanding the full impact of tangibility is only possible through a variety of studies with different research models" (p. 9), they also argue for the results to be taken within the context of the study in which they were found.

The question asked by Marshall, "Do tangible interfaces enhance learning?" is broad and answering it fully goes beyond the scope of a single thesis. Moreover, one could question the validity of such a question based on the 'media effect myth' (Dillenbourg, 2008). The media effect myth states that technology or any other kinds of media can have an intrinsic influence on learning, but that its influence on learning depends on how its specific features are used for learning. This is for example what has been observed in the CSCL field, where the research questions evolved from "can computers help students learn?" to more subtle questions on what social interactions should be triggered to improve learning and how computers could be used to trigger these interactions. Similarly, Marshall's question can be rephrased into "what processes can tangible interfaces trigger, under what circumstances?" and "what effects do these processes have on learning?".

2.4.2 Lack of deployment in the classroom

Another issue with the research on TUIs so far is that the works that go beyond prototyping and end up actually being deployed in a classroom are few and far between. The classroom is a complex environment with many constraints and fitting technology into a classroom is no trivial task. Some studies have reported on deployment of tabletop or tangible technologies in the classroom (Piper and Hollan, 2009; Zufferey et al., 2009; Do-Lenh et al., 2010; Parkes et al., 2008; Blikstein, 2008) and the CSCL community has recently increased its focused on the classroom environment, as the interest in the concept of classroom orchestration shows. However, there is still a lack of guidelines to inform the design of tabletop TUIs for the particular context of the classroom.

A typical teacher will use many learning scenarios during his class, and it is unclear whether TUIs can be introduced in existing learning scenarios or if there is a need to design new learning scenarios to make the best of them. Also, because of the lack of deployment in the classroom, practical problems such as the cost or the logistics of the deployment have not received much attention.

2.4.3 Lack of scalability

One of the reasons why TUIs are not more well-spread nowadays is because they are hard to scale up. Scaling up can be understood in many ways: broadening the user basis, extending a system to new application domains, or allowing a usage of the application at more places. Issues when scaling up tangibles include the cluttering of objects on a limited physical space, the creation of many physical objects and the potential ensuing costs, or the difficulty of saving and restoring a given state (Shaer and Hornecker, 2009). Many TUIs also require a dedicated (and often expensive) hardware environment, which limits their geographical dissemination.

Studying in more detail the constraints linked to scalability and how they can be dealt with is worth pursuing if we want TUIs to scale up one day. While some aspects of scalability, such as the cluttering of objects, are inherent in the tangible nature of the interface, other aspects can certainly be improved.

3 Context

The previous chapter reviewed the existing body of research on TUIs for learning and highlighted the need for more research to understand their impact on learning. The underlined limitations are quite general, and while the goal of every research is admittedly to draw as general conclusions as possible, one needs to start with a specific learning environment and learning topic to be able to empirically evaluate an interface.

This chapter first describes the broader educational context in which this work took place: initial vocational education and training (VET) in Switzerland. It then goes on to explain in more detail the role of carpenters, and the challenges that they face in their training.

3.1 Vocational Education and Training

In Switzerland, a majority of the teenage population is involved in vocational education and training. As of 2009, more than two thirds of the population starting a new training after mandatory school enrolled into VET (Gaillard, 2011, p.48). In many countries, vocational education is often thought of being a last resort and looked down on. However, this is not the case in Switzerland, where vocational training is quite popular and highly regarded, with 5.7 times as many vocational certifications delivered as high school degrees (Gaillard, 2011, p.48). About 30% of Swiss companies take part in the vocational training, and on average the benefits of training apprentices slightly exceed the costs inherent in it (OPET, 2012).

Almost 200 professions, including car mechanic, hair stylist, horticulturist, precision mechanic, or carpenter, can be learned through VET. The most common training is organized as an apprenticeship, which lasts from two to four years depending on the profession. Upon successful completion of their apprenticeship, apprentices are considered as qualified workers in their field and receive a Federal Certificate of Capacity. There is also a possibility to continue the training further into higher education schools after having completed the initial training. This participates to the attractiveness of the vocational training path, as it provides an early work experience without shutting the door of future progression.

This work focuses on the initial training, i.e. the apprenticeship. Two systems coexist for the training of apprentices: full-time school training, and the dual one. The dominant approach is the dual one (88% of all apprentices): apprentices go to school one day a week and work within a company for the remaining four days. The day spent at school is used to teach profession specific knowledge (5 lesson hours) as well as general knowledge (3 lesson hours). The practical knowledge is acquired during the remaining four days spent in the company and during intensive practical courses that generally occur once a year over 2 to 3 weeks. In order to be allowed to take on an apprentice, a company must provide a supervisor to the apprentice. The supervisor is typically a worker with several years of experience, who was trained in the same field as the one pursued by the apprentice, and followed a short training about supervision.

The role of the school is to make sure that every apprentice acquires a certain level in their profession and an exposure to all facets of their job. Indeed, work practices can differ across companies based on various factors, such as their size or their geographical location, and depending on the company they work for and their role in the company, apprentices may get limited exposure to important aspects of their profession. For instance, a carpenter apprentice working in a small village in the mountains might only get to build Swiss chalets, or an apprentice working in a larger company might not get exposure to the various machines to cut wood because the cutting is done automatically by a computerized numerical control (CNC). In theory, companies must expose their apprentices to all the aspects of the profession, but the school guarantees further a certain homogeneity and common basic skills to all apprentices of a given trade.

At first sight, the dual system seems pedagogically ideal, with the authentic practice within the workplace being complemented by the conceptual and theoretical knowledge gained at school. However, the workplace and the school are two different environments, with different constraints, and this can lead to a divide between them. On the one hand, companies evolve in a fast and competitive environment and are required to constantly adapt their practices based on technological progress and industry standard changes. On the other hand, schools favor continuity based on a curriculum that does not change often. Moreover, many teachers are not connected to the professional world anymore and their knowledge may get outdated. For apprentices, this leads to a gap between the practical work in the company and the theoretical teaching at school, making the school sometimes look irrelevant to the tasks they accomplish in the company. The next section explains how this gap takes place in the training of carpenter apprentices.

3.2 Training for Carpenters in Switzerland

3.2.1 Overview

In the Swiss vocational system, there are two different training paths for construction carpentry ("charpentier" in French) and furniture making ("menuisier"). This thesis focuses on construction carpenters, who construct large items, such as roof structures or even entire buildings (see Figure 3.1).

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 extremely low number of women (about one percent) is explained by cultural and historical factors.

(a) (b) (c)

Figure 3.1: Typical structures built by carpenters.

3.2.2 Tasks

The job of a carpenter is to prepare, cut, and assemble wood pieces to create the frames and roofs of buildings. 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.

At the end of his apprenticeship, a carpenter is considered trained and must be able to perform all these steps. In order to achieve this level of competency, the school study plan for carpenters is built around three main disciplines: material and technology knowledge (20%), drawing (60%), and arithmetic (20%). All three subjects are taught in the specific context of carpentry.

The carpenter apprenticeship currently lasts for three years (it will be extended by one year in 2014). Following the general VET scheme, the training happens in 3 places: the professional school (1 day a week), the company (4 day a week), and the inter-company courses, a few weeks per year for a total of 32 days over the 3 years.

The inter-company courses are supported by the companies through the central carpentry association. Their goal is to allow the apprentices to develop their practical skills under the supervision of a teacher, and away from the stress and pressure of a company environment. During these courses, apprentices come to a workshop (Figure 3.2) and build wood items that require them to go through the main steps of carpentry: read plans, produce working drawing, cut wood, and assemble pieces together. In the rest of this thesis, I will refer to the school where the inter-company courses are held as the "practical school" and to the one where apprentices go on a weekly basis as the "theoretical school".

(a) The main workshop.

(b) A drawing board.

(c) The machine room.

(d) An apprentice production.

Figure 3.2: The school where the inter-company courses take place.

3.2.3 Contextual inquiry

During the first few months of my PhD project, my supervisor I conducted a contextual inquiry to learn more about the Swiss apprenticeship program of carpenters. We visited two schools and five companies, and interviewed several teachers, directors of companies, apprentices' supervisors, and apprentices. We had had no prior exposure to the world of carpentry. These visits were therefore useful to gain knowledge about the profession. They also allowed us to discover the disagreements between the various stakeholders, and especially those between the director of companies and the school teachers.

The visits to the companies lasted approximately half a day. They happened either at the workshop or at a construction site, sometimes both successively. When at the workshop, we first asked the apprentice to give us a guided tour of his workplace and then conducted a semi-structured interview with him. During the interview, we asked questions about the type of activities his job entailed, at what frequency, which ones he preferred, and how they related to what he was doing at school. When the visit happened on a construction site, it was not possible to talk to the apprentice for a long time as he was expected to work, and keeping him out of the workforce for too long could affect the progress of the construction. Every company visit also included an interview with either the apprentice's supervisor or the director of the company. Additionally to the company visits, we also observed the activities at a theoretical school on several occasions and interviewed two teachers. We also visited a practical school and interviewed the teacher there too.

3.2.4 School and companies: two divergent views

The visits to the schools and to the companies revealed the different perspective that stakeholders have on how to train professionals, depending on whether they are active in companies or in schools.

By visiting the schools and studying the learning material provided to carpenter apprentices, we established that about 60% of the profession-specific teaching time is dedicated to drawing. This includes learning the basics of descriptive geometry and how to draw carpentry plans. The remaining 40% are split between selected mathematics subjects applied to carpentry, studying physics building and building materials, and structural mechanics.

The bottom line from the company stakeholders' was that the school had not adapted to the changes that the profession had undergone and that the training it was dispensing was mostly irrelevant. They pointed out that with the recent advent of new technology such as computer-aided design (CAD) software and computerized numerical control (CNC), and with the new construction standards, the needs for training of new carpenters had drastically changed. According to them, the need to learn drawing for new generations of carpenters is much weaker nowadays than it was before, the same way that the need for them to know how to use a hand saw is weaker nowadays than it was before the advent of the electric saw.

The directors further argued for the reduction of the time dedicated to drawing at school, emphasizing that drawing was merely a legacy from the times when carpenters needed to draw complicated structures by hand, and that few carpenters drew plans by hand nowadays. Insisting on the metamorphosis of the profession, they explained that most of the work of a modern carpenter is to perform other tasks than wood work, such as applying insulation to the roof, or roof covering. In their opinion, a significantly larger part of the training should be dedicated to physics building and structural mechanics instead of drawing.

Apprentices confirmed that, in practice, they almost never draw anything, since most of the plans are made by a more experienced co-worker, be it with a CAD software or by hand. While not all of them disliked drawing, most of them said they did not understand why they have to spend so much time learning how to draw when drawing will most likely never be part of their work tasks.

The teachers, on the other hand, had a different take on the subject: they insisted on the fact that drawing was the basis of the profession and that it should definitely not be abandoned. They acknowledged that drawing was not per se used in the professional environment anymore, but emphasized that it was key to learn the concepts of the profession, helped apprentices learn to read plans, and helped develop their spatial skills.

The latter turned out to be the only point on which company directors and teachers agreed: being a carpenter requires excellent spatial skills. Indeed the information regarding the physical, three-dimensional (3D) items that carpenters have to build is conveyed by means of two-dimensional (2D) paper plans. Tasks such as transitioning from the paper plans to the final object, imagining how several beams are going to fit together, are carpenters' bread-and-butter and require excellent spatial skills.

3.3 Conclusions

In this chapter, the basics of VET in Switzerland, as well as the role of carpenters, were explained. The results of a contextual inquiry revealed the existence of a gap between the skills needed to be a carpenter, as perceived by the companies on the one hand, and the skills that are developed at school, on the other hand. One subject on which the school and the company stakeholders agreed was the need for excellent spatial skills in order to be a professional carpenter.

Carpenters' spatial skills have never been explicitly trained before, because they are assumed to be developed through the drawing classes. There are two implicit assumptions here. First, that spatial skills are trainable. Second, that the drawing done by carpenter apprentices indeed trains their spatial skills. The validity of these two assumptions can be captured through the two following questions, which will be addressed in the next chapter through a literature review: (1) are spatial skills trainable?; and (2) can they be trained in a different way than through drawing exercises?

These two questions will not be the main research questions of this thesis. The training of spatial skills in the context of carpentry training is merely the learning subject that was needed in order to be able to assess the effectiveness of a tangible user interface on learning. The main research question that this thesis will explore, and that is detailed in Chapter 5, is the following one: Can a tangible user interface help acquire spatial skills, and if so, how?

When we started this project, our goal was not to eradicate drawing from the carpenters' curriculum. Drawing remains part of the professional culture of carpenters. Rather, our vision is to keep drawing activities as a way to legitimate the apprentices position in their community, but to give these activities a time budget that is better adapted to the functional importance of drawing skills in their practices. The tension between tradition and adaptation is common to many professions. The solutions explored in this thesis are specific to carpenters, but the underlying approach could be adapted to many contexts.

4 Spatial Skills

The previous chapter highlighted the importance of spatial skills for carpenters. This chapter expands on spatial skills, and on spatial skills in the context of carpenters. It first presents spatial skills in more details, and then describes the results of a field study conducted to assess spatial skills among carpenter apprentices. For the data collection of this study, I received help from Christoph Arn and Engin Bumbacher.

4.1 What Do I Mean by "Spatial Skills"?

Spatial cognition addresses how humans acquire and encode spatial information, and how it is represented in memory and manipulated internally (Quarles et al., 2008; Hegarty et al., 2006). It is a specific and important kind of cognition, to the point that Gardner (2006), in his original theory on multiple intelligences, identified spatial intelligence as one of the seven types of intelligence that humans possess.

In his work, Piaget distinguished three types of spatial skills, which are developed in three successive stages: topological, projective, and euclidean skills (Bishop, 1978; Sorby, 2009). Up to age 5, children develop topological skills, which are 2D skills that allow them to assess an objects' closeness to others. During the second stage (typically up to the beginning of adolescence), individuals develop projective skills, which involve the ability to visualize 3D objects and imagine how they will look like from different points of view. Finally, euclidean skills combine projective skills with concepts of measurements, such as distance, area, and volume. Many individuals never acquire these skills.

The development of spatial skills is contentious and Piaget's classification is only one among others. Nevertheless, a separation between two main dimensions in spatial ability is usually accepted: visualization (e.g. mental rotation, paper folding) and navigation (e.g. map orientation). These two dimensions have also been defined as "scales" of spatial ability (Hegarty et al., 2006), visualization being considered small-scale and navigation large-scale. Ability in each of these two dimensions may be independent. Uttal et al. (2012) have defined four other classifications that are based on the type of information that a task requires (extrinsic

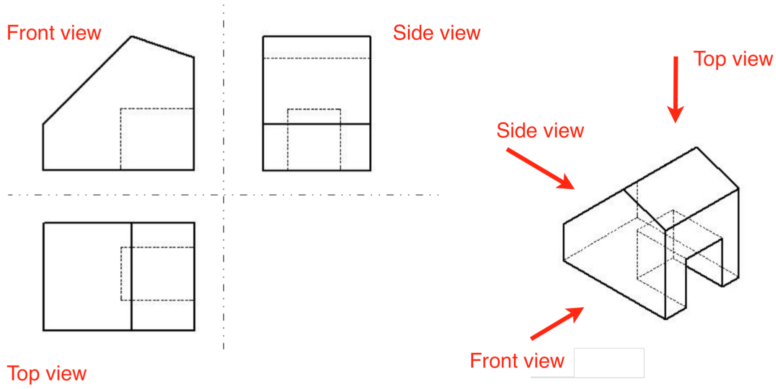


Figure 4.1: Orthographic projections.

versus intrinsic) and whether or not it is dynamic or static. For example, a mental rotation task would be classified as dynamic (one has to mentally turn the object) and intrinsic, since it requires a comparison of the parts of an object, as opposed to a comparison of the object with its environment.

For carpenters, while all types of spatial skills are certainly useful, the most needed one is visualization (or intrinsic, both static and dynamic). This is the skill that allows them to go back and forth between the 2D representation of an object, given by the plan, and the 3D object itself. With good spatial skills, they can imagine, from a plan, how a building will look like, how beams will be assembled, or check on a plan whether what they built is correct. In the rest of my thesis, when I refer to "spatial skills" I mean the spatial visualization skills that carpenters use to imagine a 3D object from a 2D paper representation, and the other way around. For carpenters, the 2D representation is given by orthographic projections, which are a kind of parallel projection where all lines are orthogonal to the projection plane. Figure 4.1 shows the orthographic projections and a perspective 3D representation of it. The 3D object can be found based on the 3 orthographic projections if there are no coinciding projections of vertices (Lafue, 1976).

4.2 Importance of Spatial Skills

As pointed out by Sorby (2009), "spatial skills have been a significant area of research in educational psychology since the 1920s or 30s". Educational systems, though, have often neglected and looked down on visual thinking, favoring instead other cognitive skills (Sommer, 1978; Arnheim, 1980; Smith, 1964; Gardner, 2006). In the last decades, the interest for spatial skills has increased, leading to a growing body of research on understanding, assessing, and developing means of training spatial skills.

In a study involving 400,000 participants of grades 9 to 12 over 11 years, Wai et al. (2009) studied the link between spatial ability and STEM domains (science, technology, engineering, and mathematics). Their results showed that spatial ability has a strong influence on the development of expertise in STEM fields. For example, individuals holding at least a Bachelor's degree in engineering have spatial skills that are more than one-and-a-half standard deviation higher than the rest of the population. Other researchers also suggested that spatial ability assessment could be used to improve the detection of teenager talents for the STEM domains (Shea et al., 2001), or that improving mathematics and science skills could be achieved by improving spatial thinking (Newcombe, 2010; Sorby et al., 2013). Other researchers also hinted that improving spatial skills could reduce the number of university dropouts in engineering studies (e.g. Sorby, 2009).

Spatial skills are not only important for school subjects. In the professional world, spatial visualization skills and mental rotation abilities are important for technical professions (Maier, 1994). The professions that come first to mind are those that require a high level in STEM domains, such as mechanical engineer and architect. However, there are also some more unexpected professions for which spatial skills play an important role. For instance, Hegarty et al. (2009) showed that dentistry students develop mental models of teeth. Hambrick et al. (2011) found that novice geologists' spatial ability was correlated with their score at geology mapping tasks. Several studies showed the importance of spatial thinking for chemical sciences (e.g. Carter et al., 1987). Hamlin et al. (2006) showed that a person's spatial ability influences their ability to learn and use 3D modeling software.

For those professions that require spatial skills, it could be relevant to include spatial skills training early in the professional training. This could be of particular importance for young professionals. Indeed, whereas more experienced workers can rely on a greater knowledge of relevant spatial structures and dedicated mechanisms, making the need for spatial skills less urgent (Hambrick et al., 2011), young professional could benefit from explicitly training spatial skills.

This relies on the assumption that spatial skills are trainable and goes against a wide spread belief that one is born with a given potential for spatial thinking and that this potential is fixed. In fact, skepticism about the malleability of spatial skills was persistent for many years even among researchers, with many of them arguing that training spatial skills only leads to short improvements, and only in cases where the training and measurement tasks are very similar (e.g. Sims and Mayer, 2002). However, in the most recent and comprehensive meta-analysis to date on spatial skills training, Uttal et al. (2012) found that spatial skills were moderately malleable and that training, on average, improved performance by half a standard deviation.

4.3 Spatial Skills Training

4.3.1 Results from training

One outstanding demonstration of the malleability of spatial skills and their influence on STEM domains is the work by Sorby (2009), who designed a course for developing spatial skills of first year university engineering students. One peculiarity of spatial skills is the gender difference: men consistently outperform women (e.g. Linn and Petersen, 1985; Robert and Chevrier, 2003). It is not clear whether the difference between genders is biological or environmental (McGlone and Aronson, 2006; Newcombe, 2010). Sorby's hypothesis was that women's weaker spatial skills was the reason why they had a higher dropout rate than males in engineering programs. During more than a decade, she offered a remedial course on spatial skills to first year university engineering students with weak spatial skills. She showed that the retention rate in the engineering program improved significantly for female students attending her course. On a shorter period of time, she then also trained middle school and high school students. In all three cases (middle school, high school, and university), the training consistently and positively impacted the success rate of all students, and especially that of girls.

Sorby's results indicate that both genders respond equally well to training. Uttal et al. (2012) confirmed that neither the gender, nor the age, has a significant impact on the response to training. However, they showed that initially low-performing students improved slowly at the beginning and more later on in the training, whereas students with a higher initial performance improved most early in the training (Terlecki et al., 2008).

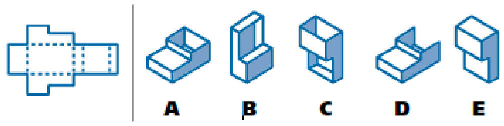
From their meta-analysis, Uttal et al. (2012) found no indication that the effect of the training was only fleeting, as has often been claimed. Their results also showed that the training could transfer across tasks, although it did not happen in all the studies. In summary, despite many controversy over the years, the most recent results indicate that the training of spatial skills can be effective, durable, and transferable.

Research is still needed to understand spatial skills more thoroughly. In particular, it is unclear how large the training of spatial skills should be in order to have an effect on other domains. Nevertheless, the fact that spatial skills can be improved and that their development may be correlated to higher performance in STEM domains has raised great interest. The idea that spatial skills should be the subject of a formal training has thus gained support in the recent years and several training courses have been developed.

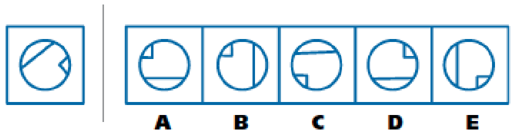
4.3.2 How to train spatial skills

Overview of training methods

In their review on spatial skills, Uttal et al. (2012) distinguished three training methods: courses, video games, and task specific training. They did not find any significant differences between these three types of training, and suggested that many different training procedures can be used to achieve durable and meaningful improvements. This was confirmed by Newcombe (2010), who, from the literature, identified three main ways of training spatial skills: dedicated spatial thinking tasks, such as those shown in Figure 4.2; the use of a language containing well-conceived symbolic representations and analogies; and the use of gestures (both for teacher and for the learner).



(a) Three-dimensional spatial visualization: If the object on the left is folded at the dotted lines, which object on the right will it form?



(b) Two-dimensional (2D) spatial visualization: Which one of the five objects on the right matches the object on the left?

Figure 4.2: Example of dedicated spatial thinking tasks (from Newcombe, 2010).

Traditionally, researchers have shown that activities requiring eye-to-hand coordination lead to development of spatial skills. Typical such activities include playing with construction toys at a young age, attending classes of drafting or mechanics, and playing 3D computer games. Newcombe (2010) suggested many ways for teachers to help students improve spatial thinking, such as highlighting spatial elements in mathematics classes, using maps and models of the world, asking children to imagine where things will go when moved, and doing jigsaw puzzles.

One activity that has repeatedly proved effective for learning is sketching (McKim, 1980; Gerson et al., 2001; Sorby, 2009). Despite its effectiveness, it is not always popular among students. For instance, the exercises developed by Sorby (2009) for her course were available both as a paper exercises, which involved sketching, and as a multimedia software, which did not involve sketching. Students stated that they preferred using the software rather than the paper workbook, although students who performed sketching improved more than those that did not.

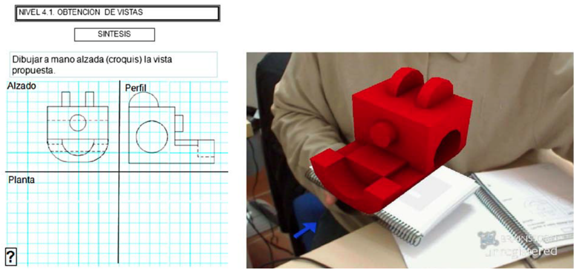
This short literature review allows us to answer the two questions asked about spatial skills in the previous chapter. First, spatial skills are indeed trainable, although there are still questions about the lasting effect of training and its transfer. Second, while drawing (sketching) can be used to train them, it is only one of the various ways in which they can be trained.

Training with technology

There has been many attempts to develop multimedia software to develop spatial skills. In 1979 already, Rankowski and Galey (1979) had shown that the use of television could increase the spatial ability of students. Recently, much research focused on multimedia tools to train spatial skills, most of them targeted at engineering students (Contero et al., 2005; Martín-Gutiérrez et al., 2010; Wang et al., 2007; Sorby, 2009). These systems usually involve a series of short “drill tasks” similar to those done on paper (such as those shown in Figure 4.2, for example).

Martín-Gutiérrez et al. (2010) designed an augmented-reality system to train spatial thinking for engineering students. The system features virtual 3D objects that can be used in small visualization tasks, as shown in Figure 4.3. The authors created a course of five sessions of increasing difficulty for a total duration of nine hours. Twenty-four first year university engineering students followed this course, and their performance, compared to that of a control group, showed that the system had a positive impact on the development of spatial abilities of its users. Furthermore, a survey revealed that the satisfaction with the system was quite high.

(a) A user manipulating the tags that represent the object.



(b) A screenshot of the application.

Figure 4.3: The augmented-reality system developed by Martín-Gutiérrez et al. (2010).

As already mentioned in Section 2.2.2 tangible interfaces have also been developed to help train spatial skills or help users with low spatial ability, with positive outcomes (Kim and Maher, 2008; Quarles et al., 2008), but those efforts remain few and far between.

4.3.3 Limitations

I see three main limitations to the current research on technology for spatial skills training. The first one is the lack of variety in the population targeted by these systems. Indeed, most of them are aimed at university engineering students. Although there are some studies reporting on the training of other populations, such as dentistry or middle school students, such attempts remain rare and, moreover, the course material used to train them is not necessarily adapted to the population's specific needs.

This leads us to the second problem: most of the training systems are not tailored for the population and context in which they are meant to be used. For example, Sorby (2009) used the same course material and multimedia system for university engineering students, high school students, and middle school students. The material of her course, as is the case for many others, consists mostly of small "drill tasks", such as finding a missing orthographic view, finding a match for a rotated object, or choosing the correct cross-section of an object among a number of candidates. These tasks are not bad per se, but they are not embedded in the larger educational task pursued by the students. This can result in a decrease of motivation from the students if they fail to see the relevance of these generic exercises for their current course of studies.

The last criticism has to do with the use of technology. Looking at the systems developed so far, it seems that most of them mostly offer a mere digital version of the material that was used for pen and paper training material. There are some exceptions, such as the work by Martín-Gutiérrez et al. (2010) and the two tangible interfaces mentioned, but those represent a minority. Spatial visualization involves making sense of and relating 2D and 3D representations, and one natural leaning would therefore be to use a real 3D object for learning .

4.4 Spatial Skills and Carpenters

The importance of spatial skills reported by both the carpentry teachers and the directors of carpentry companies is coherent with the finding that technical professions require some well-developed spatial skills. However, I did not find any data specifically on carpenters' spatial skills, and I therefore conducted a study with the goal of finding (1) whether carpenters indeed have high spatial skills, and (2) whether their spatial skills improve throughout their training. Formally, this means that we have two hypotheses:

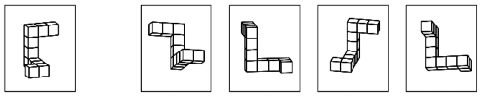
- H_A The spatial skills of carpenter apprentices are higher than those of similar individuals that do not attend a carpentry training.
- H_B The spatial skills of carpenter apprentices improve over the course of their apprenticeship.

To verify these two hypotheses, the first step was to create some testing material (see next paragraph). Answering H_A required comparison populations. Two target populations were chosen so as to represent various samples of the population: apprentices of another profession (logistics apprentices with whom we had worked in a previous research project), and more academic oriented subjects (high school students). The veracity of H_A is discussed in Section 4.4.2. As for H_B, there are two ways to test it: compare the performance of apprentices enrolled in different years of the apprenticeship, or compare the performance of the same apprentices several times over time. The results of both approaches are reported in Section 4.4.3.

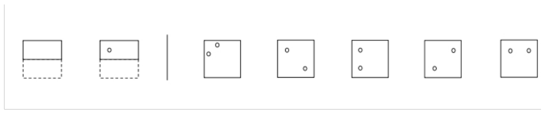
4.4.1 Test settings

Testing material

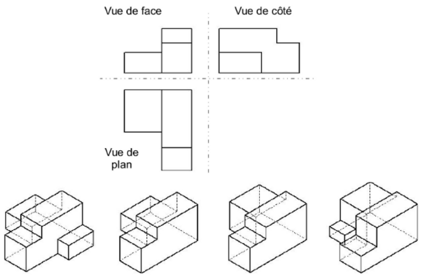
There exist several tests to measure spatial skills. The one used in this study is made of three parts: mental rotation, paper folding, and orthographic projection. Mental rotation (MR) and paper folding (PF) are two widely used tests to measure respectively mental rotation and spatial visualization abilities (Peters et al., 1995; Ekstrom et al., 1976). The mental rotation test is made of two series of 12 questions, and the paper folding test is made of two series of 10 questions. The MR and PF tests were chosen for this study because they are the main tests used in the literature. The orthographic projection part (OP) was designed specifically for this test and is composed of 6 questions that require the participant to match a 3D perspective view of an object with a 2D one. These additional items were added so that the test be closer to the practice of carpenters.



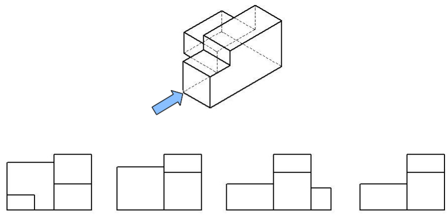
(a) Mental rotation



(b) Paper folding



(c) Orthographic projections (type A)



(d) Orthographic projections (type B)

Figure 4.4: Example questions for the three parts of the test used to assess spatial skills.

Example questions of each of the three parts are shown in Figure 4.4. For the mental rotation test, two of the four figures on the right are a vertical rotation of the left image (Figure 4.4a). The participant must identify these two figures. For the paper folding test, a square piece of paper is folded and then punched as shown on the left (Figure 4.4b). The participant must find which one of the five figures on the right matches the piece of paper, once it is unfolded. For the orthographic projections, there are two types of questions. In the first type (Figure 4.4c), the participant must tell which one of the four 3D model matches the three orthographic projections. In the second type (Figure 4.4d), a perspective representation of the model is shown and the subjects must determine which one of the four 2D drawing matches the 3D model when seen as shown by the arrow. All test questions (except for the MR questions, which are not publicly publishable) are shown in Section A.2.

Additional information collected

Besides the result of the test, additional data was collected through a small questionnaire that participants filled out before doing the test. The questionnaire asked the gender and the age of the participant, whether the participant had a prior diploma in another field, and how frequently they played video games. The test was carried out in two different regions of Switzerland, one French speaking and the other German speaking, and the region was recorded for each participant.

Participants

A total of 726 subjects (98 females) were tested. The subjects were either carpenter apprentices, logistics apprentices, or high school students. The youngest and the oldest participants were 14 and 40 years old, respectively, and the average age of all participants was 18.

Tables 4.1 and 4.2 summarize the data by year and by gender for the three separate types of curriculum. The duration of each curriculum is three years, but practical constraints allowed us to gather third year data for carpenters only. For the years 1, 2, and 3, the tests were taken at the very end of the academic year, meaning that year 1 students already had an entire year of training behind them. Subjects referred to as in year 0 were tested two weeks after the start of their first year of training. This was the most practical way to get a given population of subjects together while minimizing as much as possible the effect of their training.

	Year				
Curriculum	0	1	2	3	all
Carpenter	77	150	148	65	440
Highschool	67	38	48	0	153
Logistician	35	68	30	0	133

Table 4.1: Number of subjects by year and by curriculum.

	Gender	
Curriculum	Female	Male
Carpenter	1	439
Highschool	79	74
Logistician	18	115
All	98	628

Table 4.2: Number of subjects by gender and by activity.

Testing procedure

The test was done on paper, lasted for 35 minutes and was taken in the classroom by all students of the class at the same time. The teacher was present during the test. Participants were not allowed to communicate for the duration of the test. A timed PowerPoint presentation (shown in Section A.2.1) displayed the instructions and acted as time keeper for the test to ensure equality of treatment among the subjects in the various classes. The instructions included sample questions with answers to ensure that participants understood the task. For the parts of the tests that had two series of questions (MR and PF), a 90 second break was given to the participants between the two series.

Scoring

The final score of the test was computed as an average of the percentage score of of each part of the test, with an equal weigh of one third for each of them:

$$\text{score} = \frac{1}{3} \cdot S_{MR} + \frac{1}{3} \cdot S_{PF} + \frac{1}{3} \cdot S_{OP}$$

where S denotes the percentage score of the given part of the test.

Data pre-processing

The data needed to be processed in two ways to allow comparison across the three populations. First, given the correlation between spatial skills and gender reported in the literature girls were excluded from the data (in accord with the literature, our results showed a significant gender difference ($F[1,724]=24.1$, $p=.000$)). Second, only carpenters had data for the third year. To avoid a potential year effect, the third year carpenters were removed from the data for the population comparison. They will be reintroduced later in Section 4.4.3 when comparing the progression over time within the carpenter sample.

A note on statistical testing and reporting of the results

Throughout my thesis, I will report the results in the same graphical way whenever possible. I found a slightly modified box plot representation to be the most informative one. It shows the median, and the first and third quartiles. Additionally, it also displays the mean and the 95% confidence interval. The quartiles and median information give a visual indication on the normality of the distribution, and the mean and confidence intervals give an idea of the significance of the results. An example of such a graph is shown in Figure 4.5. For each of the results reported, unless mentioned otherwise, I 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. The statistical tests use a 5% confidence interval.

4.4.2 Population differences

Comparing populations on the overall score

A comparison of the overall score of the three populations with all males enrolled in year 0 to 2 shows that there was no significant difference between the carpenters and the high school students, but that there was a significant difference between the logisticians and the rest of the subjects ($F[2,560]=56.23$, $p=.000$), as shown in Figure 4.5.

Spatial skills are often at least partially correlated with the general school level, and the general school level of carpenter apprentices is admittedly lower than that of high school students and comparable to that of logisticians (Stalder, 2011). As can be seen in Figure 4.5, the carpenter's performance is significantly higher than the logisticians' and identical to the high school students', which suggests that the carpenters apprentices' spatial skills are well-developed.

Noteworthy is the fact that the spatial skills of carpenters were already strong at the beginning of their training. In particular, as can be seen in Figure 4.5, their performance in year 0 was similar to that of high school students and significantly higher than that of logisticians ($F[1,110]=9.6$, $p=.002$). This indicates that there is a selection effect prior to the start of the apprenticeship, i.e. people who choose to start a carpentry apprenticeship generally have well-developed spatial skills.

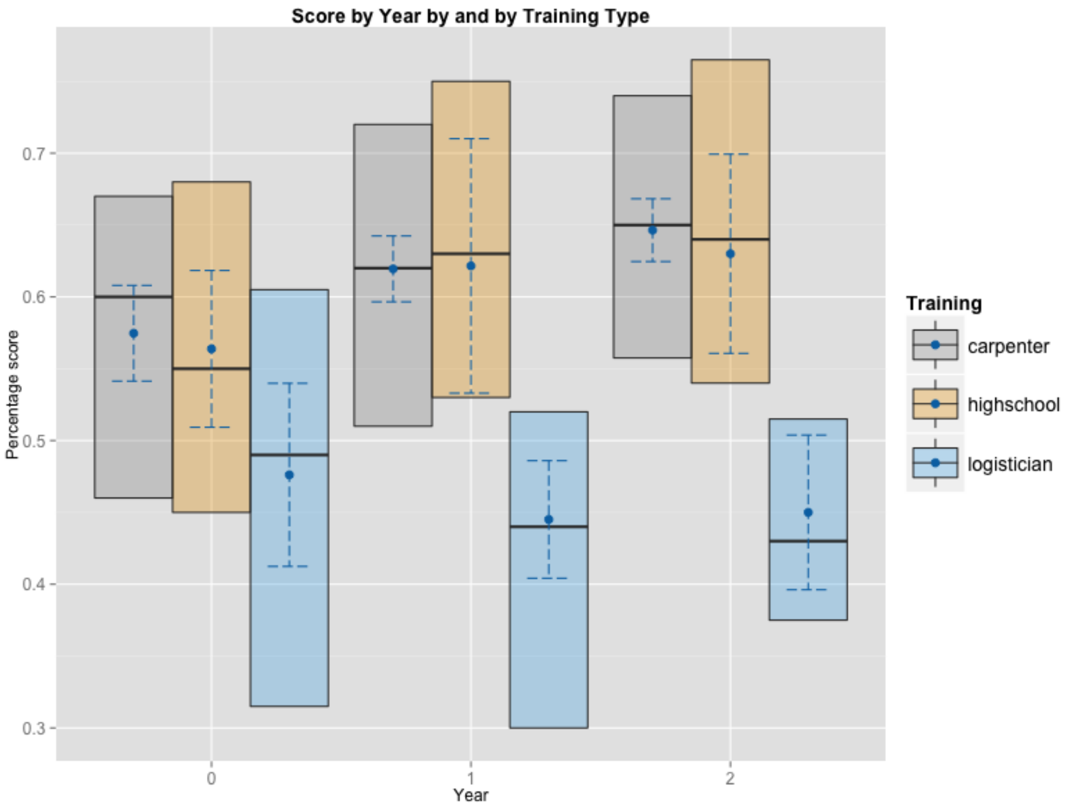


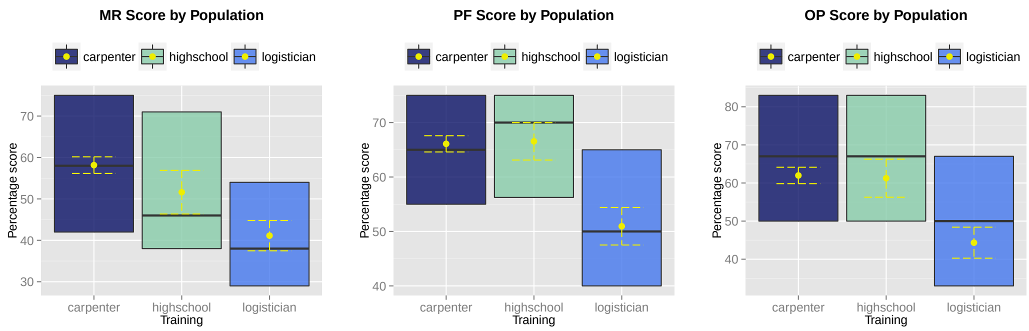
Figure 4.5: Scores for each curriculum for year 0, 1 and 2. The dot shows the mean score and the dashed lines the confidence intervals at 95%.

Comparing populations for each part of the test separately

Grouping all individuals of the same population independently of the year, but separating the results for the three parts of the test, gives an indication of the source of the difference between the different populations (Figure 4.6). The carpenters significantly outperformed the logisticians in all parts of the test. The high school students' and the carpenters' scores are similar for the paper folding and the orthographic projection parts, but the carpenters' score is significantly higher for the mental rotation part ($F[1,446]=6.42$, $p=.01$). Focusing on those two populations and splitting their score by year, the comparison shows no significant difference, except for year 0, which offers a borderline significant difference ($F[1,109]=3.93$, $p=.05$). However, because the population of high school students is rather small when split by year (34, 13, and 27, for year 0, 1, and 2 respectively), the results obtained when comparing the performance of high school students in a given year to carpenters' performance in the same year should be taken with a grain of salt.

The effect of age

There was a significant age effect: the older the participants, the lower the performance ($F[1,626]=5.51$, $p=.02$). This is surprising because there has been no report in the literature



(a) Mental rotation (b) Paper folding (c) Orthographic projections

Figure 4.6: Score by population for each part of the tests: (a) mental rotation, (b) paper folding, and (c) orthographic projections.

of spatial skills decreasing between the teen years and the age of 40 (the age of the oldest participant). A likely explanation is that older participants were distributed differently than younger ones: logisticians accounted for 46% of the population of the participants aged 25 or more, whereas they were only 17% under 25. However, even when adding the profession as a covariate, the age remains significant, indicating that there is an effect of the age itself (F[1,622]=6.5, p=.01).

Neither the region, nor video games, had a significant effect, be it on the global score or for any of the parts of the test. Having a previous training or significant work experience (3+ years) did not impact significantly the results either. However, there were also 21 carpenter apprentices who had previously been trained as furniture maker ("menuisier") and were then enrolled in another training (19 in carpentry, 1 in high school, and one in logistics). Those 21 participants globally performed significantly better than the rest of the population (F[1,626]=5.3, p=.02).

Comparison with previous research

Comparing the carpenters' scores on those two parts of the tests with previous results reported in the literature can give an indication on the carpenters' performance. There are many studies in which scores on the mental rotation test have been reported, almost all of them with undergraduate students as participants. For paper folding, there is a greater variety of tests, and therefore fewer studies with which the results of the current study can be compared. Moreover, some studies do not report the raw results required for a statistical comparison (mean and standard deviation, separated by gender). In the end, I selected three studies for mental rotation and two for paper folding.

For mental rotation, Peters et al. (1995) tested 237 males enrolled in a Bachelor program at the University of Guelph, in Canada, 102 in an Arts program, and 135 in a Science program. Their average age was 21.3. Students in the Arts program scored 50.4% (standard deviation (SD): 20.0%) on average, and those in the Science program got a mean score of 61.7% (SD: 20.0%). In comparison, the 439 male carpenter apprentices tested in the current study got a mean score of 58.1% (SD: 19.8%). According to a t-test, the carpenters' performance is significantly higher than the Arts students' (t[539]=-3.53, p=.000), but not significantly different than the Science students' (t[572]=1.84, p=.66). Another study took place in St. Francis Xavier University, in Canada (Voyer and Saunders, 2004), and involved 139 male introductory psychology students. The mean age of the participants was 19.8, and their average score, 60.7% (SD: 18.9%), which is not significantly different from the carpenters performance (t[576]=1.36, p=.17). A third study (Gouchie and Kimura, 1991) tested 42 volunteer males (mean age 21.0), mostly undergraduates at the University of Western Ontario, still in Canada. They scored 45.8% (SD: 39.0%), a score significantly lower than the carpenters' (t[479]=-3.44, p=.001).

The same study also tested the participants on the paper folding test (although they used only one of the two parts of the test used in the current study, i.e. 10 questions instead of 20). Their performance (mean=62.9%, SD=20.8%) was lower than the carpenters' (mean=66.4%, SD=14.6%), although not significantly (t[479]=-1.42, p=.16). Still at the University of West Ontario, Kimura (1994) measured the performance of 24 undergraduate male students (no mention of age) to 66% (SD: 24.0%). The assessment was again done using only 10 questions of the paper folding test. The performance was not significantly different from the carpenters' performance (t[545]=-.22, p=.83).

These comparisons with previous research show that the carpenter apprentices' performance on either the mental rotation or the paper folding test was as high, and even higher in some cases, than first year university students' performance. This result goes along the same line as the findings from the comparisons with high school students and logistics apprentices, and answers the confirms the first hypothesis made at the beginning of this section (H_A): carpenter apprentices have well-developed spatial skills. We now turn to the second hypothesis (H_B), which can be reformulated as a question: are these skills high by nature, or do they develop over the course of the apprenticeship?

4.4.3 Improvement over time

One way to look at the progression over time is to compare the performance of the apprentices in year 0 with the performance of the apprentices in year 3. However, this makes the assumption that, on average, each yearly batch of apprentices started their apprenticeship with the same spatial skills level. One way to get rid of this assumption is to measure the actual skill improvement over time by testing the same apprentices twice: once at the beginning of their apprenticeship, once towards the end of it. This has the disadvantage of introducing a retest effect. I will present the results of both approaches here.

Different students on different years

The 439 male carpenters were spread over the 4 years (see Table 4.1). Figure 4.7 shows that carpenters tend to improve slightly over the course of their apprenticeship. According to a linear model fitted by regression (shown in the Figure as a line), the average score improvement for each year of training was 2.2%, which is a significant improvement (F[3,437]=4.65, p=.003).

Although the 2.2% average improvement per year is statistically significant, comparing the results of the years two by two with a pairwise t-test tells a slightly different story. Indeed, while the score for year 0 is significantly different from that of the three other years, the year-to-year comparison of the performance in year 1, 2, and 3 shows no statistical differences.

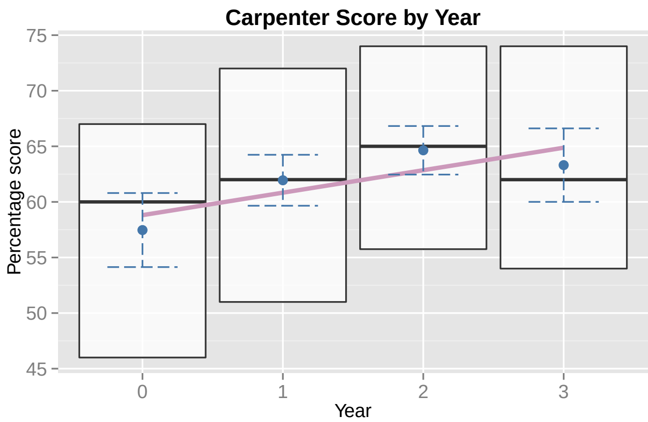


Figure 4.7: The overall score for carpenter participants split by year. The purple line represents the linear regression model.

This is a contrasted result: on the one hand there is a significant improvement over the four years tested, but on the other hand the improvement seems to come mostly from the lower result of participants tested in year 0. This could indicate that spatial skills improve strongly during the first year of the apprenticeship, and then only marginally. It could also be due to the drastic selection that occurs in the first year of the carpenters' training, where about one third of the apprentices fail. The significant improvement over time found for carpenters does not appear for the two other populations ($F[1,72]=2.62$, $p=0.11$ for high school students, $F[1,113]=0.44$, $p=0.50$ for logisticians), although the non significance of the statistical result might also be due to the smaller sample size of those two groups.

One odd result is the slightly lower (though not significantly, $t[211]=.67$, $p=.50$) performance of year 3 carpenter apprentices compared to that of year 2 carpenter apprentices. Although it is possible that the spatial skills level could stagnate after two years of training, it is difficult to justify why it would go down. One explanation could come from the conditions in which the test was taken: for year 1 and 2, the test was taken somewhere during the last two weeks of the school year. For year 3, because of the final exams, participants took the test on the very last day they had to come to school. They were asked to do their best, but it might have been that their level of application was negatively influenced by the very close perspective of vacation.

Same students two years apart

A second way to measure whether the spatial skills level of carpenter apprentices improves during their apprenticeship is to test them twice during their apprenticeship: once at the beginning, and once towards the end of it. In the spring of 2012, two years after the first round of tests, a subset of the same carpenter apprentices – those who were still doing their apprenticeship – were administered the same test a second time.

Data. In total, 124 male apprentices took the test twice. In order to be able to make cleaner comparisons, 9 apprentices who had failed a year and 4 apprentices who were on a fast track 2-year apprenticeship were removed from the data set. The final sample size was 111, with 41 apprentices in year 0 in 2010 and in year 2 in 2012, and 70 apprentices in year 1 in 2010 and year 3 in the 2012. In the rest of this section, I will refer to those two groups by the year in which they were enrolled in in 2010, i.e. year 0 and year 1.

Results. Table 4.3 shows the 2010 and 2012 performance for three different groups: all students together, only year 0 students, and only year 1 students. For each of these groups, the detail for each part of the test is also shown. Additionally to the percentage scores of 2010 and 2012, the table shows the relative learning gain (RLG) between 2010 and 2012, and the statistical significance of the difference between the two scores, computed with a two-sample paired t-test. The RLG was computed as shown in Equation 4.1, where *pre* and *post* are the percentage score for the test in 2010 and 2012, respectively.

$$RLG = \begin{cases} 100 \times \left(\frac{post-pre}{100-pre}\right), & \text{if } (post-pre) > 0 \\ \\ 100 \times \left(\frac{post-pre}{pre}\right), & \text{if } (post-pre) \leq 0 \end{cases} \tag{4.1}$$

There are other ways to report learning gain (score difference, percentage increase, etc.), but the RLG has the advantage of not penalizing students who scored well on the pre-test, because it measures the improvement achieved on the possible improvement from the pre-test score.

Improvement over the two years. There was a significant and positive global RLG between the 2010 and the 2012 scores (26.4%). The mean percentage score went from 62.8% to 72.6% (F[1,110]=95.78, p=.000). Even when looking at the parts of the test separately, each of them displays a strong and statistically significant improvement when measured on all students. The improvement was stronger than for the results reported in Figure 4.7: it went from a 2.2% average yearly improvement to a 4.9% average yearly improvement. This could be the result of the selection effect incurred by the removal of students redoing a year from the second set of data. It is also due to the lower score of students measured in the end of year 3 and mentioned above. Indeed, when taking into account only the years 0, 1, and 2 from the results in Figure 4.7, the yearly improvement goes up from 2.2% to 3.6% (57.5% in year 0 to 64.6% in year 2).

Improvement for each part of the test. When further splitting the students per year, all but one improvement still shows statistical significance, with the only non significant result being the orthographic projection part for the year 0 group. In parallel, on the same part of the

Group	Part	2010	2012	RLG	Significance
year 0	MR	60.0	70.1	29.3 %	t[40]=-3.32, p=.002
	PF	61.3	73.9	30.7 %	t[40]=-3.26, p=.000
	OP	63.0	67.9	20.6 %	t[40]=-1.27, p=.21
	all	60.9	71.4	27.4 %	t[40]=-6.11, p=.000
year 1	MR	60.6	71.8	27.5 %	t[69]=-5.53, p=.000
	PF	68.8	76.4	27.4 %	t[69]=-5.31, p=.000
	OP	61.0	69.8	25.6 %	t[69]=-3.54, p=.001
	all	63.9	73.4	25.8 %	t[69]=-7.60, p=.000
year 0 and 1	MR	60.4	71.2	28.2 %	t[110]=-6.37, p=.000
	PF	66.0	75.5	28.7 %	t[110]=-8.19, p=.000
	OP	61.7	69.1	23.8 %	t[110]=-3.49, p=.001
	all	62.8	72.6	26.4 %	t[110]=-9.79, p=.000

Table 4.3: Percentage scores for 2010 and 2012, RLG and statistical significance of the difference between the two years, separated by groups of students and by the parts of the test.

test, year 1 apprentices did improve significantly. This suggests that the improvement on the carpentry specific part is linked to the second and third year of the apprenticeship, rather than the first year of it. It is surprising, because what was tested in the carpentry part (orthographic projections) is already taught in the first year. There are several explanations for this. One is that apprentices need some time to master orthographic projections, and longer exposure to them, both at work and at school, increases their performance. The fact that the performance on the other parts of the test improved could reflect the higher degree of specialized knowledge required for the orthographic projections. By comparison, the MR and PF tasks are purely spatial reasoning tasks and do not require extra knowledge. Another reason could be that the test did not capture improvement as well as the MR and PF parts, which had more questions and which benefited from more effort and meticulousness from psychologists in the design of the questions. Finally, the similarity between the OP part and the school exercises might have reduce apprentices' motivation.

Differences between years. The two groups of students (those in year 0 in 2010, and those in year 1 in 2010) performed similarly. Their performance at both the 2010 and 2012 tests was similar (t[89.3]=1.11, p=.26 and t[77.4]=-.83, p=.41). Their relative improvement from 2010 to 2012 was also similar (27.4% versus 25.8% in favor of year 0). Comparing pairwise each part of the test for both years, the only comparison that offers a significant difference between the two groups is the 2010 paper folding test (t[73]=-2.5, p=.02, average score of 61.3% versus 68.8% in favor of the year 1 group). The score in 2012 is still a bit higher for year 1 apprentices, but not significantly so (t[87.9]=-1.02, p=.31, average scores of 73.9% versus 76.4%), hinting that year 0 students caught up with year 1 students during those 2 years. This could indicate that years 1 and 2 of the apprenticeship are more effective for paper folding activities than years 2 and 3, or that there is an asymptotic limit to the paper folding performance and that this limit is reached within one year of training.

4.4.4 Limitations of the current study

As already mentioned, the measure of the progression done by comparing different students enrolled in different years of the apprenticeship makes the assumption that the four batches of students started their apprenticeship with similar spatial skills. This assumption, although reasonable, cannot be verified. When measuring the progression as the score of the same students two years apart, there is a non measurable retest effect. Arguably, the retest effect two years apart is weak, but ideally, there should have been a control group to measure this retest effect.

Everything was done to ensure that the test conditions were the same for all subjects, especially time-wise, but practical constraints made it impossible to have the exact same conditions for all subjects. Subjects were tested on different days of the week, at different times, and with different surrounding school environments. As mentioned, this was especially a problem for the year 3 carpenter apprentices, who were tested on their last day of school and were not as diligent as the other subjects.

In the design of the test, a set of questions specific to the carpenters' curriculum (OP) was added to two widely used tests (MR and PF). In order to keep the total duration of the test under the duration of a lesson, the number of questions in the OP part was limited to six. Ideally, to be able to pinpoint exactly what difficulties participants had with orthographic projections, the OP part should have comprised about 20 questions. However, this would have pushed the test over the duration of a lesson, which would in turn have made the administration of the test difficult. The information brought by the OP part is therefore limited.

Finally, the way in which the score was computed could be discussed. For instance, no difference was made between a wrong answer and a missing answer, as it is sometimes the case in the literature. Also, the final score was computed as a weighted average of each of the three parts, with each of the three weights being equal. Other ways to compute the final score would have been to use weights proportional to the number of questions or to the time allotted to complete a specific part of the test. We chose to give equal weights to each part of the test so that the various skills tested by each part of the test would be taken into account.

Despite its limitations, this study fulfilled its goals by providing useful information on carpenters' spatial skills. In particular, it makes it possible to answer the two questions that were raised at the beginning of this section: whether carpenters have better spatial skills than other similar populations, and whether their spatial skills improve over the course of their apprenticeship.

4.4.5 Conclusions from the study

Carpenters' spatial skills are high...

The first hypothesis (H_A) was that carpenters had higher spatial skills than similar populations. This hypothesis was verified in two ways: through a first-hand comparison with high school students and logistics apprentices, and through a comparison with previously reported results of college students from Canada. The results further showed that carpenter apprentices' spatial skills were already high at the beginning of their apprenticeship, indicating that there is a selection bias in the way carpenters are chosen or choose themselves their profession.

... and yet they improve

The second hypothesis (H_B) was that carpenters' spatial skills improved over the course of their apprenticeship. The measurements were done in two ways, but showed similar results: carpenter apprentices' spatial skills performance increases over the years of the apprenticeship, and especially during the first year of their apprenticeship. We can therefore accept H_B, with the reservation that part of this improvement may be due to a selection bias (weaker students leave).

The verification of the two hypotheses complements the results of the contextual inquiry reported in the previous chapter, which established that spatial skills are key for carpenters. The importance of spatial skills is evident in the selection process – whether it is a self-selection or a company-based selection – as well as in the training received by the apprentices.

4.5 Conclusions

This chapter first introduced spatial skills in general, and then spatial skills in the context of carpenters. Spatial skills have been a research topic for many decades, and the interest towards them has increased recently because of their potential to improve performance in STEM domains. Several studies showed that spatial skills are trainable through various activities such as video games, sketching, and dedicated drill exercises. Although large by the number of studies, the existing body of research has some limitations. It lacks variety in the populations tested. The means developed for training are most of the time generic and do not take into account the context for which spatial skills are taught. Finally, the computerized ways of training spatial skills are often a computer implementation of paper exercises, and they do not take advantage of the specific features of the computer.

Two main facts were established concerning carpenters and spatial skills. First, carpenters do have superior spatial skills than would be expected given their general school level. In particular, their spatial skills were found to be similar to those of high school students and first year university students enrolled in a Science Bachelor program, and superior to those of logistics apprentices and first year university students enrolled in an Arts Bachelor program. Second, two different comparisons showed that carpenters' spatial skills improve over the course of their apprenticeship. These findings confirm that spatial skills are trainable and suggest that the high spatial skills level of carpenter apprentices is due to a selection bias as well as to the training that they receive during their apprenticeship.

The situation described so far suggests that there is an opportunity for a tangible interface to train spatial skills for carpenters. Carpenters need to develop their spatial skills. The widely accepted assumption today is that they currently do it by performing descriptive geometry pen and paper exercises. These exercises are useful and should not be removed, but they also have some limitations, and complementing them with a computer interface could help overcome these limitations.

Most of the computer programs created so far for the training of spatial skills have reproduced the paper exercises in a digital format. Few have tried to use tangible interfaces, although it seems that the inherent spatial component of TUIs could fit well the training of spatial skills. However sensible this seems, it remains to be explored whether, and if so how and why, a TUI can be helpful to train spatial skills. This will be one of the questions studied in this thesis (Chapter 6).

Another challenge that emerged from the literature review is the lack of contextualization of the existing training material. Spatial skills were and are often taught in extra-curriculum remedial classes, and the exercises used in those classes are usually quite generic. While this can lead to some results, it is also important – and especially so in a vocational training – that the learning material be tailored to the context in which the learned skills will be used. Therefore, a second challenge will be to create an interface that can be used in a way that makes sense in the context of carpenters (Chapters 6 and 7).

5 Research Questions and Method

5.1 Research Goals

As explained in the previous chapter, spatial skills can play a key role in many school subjects as well as for many professional domains. The recent interest for spatial skills has been accompanied by the development of remedial courses to train spatial skills. Few of these courses use computer technology as a support, and when they do, they often fail to make use of its potentialities. In parallel, as explained in Chapter 2, there has been enthusiasm for using TUIs in educational contexts, although empirical evidence of the impact of TUIs on learning is still lacking. TUIs are a promising means to train spatial skills, because the physicality of TUIs could help make the link between 2D and 3D objects.

The ensuing research question is: "Do TUIs enhance the acquisition of spatial skills?". While answering this question is quite appealing, it is too broad and vague to be answered empirically. It is indeed rare that one technology be helpful to develop a specific skill in all cases. Rather, I propose to split it into the following questions, which focus on *how* TUIs could be used, if at all, to train spatial skills:

Q_F Which features of a TUI support the acquisition of spatial skills?

Q_{PS} What classroom pedagogical scenarios do TUIs support for the training of spatial skills?

Of course, these questions are still quite generic. Before one can answer them, many other questions need to be clarified, such as: What TUI are we talking about? What will the tangible elements be? What will the learning activities be? Who is the target population? What baseline should the learning performance with the TUI be compared to? The broader context of this work is that of carpenter apprentices, and the specific context will be described in each study. The research method used is design-based research, which is presented in the next section.

5.2 Design-Based Research

In educational technology research, there is a tension between the development of pedagogical theories and their applications and assessment in real learning environments. A learning environment is often complex, and no one learning environment is identical to another. Yet, one of the most spread ways of investigating educational technology is to conduct laboratory experiments, in which everything can be controlled and one or two parameters are varied, and the output of this variation studied. The problem of laboratory studies in the field of educational technology is that their conclusions are often not verified in the ecologically valid settings.

Design-based research (DBR) was developed to close the credibility gap between laboratory studies and the real, concrete, learning environments. The idea was to conduct complex intervention studies (also called "design experiments") in order to inform how the design of a learning environment can affect the learning and teaching (Brown, 1992; Wang and Hannafin, 2005). Wang and Hannafin (2005) give the following definition of DBR:

> [DBR is] a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories.

The same authors list five basic characteristics of DBR. It is *pragmatic*, in the sense that it aims at solving real-world problems through interventions. It is *grounded* in existing research and theory and in the real-world context. It is *interactive, iterative* and *flexible*: the design involves the participants and is conducted through repetitive iterations of analysis, design, and implementation. It is *integrative* of various research methods to maximize the credibility of ongoing research. Finally, a careful documentation of the process ensures that the design principles found and guidelines to apply them are *contextual*.

Whereas the goal of a traditional empirical research iteration (hypothesis, experiments, and theory) is typically to refine the original research hypotheses, the goal of a design-based research iteration is the refinement of the problem definition, the prototype, and the method (Reeves, 2000). The outcomes of DBR are theories (design principles, design frameworks, design methodologies) and practical interventions, which can include, as is the case in this thesis, the creation of a working prototype. To reach this outcome, researchers use multiple means such as surveys, observations, interviews, and document analysis.

DBR has some drawbacks. For one thing, it is time consuming, both from the researcher perspective and from the school perspective. School time is precious, especially in the dual training system where apprentices go to school only once a week, and it can therefore be problematic to do interventions in schools. Another criticism that has been formulated against DBR is its lack of strong theoretical foundation. The use of several research methods creates the risk of misusing all of the methods. Finally, the adaptation to a specific context, which is the cornerstone of DBR, can make it difficult to generalize results to different contexts.

Despite these drawbacks, DBR is an appropriate methodology in the context of this project for several reasons. The participation of teachers is key, because carpentry is a complex field, and it is hard for a researcher to build an expertise allowing him to be solely in charge of the solution design. The multiple iterations of analysis, design, and implementation, allow the researcher to be adaptive and reactive as he gains more understanding of the context and subtleties of the field. The pragmatism of DBR matches with the general culture of carpentry, in which the focus is on building concrete things. The fact that the time at school is precious, because of its scarcity, is slightly problematic, for it means that the interventions need to be short and punctual. However, the benefits – a theoretical approach tied to a real-world context as well as a working prototype – exceed the drawbacks.

5.3 TapaCarp

TapaCarp is the system that stemmed out of the DBR approach. It is a result of the work done during my PhD, in many iterations, but I present it (and the design process that led to it) in this chapter because it is needed to refine the research questions. As pointed out in the "Acknowledgements" Section at the beginning of the Thesis, several people contributed to the development effort of TapaCarp.

5.3.1 Technical setup

TapaCarp runs on the Tinkerlamp, a top-down camera-projector tabletop system shown in Figure 5.1. The Tinkerlamp used in this project is an evolution of the original one developed by Zufferey et al. (2009) to teach logistics (see Section 2.2.2). Besides the hardware that was upgraded in the new version, there are two main differences with the original version. First, a computer is now embedded in the Tinkerlamp, so that there is no need anymore for an external laptop computer. Second, instead of pointing down to the tabletop, the camera and projector face up towards a mirror that redirects the light ray on the tabletop. This has the advantage of extending the projection area (from 30 × 50 cm to 50 × 70 cm). The zone that is lit-up by the projector is called the "active" zone or the workspace, and the zone that is not lit-up, the "inactive" zone. Objects are not detected when in the inactive zone, either because they are outside of the camera field of view or because the darkness does not allow their tag to be detected by the camera.

The projector has a resolution of 1280 × 768 pixels mapped to the 50 × 70 centimeters area on the tabletop. The camera has a resolution of 1280 × 960 pixels and sees an area slightly larger than the projection zone. The system tracks objects equipped with fiducial markers, and provides visual feedback through the projector. The fiducial markers (Bonnard et al., 2013) are similar to ARTag markers (Fiala, 2005) and allow to accurately track elements placed on the tabletop surface. Only objects with a tag can be detected and interpreted by the system. TapaCarp was built using `C++` and OpenGL and the code is deployed on an Ubuntu Linux operating system. TapaCarp can be used by teams of 2 or 3 students, or by a single student,

Figure 5.1: The Tinkerlamp.

depending on the activities and the circumstances.

5.3.2 Design constraints and choices

As mentioned in Section 2.2.2, there are both advantages and drawbacks to top-down systems, compared to rear ones. Originally, the Tinkerlamp was designed as a top-down system because the ability to project on top of objects was important. This comes at the cost of higher sensitivity to light conditions, and the tag tracking algorithms had to be made robust enough so that light problems are minimal. Tracking from the top is also sensitive to objects or body parts occluding markers, but in general users understand quickly that they should not occlude the markers from the camera. Another limitation of the top-down approach is that detecting finger touch is not accurate enough, therefore not allowing touch interaction.

Using fiducial markers ensures reliability. TapaCarp is meant to be used in classrooms that are lit up normally, and it is not possible to require that the classroom be darkened, or that the room have a constant luminosity. Although detecting objects without any kinds of markers would be nicer, fiducial markers offer a greater reliability to changing light conditions.

Another important design choice was to not include a keyboard as an input means, but to keep the mouse. A keyboard is the main input means for most computer interfaces, but is rarely included in TUIs, which favor more natural input means. Including the mouse can seem more surprising. The mouse was actually not included in the original design, but one of the activities (see Section 5.3.5) required a means to do accurate selection, and after assessing various ways of doing this, the mouse proved the fastest and most accurate selection tool.

One final important design consideration was time. As already mentioned, the time that apprentices spend at school is limited. Although the goal was for TapaCarp to be used substantially in the classroom, it also had to be designed so as to consume as little of this time as possible. The activities were thus in general kept short and the interface was minimized to reduce the time needed by apprentices to get started.

5.3.3 Design process and basic components of the interface

TapaCarp was designed in several iterations with the help of carpentry teachers and their students. The teachers' main complaint about their students was that they did not make the link between the 2D representations of an object (its orthographic projections) and its 3D shape. Not making this link leads students to draw plans that are wrong, i.e. plans that can correspond to unbuildable roof structures. According to the teachers, students tend to follow descriptive geometry "recipes" to draw their plans, without understanding the link between the orthographic projections that they are drawing and the final object that they mean to represent.

The teachers' complaint can be seen as paradoxical. On the one hand, the drawing techniques that they teach to their students are very methodical and do not leave room for freedom of execution. Much of the emphasis is on the accuracy of the drawing, and simplified drawing algorithms are given to students to help them bootstrap their ability in descriptive geometry. On the other hand, they expect the students to have a deep understanding of the geometrical meaning of what they are doing. Teachers are not to blame for this paradox, for what they are asked to do is precisely to teach these drawing techniques. As for students, it is only understandable that they try to cling on to a given set of rules in order not to lose ground while learning as hard a subject as descriptive geometry. In any case, the current situation is not satisfactory, and TapaCarp could be a step forward towards improving it by providing an intuitive, yet drawing-related, way of training spatial skills. The approach is also more integrative in that it makes a clearer link between the 2D representations of the object, its 3D representation, and the object itself.

The development of TapaCarp can be decomposed in two main phases. The output of this first phase, which included several design iterations, was a basic system with an interface composed of cards, blocks, and a mouse, and offering several short activities on orthographic projections. In a second phase, the interface of TapaCarp evolved as the demands for different activities surfaced. Although the work was done in an iterative fashion over several years, the full system is presented in this section for the sake of clarity.

Blocks

The first component is a set of solid blocks made out of wood or cardboard (Figure 5.3). The blocks are equipped with fiducial markers, which allows Tapacarp to track their position and orientation accurately. Knowing the topology of a block, TapaCarp displays its orthographic projections and a perspective view of the object (Figure 5.2). The gray square on the bottom

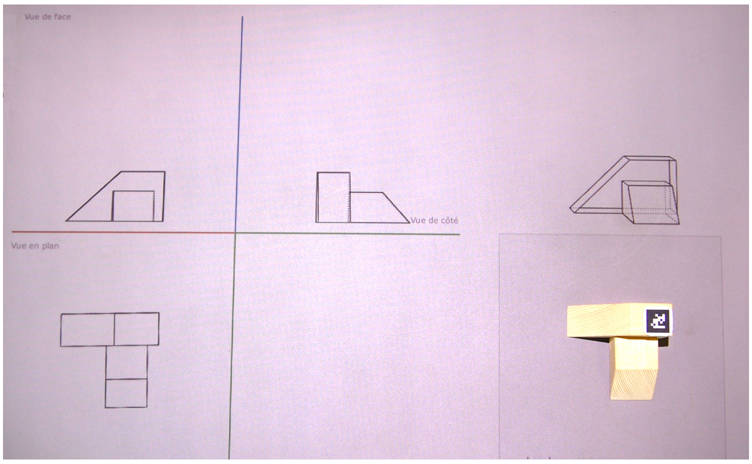
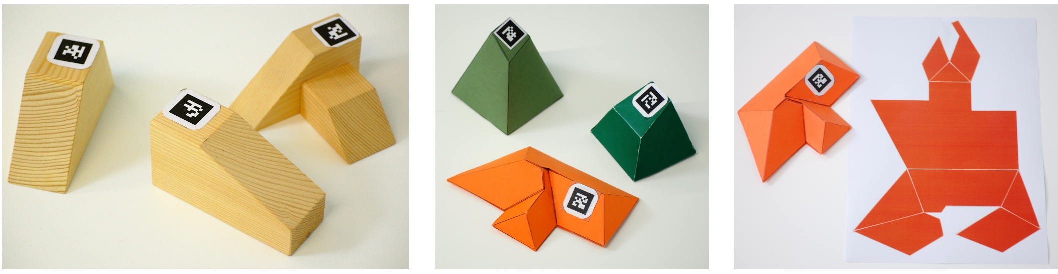


Figure 5.2: The layout of Tapacarp: the wooden block and its 3D representation. On the left, the three orthographic projections of the block.



(a) Wooden blocks. (b) Cardboard blocks. (c) The printed development of a cardboard block, and the block.

Figure 5.3: The blocks.

right is the “block zone”, a 19 × 19 cm square in which blocks should be placed. The three orthographic projections are shown on the left, and a perspective view on the top right. All representations are dynamically linked, allowing the users to explore the relationship between them by moving the blocks and seeing the effect of the movements on each view. The blocks play a key role in the TapaCarp interface: they serve both as the main input to the system and as an external representation of the projected orthographic projections.

At first, all the blocks were made out of wood and had an abstract shape, i.e. they did not look like a building. Making the blocks out of wood was beneficial from a cultural point of view, because wood is the main material of carpenters, but it had two drawbacks. First, it requires raw material and the appropriate tools to create them. Second, non-convex shapes are hard to create by subtraction, as is the case when cutting wood with a saw. For these two reasons, we introduced blocks made out of cardboard, that can be simply printed on an A4 page, and then folded and glued into a solid shape (see Figure 5.3c).

Cards

While blocks are the core manipulation handles of the interface, they do not allow the users to trigger specific actions such as “launch activity 1”, “show feedback”, or “check my solution”. We therefore introduced the second component of the interface: a set of paper cards (Figure 5.4). Cards are used to issue actions or to modify the values of some options (e.g. hide or display axes). Each card only has one function and the number of functions provided by the system is therefore proportional to the number of cards available. This makes it easy to adapt the number of features to the students’ level of expertise with the system or the learning domain, by giving to the students the appropriate set of cards. Cards were also chosen for practical reasons, such as their ease of distribution, storage, and sharing, all of which go in the direction of reducing the global orchestration load faced by the teacher. They are easy to manipulate and share between several users. Besides the icons, the other thing that distinguishes the cards is their background color (see Figure 5.4b). Cards that launch an activity are red, cards that modify the display are turquoise, animation cards are purple, etc. This is a minor detail, but it helps find cards much faster, and can contribute to decrease the overall orchestration load.

(a) A regular and a flippable cards.

(b) The background color distinguishes the families of cards.

(c) Gray rectangle in which the card must be put.

Figure 5.4: The cards.

During the first usability studies, a problem appeared with the cards. This problem may seem like a detail, but I describe it into some more detail as it underlines the iterative design approach that was followed. The problem was that, sometimes, users forgot that they had activated a card and left it in the workspace even after they were done using it. As a consequence, the action linked to the card was triggered continuously. This can lead to undesired behavior, such as the solution being checked permanently if the “check solution” card is left active.

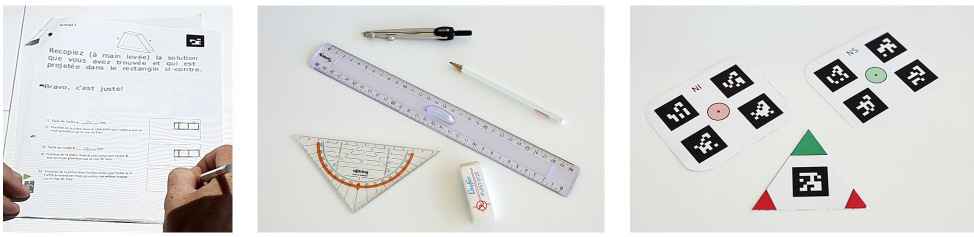
An obvious first step to avoid continuous triggering is to introduce a minimal required duration between two occurrences of the same event. The event linked to the card can only be fired after a duration d (say 4 seconds) has elapsed since the same event was last triggered. However, this does not solve everything: if the user leaves the card for longer than d, the event will still fire again unexpectedly.

“Flippable” cards solve that. These cards, shown in Figure 5.4a, are smaller than the regular cards and are two-sided. On one side, an icon and a text label indicate what the function of the card is. This side is easy to recognize for the human eye, but the machine does not detect it. The tag that triggers the action is printed on the other side of the card. The intended usage of the card is that it is by default placed with the icon side up. To use it, the user grabs it, flips it, places it briefly in the workspace, flips it back, and puts it down (with the icon face up). In that case, the card works like a click on an action button in a GUI. However, two problems led us to abandon flippable cards. First, some users had difficulties grabbing the card to turn it, instead of just moving it on the table as with regular cards. Second, users often put the card down with the tag facing up, defeating the intended barrier to continuous event triggering and making it harder to find the card later on, because the icon is hidden.

A new solution that does not involve any flipping of the card and solves the problem of event triggering is to ask the user to place the card at a given location. When a user puts a card in the active zone of the system, a gray square of the size of the card appears (Figure 5.4c). This square represents the location where the card must be placed in order to be activated. As long as the card is not put in this rectangle, it has no effect and the rectangle stays in the same location. If the card is put in the rectangle, the rectangle disappears for 5 seconds, and then reappears at another location. A minor drawback of this solution is that sometimes another object sits where the card needs to be placed and needs to be moved before the card can be placed there. But this counterbalanced by ability to simply be able to slide the card to the right location (as opposed to grab it) and the benefit of events not firing continuously.

Mouse

A third component of TapaCarp is a standard computer mouse that was used to interact with the digital models. There is no reason why modern interfaces should avoid using traditional computer input devices when these prove the most effective. We ran various small-scale usability tests to compare the mouse with tangible "selectors" (see Figure 5.5c), which showed that the mouse was the fastest and most accurate tool to select and move thin objects, such as lines. In some cases, however, for example when selections in a given activity are rare, it may sill make sense to use paper tools, such as the ones shown in Figure 5.5c, to manipulate digital objects.



(a) Activity booklet. (b) Drawing tools. (c) Paper tools.

Figure 5.5: Booklet and tools.

Booklet activity and drawing tools

Based on the blocks and cards interface, we developed a series of learning activities to help apprentices learn to link the 2D and 3D representations of an object. Some activities are described below in Section 5.3.5. The activities could be completed in a short amount of time (less than 5 minutes). Their level of difficulty was easily adaptable, for example by selecting simple versus complex blocks. However, and although they participated in the design of these activities, the teachers thought they were too far from the school curriculum, because they did not include drawing (see Chapter 7 for more detail on that).

We therefore added a new interface component, a paper activity booklet (Figure 5.5a), to include new activities in which students could perform the act of drawing on paper. They drew with their regular drawing tools (Figure 5.5b), which further satisfied the teachers. The drawing tools are not, properly speaking, part of the interface – they are not tracked by the camera and are hence not an input device – but fiducial markers could be placed on them too, for instance to check if the center of a protractor is accurately placed in the center of an angle to be measured. The booklet is composed of A4 pages, each page being a separate activity and equipped with a unique fiducial marker so that the system could augment it with instructions and feedback. The activity booklet used in one study is shown in Section A.5.4, and one activity using the drawing tools and the activity booklet is presented in Video 1(more details about the videos can be found in Section A.1).

5.3.4 Basic features

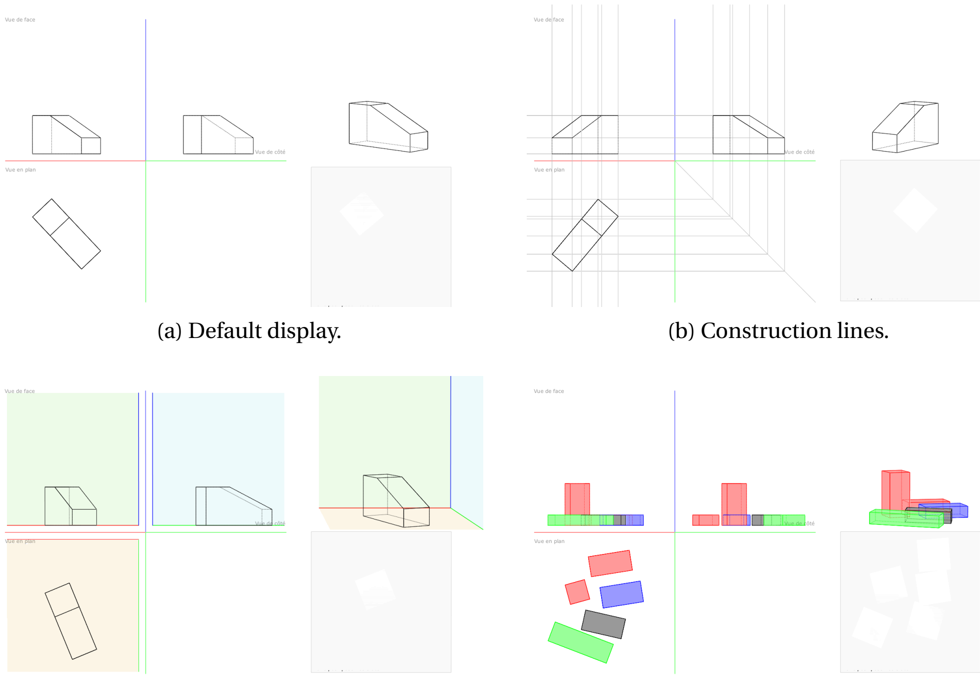
Many features were developed, tested, and modified during the iterative development of TapaCarp. The features described below are not an exhaustive list of everything that was developed, but rather a selection of the features that constitute the kernel of the TapaCarp interface and that can show the thought process that accompanied the design of Tapacarp.

Display options

The default display is shown in Figure 5.6a. It can be modified using cards which can for instance add construction lines, display axes and color background, or change the rendering of the blocks from wireframe to solid (transparent) representation, as shown in Figure 5.6. These three features and a fourth one ('freeze' feature) are presented below in more detail.

Construction lines serve to link the three orthographic projections (Figure 5.6b). Given a point p of an the object, this point can be seen in each of the three views, but the information given by each of the views is different. The top view gives the x and y coordinates of p, the front view the x and z coordinates, and the side view the y and z coordinates. The construction lines help report the coordinate from one view to another. This is useful when one of the three views is missing and must be reconstructed from the two others, which is one of the typical exercises given to carpenter apprentices.

A second display feature is to display axes and add color to the background of each of the three



(a) Default display.

(b) Construction lines.

(c) Background colors.

(d) Rendering.

Figure 5.6: Screenshots of the TapaCarp display, showing (a) the default view and three ways in which the display can be augmented: (b) construction lines, (c) colored background on the views, and (d) changing the rendering of the object from wireframe to solid.

orthographic projections. The same colors and axes are also shown in the 3D representation of the object (Figure 5.6c). This feature helps beginners remember what each of the three views represent. For example, when first dealing with orthographic projections, apprentices tend to mix the position of the side and front views, or to forget whether the side view represents the object when it is looked at from the right or the left. Matching the axes of the 3D view and the orthographic projections also helps understand the relationship between the views.

Changing the rendering from solid (with transparency) to wireframe, and vice versa, can help get a better representation of the position of the objects in some cases. In general, it makes more sense for apprentices when objects are rendered as wireframe, because this is the representation they use when drawing. However, when many objects are present in the scene, a first step is to distinguish their position relatively to each other. In this case, switching to the solid representation can help, because the objects that are in the background will blocked by the front objects (results in Chapter 8 will confirm this).

A fourth display feature is one that allows to freeze the display. In other words, once the "freeze card" is activated, the movements of the blocks are not taken into account anymore by the system. This feature was added to allow users to move blocks freely without changing the orthographic views. This is useful when users are doing other things on the tabletop that risk moving the blocks inadvertently (such as drawing, for example), or when they need to take the block and move it for some reason.

Action cards

As for the display options, actions are issued through cards. By default, the application shows the three orthographic views and a 3D view of the objects placed in the gray square. Activities can then be launched, each activity with a different (red) card. An orange card is used to quit any activity and return to the default mode.

In some activities, generic actions can be issued, such as “check solution”, “show solution”, and “ask for help”. It is not mandatory for each of the activities to implement these functions. Even if the functions are available in the activity, the teacher can easily reduce the interface available to the students by not giving them all the cards. For example, the teacher may not want the students to be able to peek at the solution. In this case, he can keep the “show solution” card and require the students to ask him if they want to look at the correct solution. This is one way to empower the teacher and shows how the cards can be used to improve classroom orchestration.

5.3.5 Activities

It is not a requirement for all the activities to use all the components of the interface. In fact, doing so can result in a cluttered workspace, as will be seen in Chapter 7. The components are the available tools to design meaningful learning activities, and it is up to the activity designer to use them for the best.

The first learning activities were targeted at first year apprentices who were just beginning to study orthographic projections. They were designed by myself based on the official learning material and discussions with the teachers. Their goal was to get the apprentices acquainted to orthographic projections. These activities did not use the drawing tools or the booklet, yet they were close to the drawing activities but tried to focus on the cognitive rich parts.

For example, a typical exercise that is done on paper is to give the students two of the three orthographic views, and ask them to draw the third one. The following similar activity was implemented in TapaCarp: while apprentice B is not looking, apprentice A chooses some blocks and places them to his liking. Using an action card, he tells the system to save the position of the blocks, removes the block, and hands them to his partner (apprentice B). The system shows the front view and the side view of the blocks placed by apprentice A. Based on these two views, apprentice B must place the blocks identically to how apprentice A placed them.

This activity shows the potential of TapaCarp for apprentices to learn in a more intuitive and less “recipe” way. When doing this activity on paper, students first draw the construction lines to link the views together and find each of the points of the object on the missing view. Drawing the construction lines is very mechanical and does not bring any added cognitive value, and yet, it is time consuming. One teacher reported that students often draw the construction lines and then call him to ask him what to do. With TapaCarp, there is no need to draw construction lines and apprentices have to start thinking immediately. To make sure that their solution is accurate, they can use the “Construction lines” card. Using TapaCarp also allows apprentices to create their own “model” instead of having a drawing forced upon them. This can emulate a competition between the two apprentices, and has the benefit of making one apprentice think of how to create a difficult exercise for his partner.

In total, 5 activities with the same philosophy were implemented. The goal of the project was not to implement many activities, neither is it the goal to describe all of them here. Two of them will be presented in the next chapter in the context of the study in which they were used. Activities targeting more advanced subjects in the carpentry curriculum will be described in Chapter 7.

5.4 Implementing Design-Based Research

5.4.1 Types of studies

The DBR approach integrates several research methods, and this was the case in this project too. We did several kinds of studies. The first study was the contextual inquiry reported in Section 3.2.3. The second one was the study of spatial skills, done through the spatial skills test. In the rest of this thesis, five more studies will be presented. Four of them were conducted in a classroom in a carpentry school, but with different approaches. Two of them were controlled studies, close to laboratory experiments, run with the goal of studying the impact of specific design choices on the learning outcome (Chapter 6). The two other studies, presented in Chapter 7, were focused on the integration of the prototype in the classroom, and were less controlled. The final study, described in Chapter 8, was an online study. Besides these five main studies, I also conducted pilot studies and interviews with apprentices and other users during the design process. Table 5.1 summarizes the main studies that will be presented in this thesis.

Depending on the classroom situation, some of the studies involved controlled groups, some did not. Some required the participants to work in pairs for the entire time, whereas in others, participants worked individually, or interacted with each other only punctually. The method and context of each study will be described with the presentation of the study.

5.4.2 Participants

Participants to the studies had ties with three entities: companies, the theoretical school, and the practical school. DBR rests on the assumption that a tight collaboration with participants can be established. In our case, convincing the various participants to come on board with the project did not cause any problem. The companies welcomed us when we visited them for the contextual inquiry. The directors of the two professional schools gave their approval after

Title	Type	Location	Participants	Chapter
Carpentry discovery	contextual inquiry	companies, schools	15	3
Spatial skills	paper tests	schools	726	4
Tangible effect	semi-controlled	theoretical school	46	6
Feedback effect	semi-controlled	theoretical school	56	6
Classroom deployment	uncontrolled	theoretical school	24	7
New learning scenario	semi-controlled	practical school	24	7
Scaling up	uncontrolled	online	340	8

Table 5.1: Summary of studies presented in this thesis.

we explained the goals of the project and demonstrated the tangible interface that had been develop for logistics apprentices. The two teachers from the theoretical school and the teacher from the practical school were enthusiastic. They displayed great availability outside of their teaching hours, but confirmed that the time with their students was precious. Therefore, as a general rule, interventions with apprentices during the teaching hours were kept short and, whenever possible, involved the apprentices that were ahead of schedule. In general, the apprentices were curious to use the system and willing to participate in the studies.

5.4.3 Data collection

During the studies, data were collected in various ways. Observations, semi-structured interviews, questionnaires, audio and video recordings, and the snapshots taken by TapaCarp were the main data supplies for the qualitative analyses. For the quantitative analysis, the data came from the log files of TapaCarp.

The learning gain was measured as the difference between a test taken before the treatment (pre-test) and another test taken immediately after the treatment (post-test). This way of measuring the learning gain has several limitations. It assumes that the treatment develops a particular skill, and that this skill is accurately measured by the tests. Both assumptions are hard to verify: it is not always obvious to identify one particular skill that is developed by the treatment. When it can nevertheless be done, measuring it through tests is problematic: either the tests are similar to the treatment task itself, in which case one can wonder whether the test measures a skill or the ability to learn how to do a task. A test with items further from the task can solve this issue, but it introduces an implicit need for learning transfer. Another limitation is that an immediate post-test does not measure how the learning holds over time. This could be addressed by having the participants take a delayed post-test, but for logistics reasons, I never went down this path in this project.

5.5 Refined Research Questions

At the beginning of this chapter, I introduced the two main research questions of this thesis in a quite generic way. With the description of TapaCarp being done, we can now refine these questions.

The first research direction is to explore whether TapaCarp can be useful for training spatial skills, and if so, which of its features are key to help acquire spatial skills. TapaCarp is neither an activator nor an inhibitor of learning per se. Instead, it offers a set of components and features and what is decisive for the learning outcome is how these components and features can be taken advantage of during the learning activities. For example, TapaCarp offers the possibility to control the virtual representation of an object with a physical object. The physical object manipulated can have the same shape as the one being virtually represented (a *literal* physical correspondence), but it can also have a different shape (a *symbolic* physical correspondence). One hypothesis is that for tasks involving spatial skills training, it is helpful to have a literal physical correspondence between the tangible object and its virtual representation, because it can help the learner build a mental model of the object. This question will be studied in more details in the first part of Chapter 6. Another feature offered by TapaCarp is the dynamic link between the physical object and the virtual representation of its orthographic projections. Being able to observe directly the impact of one's actions can be beneficial for learning, but it can also hinder reflection. The influence on learning of the dyna-linking in TapaCarp will be studied in the second part of Chapter 6.

The second research direction has to do with the pedagogical scenarios in which TapaCarp can be used in the classroom. There are many dimensions to explore regarding how TapaCarp can be used in the classroom (individual versus group versus class level, with existing activities versus with new ones, as a side tool versus as a central element in the lesson, etc.). Obviously, there is no one right usage of TapaCarp, but rather, what is interesting is to explore the opportunities and constraints offered by each types of usage in the classroom, and how each usage impacts classroom orchestration and the way in which apprentices learn. I will report on two different usages that were tested in a classroom (Chapter 7): a first one in which several TapaCarp systems were used in all students in a classroom simultaneously; and a second one in which TapaCarp was used as part of a learning activity in which some steps were done with technology, and others without it, and that encompassed the individual, the group, and the class level. Finally, in Chapter 8 I will describe how TapaCarp's price can hamper its deployment in classroom, and describe how a low-cost TapaCarp can be built that could accelerate its deployment in classrooms.

5.6 Conclusions

In this chapter, I presented the main research goals that I will pursue throughout this thesis, as well as the DBR approach and the tool (TapaCarp) that I will use to pursue them. TapaCarp was presented in a concise way, as the goal in this chapter was not to go into all the implementation and feature details of the system, but rather to give an overview of it. Other features and peculiarities will be described in the following chapters, together with the specific activities that they serve.

This chapter concludes the introductory part of this thesis. So far, I have explained that TUIs have the potential to be useful for learning, and more specifically that they could be useful to train spatial skills in the context of carpenter apprentices. Two research questions were identified: first, what features of TapaCarp could be particularly useful to train spatial skills; and second, what learning scenarios could TapaCarp favor or work well in, in terms of learning. The next chapter presents the results of two empirical studies that tackle the first question.

6 The Impact of Design Variations

6.1 Introduction

TapaCarp is built on the assumption that TUIs can help train spatial skills. This assumption is not unreasonable and stems out from former theoretical and empirical results from the literature on the potential benefits of TUIs for learning (see Section 2.2.3). Yet, it remains to be shown that TapaCarp can actually be helpful for the training of spatial skills, and what aspects of its interface are fundamental to train spatial skills.

Of course, each modification to the interface could potentially have an impact on the learning outcome, and countless experiments could be done to test the effect of each of these modifications. However, some changes have less impact in the specific case of learning spatial skills with TapaCarp. For example, changing the color or the weight of the blocks, while a change in the interface, might not have a direct impact on the development of spatial skills. This chapter presents the study of two design variations that we hypothesized to be closely related to spatial skills learning.

The first one is the physical correspondence (Price, 2008b) between the physical block and its virtual representation. Building a mental model of an object can be hard, and having a physical object identical to the virtual object (i.e. one with a "literal" correspondence) might help build this mental model. It could even make the mental model unnecessary, as the physical block can be used as an anchor for 3D reasoning. Thus, the physical correspondence between the object serving as a manipulation handle to the block and the virtual representations of the block can potentially affect the cognitive process and the learning outcome of a learner.

The second aspect is the dyna-linking between the multiple representations of the block. In TapaCarp, the block is represented as a physical object and through multiple virtual representations (the orthographic projections and the perspective view). Users expect that a movement on the block will impact the other representations immediately, and dyna-linking therefore makes for a good user experience. However, dyna-linking may also lead to a higher exploration and, consequently, less time spent on reflection. This is especially the case with tangibles, where the cost of moving objects is particularly small.

6.1.1 Similarities between the two studies

The two studies exhibit similarities. They were both run in a semi-controlled environment in the same classroom (see Figure 6.1), during a drawing class. The participants were carpenter apprentices in their first year of training, and in both cases they completed the experimental task in pairs. We chose to organize the apprentices in groups for ecological reasons, as they work in pairs most of the time at school, but the task could also be performed individually. In each experiment, participants were administered a pre-test before the experimental task, and a post-test immediately after. Each study featured two conditions. Participants were randomly assigned to one condition. In each condition, they used TapaCarp. The two studies also share some limitations, which I will discuss after the presentation of both experiments.

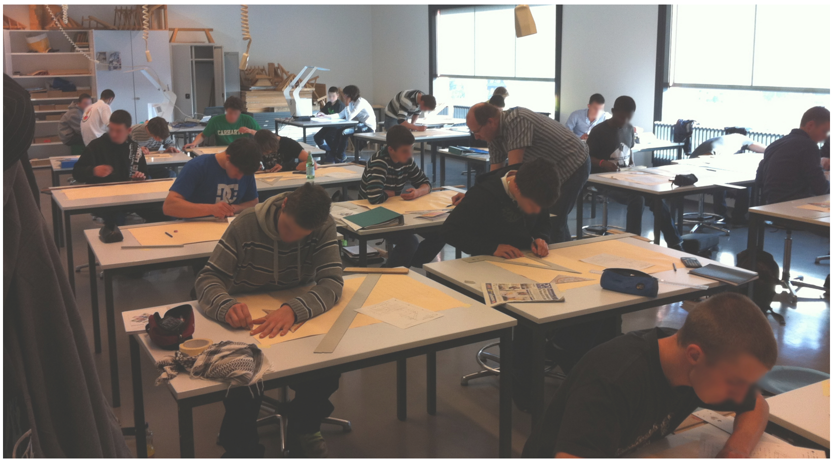


Figure 6.1: The classroom environment in which both studies were conducted.

6.2 The Impact of a 3D Shape

The main potential benefits of TUIs for learning were presented in Section 2.2.3. One of the benefits listed is physicality, which may help reduce the cognitive effort to manipulate the system and allow the learners to focus on the core of their task. For example, in TapaCarp, to change the position and orientation of the block, the users can simply apply physical rotations and translations to the physical object. This is different than in a GUI where the user would typically have to remember a combination of keys and mouse operations to use simultaneously in order to rotate and translate an object. Physicality therefore reduces the gulf of execution for the user because it is easy to move and rotate a block.

In the specific case of spatial skills tasks, physicality can also play an important role, because the haptic and proprioceptive perception of tangible representations may help to perceive 3D shapes better. Further advantages can also emerge if the physical object that serves as a manipulation handle and the virtual object have a matching shape. For example, tasks for which the learner has to imagine how an object would look like from a given point of view may become easier thanks to the ability to overlook the object and to physically rotate around it. The physical object can then serve as an anchor for 3D reasoning, whereas in the absence of such a physical anchor, the learner would be forced to build a mental model of the object and reason mentally, resulting in an increased cognitive load. For carpenters, such reasoning is typically required when making the link between the 2D representations of the object (its orthographic projections) and its 3D representation (the perspective view or the physical object). Making such a link requires mental rotation, a task that is at the core of spatial ability.

The extent to which a physical object matches the virtual object it represents has been described in the framework by Price (2008b) as physical correspondence. If the two objects have an identical shape, their physical correspondence is said to be literal; if the two shapes are far apart, their physical correspondence is said to be symbolic. The study presented in this section aims at measuring the impact of the physical correspondence between the tangible block and its virtual representations when solving a task that requires to match the 2D and 3D representations of an object.

To test this, we created two versions of TapaCarp: one in which the physical correspondence is literal, and one in which it is symbolic. The physical correspondence is the only parameter that changes between the two conditions. In particular, the mode of interaction (rotating the object and translating it) is kept constant in both conditions.

Studies examining the impact of the physical correspondence on learning are few and far between. In a tabletop system to learn about the behavior of light, Price et al. (2009) represented a virtual torch by an actual torch, i.e. a literal mapping. They found that this could be confusing for users, because the torch looked like a torch but did not behave in the same way (e.g. it beamed light while turned off and only worked when on the table).

6.2.1 Experiment and method

Participants

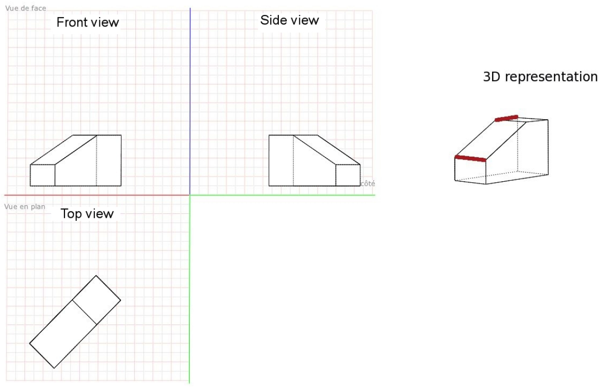
A total of 44 male apprentices in the first year of their training took part in the experiment. They had had little exposure to orthographic projections prior to the experiment. They did the experiment in pairs. Each pair was randomly assigned to one of two conditions (see the "Conditions" paragraph below). To accommodate an uneven number of participants and allow all students to participate, two participants did the experiment a second time. As they did the post-test before doing the experiment the second time, their data were kept for the analysis. In total, 13 pairs completed the experiment in the block condition, and 10 in the token condition.

Interface

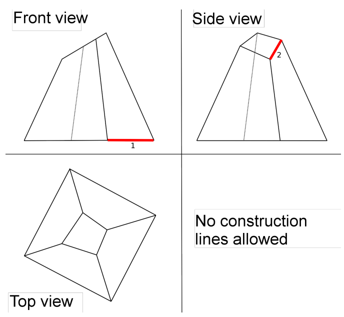
The user interface for this experiment was the standard TapaCarp interface described in the previous chapter, with the three orthographic projections on the left and the block zone on the bottom right. Instructions and feedback appeared as pop-up windows. The interface was identical for both conditions, and there was no time limit.

Task

The experimental task was the following: given a set (2 or 3) of highlighted edges on the 3D representation, identify and select these edges on each of the three orthographic projections (see Figure 6.2a for an example). The selection was done by clicking with the mouse. When two edges were superposed, a right click selected the edge in the background whereas a left click selected the front edge. When an edge was not superposed with anything else, either the left or the right click could be used to select it. The orientation and the position of the virtual object could be modified by manipulating a tangible object.



(a) Task sample: find on each of the three orthographic projections the two edges that are highlighted on the 3D representation.



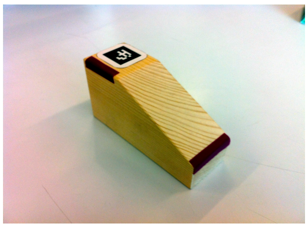
(b) Sample test question: similar to the task, but done on paper.

Figure 6.2: A sample of the task and of the test question.

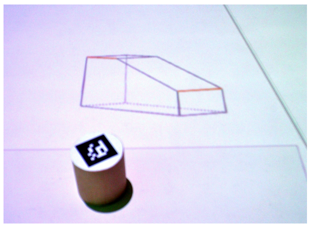
The task was designed with the teacher and was similar to the exercises presented in the curriculum. The rationale behind it is that in order to be able to identify an edge in the 3 orthographic projections, the apprentice must build a mental model of the object, and relate it to the three views. In total, the participants completed 13 instances of the task. The first one was an introductory one completed with the assistance of the experimenter and was not scored. Apprentices were then asked to complete 3 series of 4 exercises (thus a total of 12 items) each without any outside help. All exercises were of the same difficulty level.

Conditions

There were two experimental conditions: the *token* condition and the *block* one. Figure 6.3 shows the same object being represented in both conditions. In the token condition, the tangible object given to the participants to manipulate the virtual object was a small round token (Figure 6.3b). In addition to the 3 orthographic projections, a 3D perspective view of the object was shown above the “Block zone”. The highlighted edges were shown in orange on the 3D representation, as shown on Figure 6.3b. In the block condition (Figure 6.3a), the small round token was replaced by the actual 3D object (i.e. an object with a literal physical correspondence). Only the three orthographic projections were shown virtually, not the virtual 3D representation of the object, because the users had a tangible representation of the object.



(a) Block



(b) Token

Figure 6.3: An object represented in both conditions block (left) and token (right).

Thus, the mode of interaction – and especially the ease of manipulation that comes with a tangible interface – was the same in both conditions, the only difference being the shape of the manipulation handle, and, following, the addition of the virtual 3D representation in the token condition.

Environment

The experiment was conducted during a drawing class, in a classroom in which two TapaCarp systems were set up in the back (Figure 6.1). The apprentices came to the back of the classroom in pairs to participate in the experiment. The experiment was conducted over 12 days with four classes. Although it could be a potential future area of research, this experiment did not focus on the group dynamics.

Learning gain measure

The apprentices passed tests before and after completing the activity (pre-test and post-test). Both tests were completed with paper and pencil and were designed for this experiment with the teacher. The tests were done on paper to avoid measuring how much students learned how to use TapaCarp, and because paper is the common medium in the school and in carpentry. The pre-test contained 3 questions for a total of 12 points and the post-test contained 4 questions for a total of 14 points. Because the tests do not have the same number of points, we use the percentage scores to compare the pre and post performance. We also refer to the relative learning gain (RLG), computed as indicated in Equation 4.1.

A sample question of the test is shown in Figure 6.2b. All questions were identical in their form, and similar to the task questions: the 3 orthographic projections of an object were shown, but not its 3D representation. For each question, one or two edges were highlighted on one view (possibly on different views for different edges), and the task was to find each highlighted edge in the two other projections. Each edge correctly highlighted on one projection was worth one point, giving a maximum of two points per edge. To avoid a mechanical solving process, participants were not allowed to use construction lines.

6.2.2 Results

Both conditions can be compared along several dimensions. The first and most evident one is the learning performance, both in terms of learning gain measured by the tests and by looking at the performance during the experimental task. This reveals differences between the two conditions, but does not explain them. To explain them, we look into other variables, such as the number of trials, the strategies, and the average duration spent on each activity. Finally, we analyze how the shape of a block, and in particular its degree of symmetry, can influence the global performance of such a task.

Learning gain

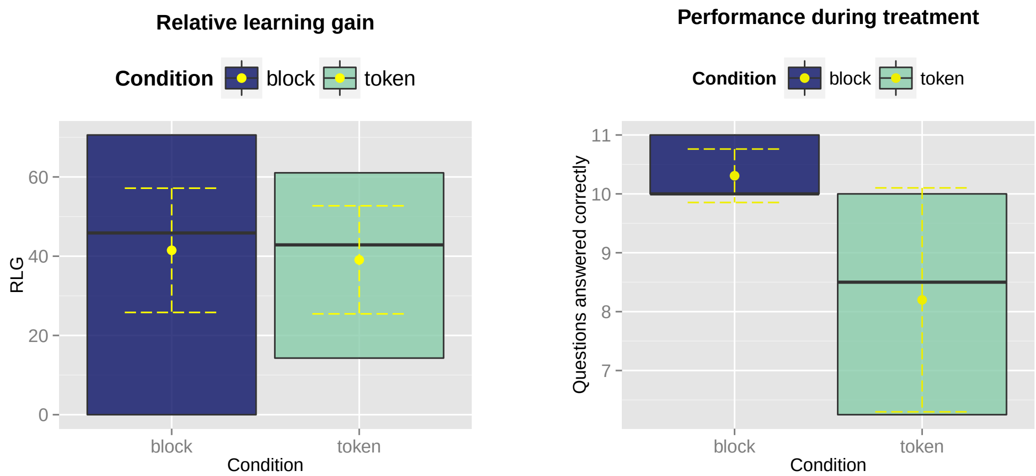
There was an overall improvement between the pre-test and the post-test in both conditions: the average score went from 52.8% in the pre-test to 73.3% in the post-test (+20.5% absolute gain, $t[43]=6.60$, $p=.000$). The improvement was significant for both conditions, but neither the difference of RLG ($F[1,42]=0.05$, $p=.82$) nor the difference in the percentage score ($F[1,42]=0.86$, $p=0.36$) between the pre-test and the post-test were statistically significant between the two conditions. However, the general trend was that participants in the block condition improved slightly more than participants in the token condition (+23.5% versus +17.4% for the difference in scores, and 41.5% versus 39.0% for RLG). This can also be observed in Figure 6.4a.

Each edge marked in the 3D model had to be found on each of the three views. Table 6.1 shows the percentage score for the pre-test and post-test by view. The score for a view is computed as the average score of all the questions on which the edge was shown in the view, or to be found on the view. For example, the question for which the reference edge was shown in the top view and to be found in the side view is included in the score of the top view and the side view. For all combinations of view and condition, the performance improved (positive RLG). For each of the three views, the improvement in the block condition was higher than in the token condition, but not significantly. In fact, none of the differences, either for the same view for different conditions, or for a same group (all, token, or block) and different views, was significant.

The percentage of correct answers for the pre-test and the post-test was the smallest for the side view, for all groups. At the same time, the highest RLG between the two tests occurred on the side view, in both conditions. This is the case even when controlling for the higher possible improvement due to the initial low score on the side view, as the RLG does. One interpretation is that the side view is naturally more difficult, and that the experimental task therefore had a larger effect on this particular view. The fact that the side view is more difficult is corroborated by previous work on mental rotation that concluded that the difficulty of mental rotation is proportional to the angle it requires (e.g. Flusberg and Boroditsky, 2011; Peters et al., 1995). The side view requires performing a 90° angle rotation whereas the two other views do not.

	All			Token			Block		
	top	front	side	top	front	side	top	front	side
pre-test	64.3%	53.8%	41.9%	67.6%	53.5%	42.6%	62.3%	54.1%	43.0%
post-test	79.1%	74.2%	69.5%	76.8%	71.5%	65.3%	80.6%	76.0%	72.2%
RLG	29.3%	39.4%	41.8%	23.0%	37.7%	40.3%	33.3%	40.5%	42.7%

Table 6.1: Percentage of questions answered correctly in the tests.



(a) RLG between pre-test and post-test.

(b) Activities completed correctly.

Figure 6.4: Average performance measured (a) by the tests, and (b) during the experimental task.

Learning performance during the experimental task

Although there was no statistically significant difference between the two conditions in the test scores, there was one in the performance during the activities. One way to observe this is to look at the number of correct answers given by pairs of participants over the 12 activities. An answer to one activity is considered correct if all the edges to be found on each of the 3 views were found correctly. As shown in Figure 6.4b, the apprentices in the block condition completed on average more activities correctly (10.3 vs. 8.2, $F[1,21]=7.5$, $p=.01$)). Another way to compare the performances during the activity is to look at the percentages of correctly chosen edges, also in questions that have not been correctly answered as a whole. Apprentices in the block condition had a higher ratio of correct edges than apprentices in the token condition (97.2% versus 92.3%, $F[1,21]=10.2$, $p=.004$).

The performance improved over time while completing the activities, as shown in Figure 6.5 ($F[1,252]=8.01$, $p=.005$). When splitting the participants by condition, the progression through-

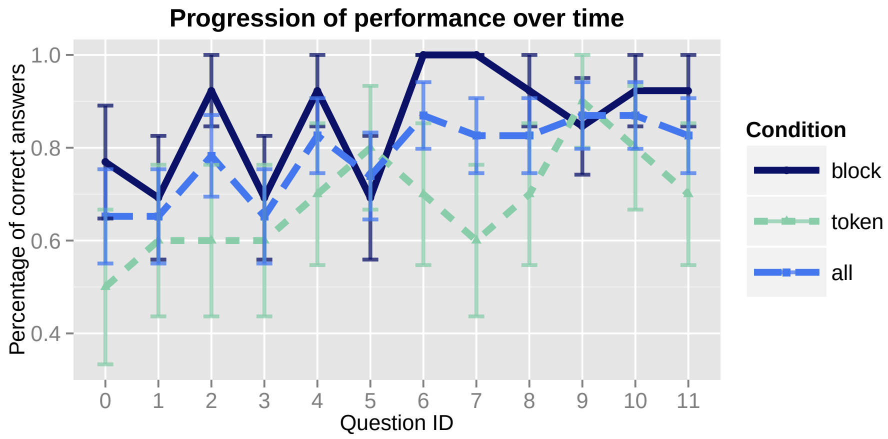


Figure 6.5: Percentage of the average number of correct answers given during the experimental task.

out all the activities was significant in the block condition (F[1,142]=4.6, p=.03), but only marginally significant in the token one (F[1,109]=3.7, p=.059, Figure 6.5).

A slight ceiling effect may be observed on Figure 6.5 and later on Figure 6.8. On the first one, one can see that the overall lowest score for a given question is close to 0.7, which means that on average 70% of all the pairs answered correctly to a question. The average ratio of correct edges per block type, shown on Figure 6.8 shows that even for the block with the lowest score, 90% of edges were found correctly. This ceiling effect indicates that the experimental tasks might have been too easy, and we will look below into other indicators (e.g. number of clicks, duration) that will give more information on the differences between the two conditions.

Similarly to the test results, the performance improved more on the side view than on the top and front ones. Looking at the number of mistakes on each view, the only combination of view and condition for which the number of mistakes went down significantly is the side view in the block condition (F[1,142]<5.1, p=.025). There is a similar trend for the side view in the token condition, although it does not reach statistical significance (F[1,109]=3.17, p=.078).

About 15% of the answers given by the participants contained a mistake on the side view (either a missed edge or an extra edge selected). This is to compare to 6.8% and 5.8% for the top and side views, respectively. Figure 6.6 shows the number of mistakes made on each view by the groups in each condition. Only for the side view is the number of mistakes significantly different between the two conditions (F[1,21]=7.34, p=.013). Indeed, pairs in the token condition made on average 9.7 mistakes on the side view, whereas pairs in the block condition only made on average 3 mistakes on the same view.

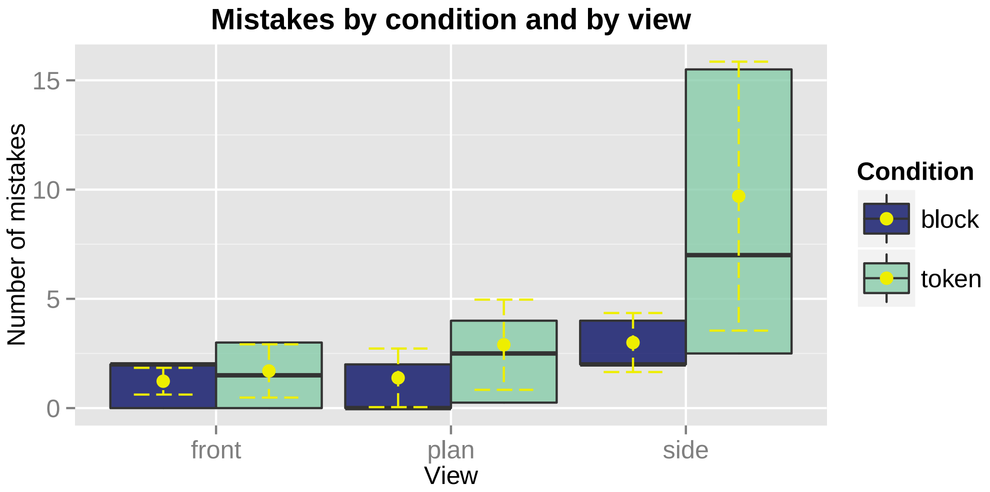


Figure 6.6: Average number of mistakes made per view during the experimental task.

Number of trials

The number of selections and deselections indicates the number of self-corrections performed by the participants, and can be interpreted as the difficulty that participants had to find the correct answers.

Per condition. Based on the ratio of the correct answers given during the experimental task, it seems that it was harder to complete the activities with the token than with the block. The number of edges selected and deselected throughout the activities confirms that. The minimum number of selections for each view corresponds to the number of edges to find. Since there were 9 questions with 3 edges to find and 3 questions with 2 edges to find, the minimum number of selections was 33 for each view. Because there were three views for each of the questions, the total minimum number of selection per group was 99. The average number of selections made by a pair in the block condition was 125.3, which corresponds to an overhead of 25%. The overhead was higher for pairs in the token condition: 145.8 selections on average, which represents an overhead of 45% ($F[1,21]=5.4$, $p=.029$).

Per view. The number of selections and deselections also varied according to the view. For all views, there were more clicks than the minimum required to complete all the exercises correctly. However, the overhead differed in its proportion: the top view showed an overview of 10.8%; the front view, 45.2%; and the side view, 61.0%. This confirms that it was harder to find edges on the side view than on the other views.

Strategies, collaboration, activity duration

Strategies. Most of the groups adopted a "group-by-view" strategy, meaning that they searched for all the edges on one view before moving on to the next one. The most common strategy was to find the edges on the top view, then the front view, and finally on the side view (18 groups, 9 in each condition). Three groups, all in the block condition, adopted the exact reverse strategy: side view first, then front view, and finally the top view. Participants never used a "group-by-edge" strategy, hinting that the cognitive cost of switching view was higher than the cognitive cost of switching edge.

Collaboration. The aim of this study was to examine the effect of the block compared to the token, and we did therefore not focus on the type of collaboration within a pair, nor did we enforce any balance in the collaboration. What we noticed was that apprentices collaborated well most of the time. There was only one mouse, so only one participant could click on the edges, but the participant without the mouse often pointed (with a finger) where to click. In a few groups, participants exchanged places every now and then. Another behavior worth noting was that some participants got up and went to the side of the table in order to be able to see the side view and help with the task on this particular view.

Activity duration. The average duration to complete the activities was 20 minutes and 30 seconds (SD: 5 minutes and 35 seconds). Completing the activities in the block condition took sightly less time than in the token condition, but not significantly (19 minutes and 30 seconds versus 21 minutes and 54 seconds, $F[1,21]=1.08$, $p=.31$). As indicated by the rather large standard deviation, there were dramatic differences in speed between the groups, especially in the token condition: the slowest group completed the activities in 34 minutes, while the fastest group only took 12 minutes (SD: 7 minutes for token versus 4 minutes and 11 seconds for block). The average activity duration slightly decreased over time. However, the most noticeable difference was between the first activity and the rest of them, hinting that during the first one or two activities, participants had to get acquainted to the system and the task.

Performance by block

In total, 7 different blocks were used over the 12 activities. The blocks are shown in Figure 6.7. The performance by block is shown in Figure 6.8 as a ratio of correct edges. The ceiling effect mentioned before explains why there is no box on the graph for some of the blocks.

One block that stands out is the "B" block, for which the performance was lower than for the other blocks. There are two possible reasons for the lower performance with this block. The first one is that it was used only once over the 12 activities and was placed in second position. As explained above, the performance improved with time, so the early placement of this block might explain part of the lower performance. A second reason is that this block has a high degree of symmetry (3 symmetry planes). A high degree of symmetry may make it harder to find reference points when trying to map 2D and 3D representations of the object.

The influence of the symmetry of a block on the performance is shown in Figure 6.9. The degree of symmetry is the number of symmetry planes on a block. There was an interaction effect between the symmetry and the view on which the edge must be found ($F[8,659]=4.9$, $p=.000$). This interaction effect is mainly due to the performance drop on the side view for

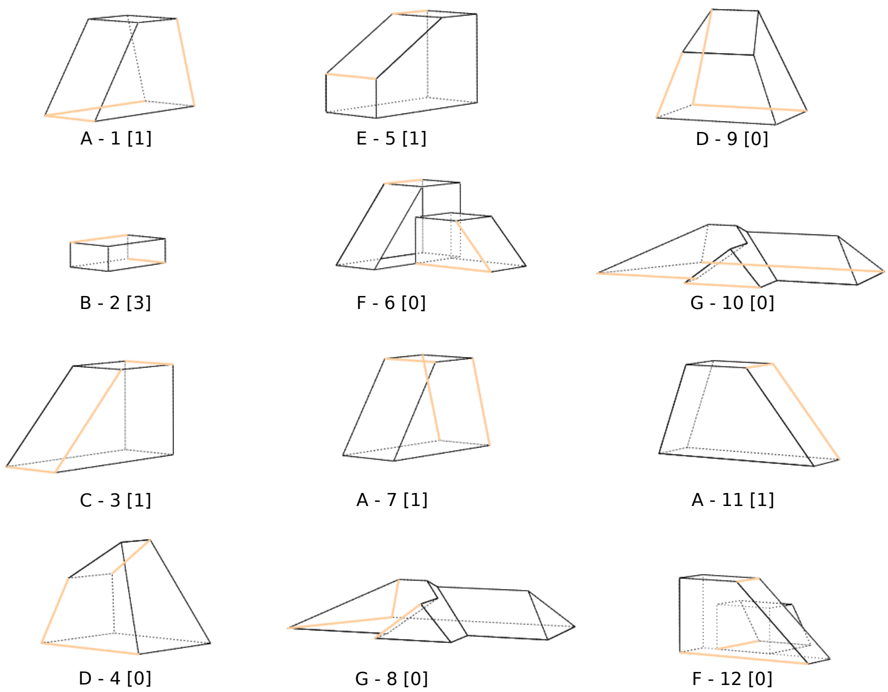


Figure 6.7: The 12 tasks solved by the participants. The letter under a figure indicates the block identifier. The first number is the rank of the question and the number in brackets is the degree of symmetry of the block.

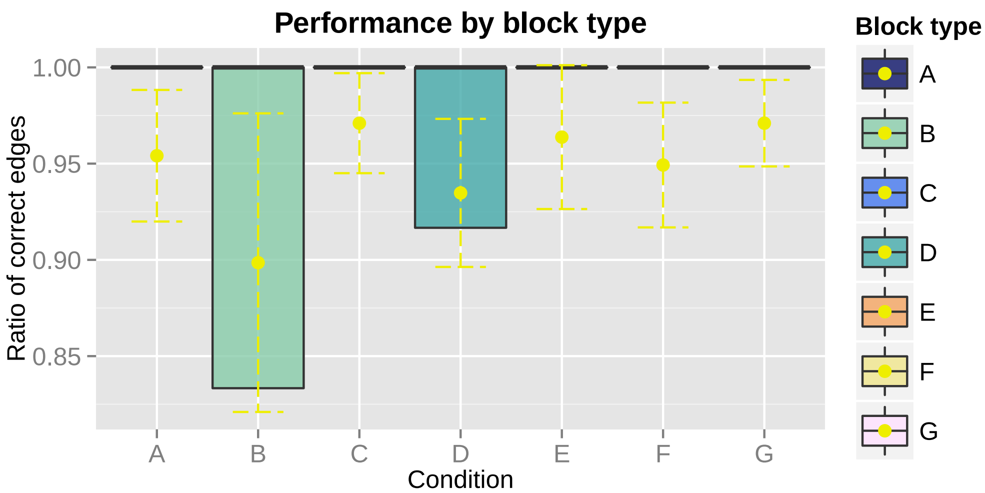


Figure 6.8: The ratio of correct edges per block type.

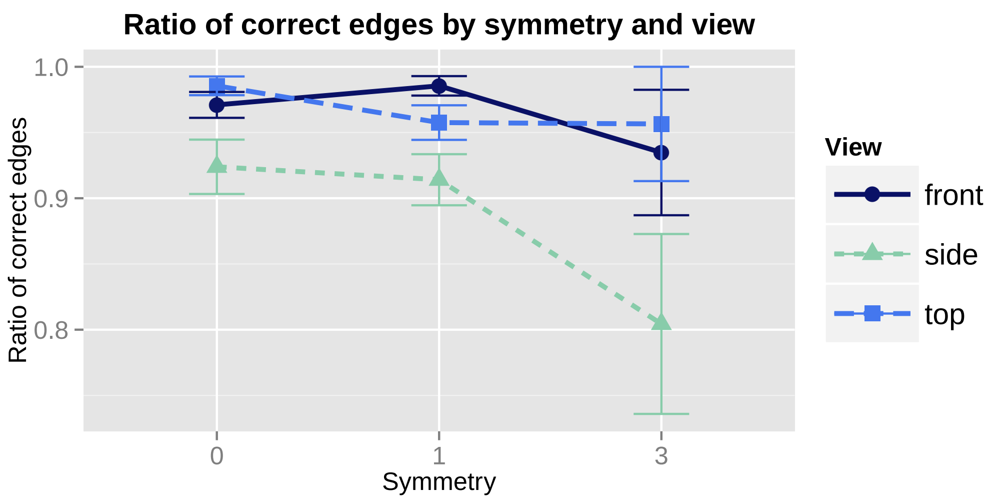


Figure 6.9: Ratio of edges found correctly by the degree of symmetry of the block to which the edge belongs and by the view on which an edge must be found.

a degree of symmetry of 3, whereas the performance for the other two views is almost flat. Nevertheless, in general, a higher degree of symmetry led to a significantly lower performance ($F[2,136]=3.2$, $p=.04$).

6.2.3 Discussion

There are three main outcomes from this study: (1) the strong positive learning outcomes in both conditions, and the difference between them in terms of performance during the task; (2) the difference of difficulty between the 3 views; and (3) the link between the degree of symmetry of a block and the difficulty to find an edge on the block. The implications of each of these findings are discussed below.

Differences between the two conditions

The main goal of this study was to examine the impact of varying the physical correspondence between a tangible object and its virtual representations. There was a performance difference between the token and the block condition during the activities. Participants in the block condition completed more activities correctly, and found more edges correctly overall. Their scores significantly improved over the course of the activities, whereas participants in the token condition did not.

Despite the differences observed during the experimental task, there was no statistically significant difference in the learning gain between the two conditions, as measured by the pre-test and the post-test. In both conditions, the participants performed significantly better in the post-test than in the pre-test. Those in the block condition improved slightly more than those in the token condition, but not significantly. The slight ceiling effect observed could explain the lack of statistical significance. Thus, it may be that the differences observed between the two conditions would have been larger with more difficult questions and a longer experimental task.

In both conditions, participants used the same “group-by-view” strategy. Participants in the block condition were slightly faster at solving activities and did significantly fewer trials (clicks). This, together with the higher success rate of participants in the block condition, indicates that doing the exercises in the block condition was easier than doing them in the token condition. Some participants who informally tested both conditions after the experiment confirmed that.

Making a task easier reduces the cognitive effort required from the student, but does not necessarily benefit learning. On the one hand, learning might be inhibited if the task is facilitated such that the cognitive mechanisms which trigger learning are unsolicited. On the other hand, learning might be increased if the process of accomplishing the task is made easier but the cognitive mechanisms essential for learning remain. In this study, although the block condition was easier, it did not inhibit learning: the improvement in the block condition was larger during the experimental task than in the token condition, and the learning gain between pre-test and post-test was similar in both conditions.

Why did the block help participants? Finding a definitive answer would require doing more experiments, but we can make some hypotheses. The literature has shown that once one has built a mental model of an object, imagining the rotation of the object with a gesture, or even imagining pulling a string to rotate the object, helps perform mental rotation (Shepard and Metzler, 1971). Thus, one hypothesis is that the block helped participants when build a 3D mental model of the object. The block contains information that the eye can perceive and integrate when building a mental model of the object, such as lighting nuances and shadows, and the spatial information that is readily available. By contrast, the wireframe representation displayed in the token condition can be harder to translate into a mental model. Having a more accurate mental model of the object would then allow students in the block condition to perform better when having to imagine how the object would look like from another point of view. One way to test this further would be to replace the block in the block condition by a physical wireframe of the block. This finding could be of importance for TUIs, since – if confirmed – it would mean that tangibles could be beneficial in tasks that require the learner to create a mental model of an object.

In the pre-test and post-test, the students did not receive any perspective view of the objects on which they had to select edges. They were also not allowed to build construction lines between the views. The intent was to force them to create a mental model of the 3D object before being able to do the task. The absence of difference in the learning gains between the two conditions indicates that if the block indeed allowed students to create a better mental model of the object during the experimental task, this did not transfer to the post-test, i.e. when the block was removed. In a different context, Glenberg et al. (2004) showed that, while manipulating objects helped improve the young children’s reading comprehension, *imagined* manipulation was even more effective. More specifically, the effects of imagined manipulation on reading comprehension carried on even once the objects were removed, whereas the effect of manipulation disappeared once the objects were removed. It would therefore be worth exploring if telling participants explicitly to imagine mentally the object in our task, too, would result in a larger effect once the 3D representation of the object is removed.

Another reason that could explain why the block helped students is that it was a more effective anchor to bring the multiple representations of the block together, than the token or the wireframe perspective. When identifying a particular edge in an orthographic projection, students can map the orthographic view of interest either to the 3D representation of the object (be it the block or the perspective view), or to another orthographic projection in which they have already identified the edge. In both cases, linking these different representations together can be hard for novices, and being able to relate one view back to the physical block could help one get his bearings.

Particularity of the side view

An analysis of the performance for each of the 3 views (top, front, and side) revealed that the side view led to more difficulties than the other two views. This was the case both in the tests and during the 12 activities of the experimental task. This is consistent with the existing literature on spatial skills, which states that the difficulty of a mental rotation is proportional to the amplitude of the rotation (e.g. Peters et al. (1995); Flusberg and Boroditsky (2011)). Because the side view requires mental rotation of at least 90°, it is expected to be more difficult than the two other views, which require a smaller mental rotation for the participant to mentally align the 3D representation of the object with the corresponding orthographic projections.

While the hardest one, the side view was also the one on which participants had the larger RLG. It was also only on the side view that the number of mistakes during the experimental task decreased significantly. This is a meaningful result from a pedagogical point of view, as it shows that TapaCarp can help students improve where they need it most.

The side view further crystallized the differences between the two conditions: participants in the token condition made significantly more mistakes on the side view, whereas the number of mistakes in the other views was comparable for the two conditions. There are two plausible explanations for this. One is that in the block condition, participants could pick up the block and move it in space, in other words they could do the rotations physically instead of doing them mentally. However, such interactions did not occur very often (2 groups). Another explanation is that the block gave the participants a better representation of the object, leading to an easier representation of what the corresponding orthographic projection would be. This second explanation is supported by the fact that only participants in the block condition started by looking for the edges on the side view first, whereas all participants in the token condition proceeded with the top view first.

Symmetry of the object

Some blocks were harder to deal with than others. The higher the degree of symmetry of a block, the lower the performance on this block. This is not necessarily intuitive, as in other types of geometrical exercises, the difficulty is often increased by using a block of higher geometrical complexity (Vetter et al., 1994), but it can be explained by the fact that the higher the degree of symmetry, the more difficult it is to find reference points on the object. Therefore, it becomes harder to figure out the orientation of the object, and by extension, harder to find a specific edge. This was especially the case for the side view.

Summary

This study investigated the effect of varying the physical correspondence between the tangible representation of an object, and its virtual ones. In one condition, the tangible representation of the object had the same shape as its virtual ones (literal physical correspondence), whereas in the other, it was a simple round token (symbolic physical correspondence). In both conditions the participants' spatial skills improved significantly after solving exercises on TapaCarp. This hints that TUIs can benefit learning, although a comparison with a control group would be needed to assert this. There was no significant learning differences between the two conditions, but during the experimental task, participants using the object with literal physical correspondence performed significantly better than those using the round token. This was especially the case when the difficulty of the problems increased – i.e. when the degree of symmetry of the block was higher or that the rotation required was larger (side view).

6.3 The Impact of Feedback

The previous study showed that a variation in the interface of the TUI can impact the learning experience in the context of spatial skills. In this second study, we focus on a change in the *activity features* rather than in the interface. More precisely, we examine the information that is provided by the system to the users in reaction to their input, i.e. the *feedback* given by the system. When they move a block, TapaCarp users expect that the movement that they apply to it be immediately passed onto the virtual representations of the block. Indeed, coupling the tangible input with a visual output and dynamically linking multiple representations of the same object is the cornerstone of the high usability and engagement of tangible interfaces (Ainsworth, 2006; Manches and O'Malley, 2012; O'Malley and Stanton Fraser, 2004). This immediate feedback (also referred to as dyna-linking or coupling), participates to the positive user experience that TUIs can foster. It also encourages exploration. For example, in TapaCarp, it is natural for a novice user to move the block around and observe the consequences of the movement on each of the three orthographic projections.

However, for effective learning, exploration needs to be balanced with other activities, such as reflection. The risk with a continuous and immediate feedback is that it leads to too much exploration and acts as a kind of prosthesis that prevents the activation of the targeted cognitive processes required for learning. This is emphasized with tangibles, because their ease of manipulation reduces the cost of trial-and-error. It can therefore be important, as noted by other researchers (e.g. Do-Lenh, 2012; Price et al., 2009), to slow down the manipulation of objects, so that students spend more time reflecting.

Feedback is a good way to influence the manipulation and exploration done by students. The way in which feedback is provided to learners can direct influence their manipulation behavior. Delayed feedback, provided when learners have completed a task, gives them opportunities to fail and to trigger cognitive processes to repair their errors. On the other hand, when feedback is too delayed, learners may fail to see the link between the feedback and their behavior. The kind and timing of feedback is thus crucial, and this is what this study aims to assess, by exposing participants to a TapaCarp in two conditions: some of them were provided with coupling, while others were not. They completed tasks related to spatial skills training, and learning was measured as the improvement of spatial skills specific to orthographic projections. We examine the impact of the presence or absence of coupling on the behavior and learning of participants.

6.3.1 Dyna-Linking and feedback

Dyna-linking

As mentioned in Section 2.2.3, multiple external representations (MER), and the related notion of dyna-linking, are often cited among the potential benefits of TUIs for learning. The main arguments are that dyna-linking could reduce cognitive load through offloading, or could encourage students to explore more. However, there is little empirical data to support these benefits, and dyna-linking could also play a negative role for learning. For example, dyna-linking could allow students to explore the links between representations without reflecting or planning their problem-solving moves. In cases where the solution is shown in one of the representations, the learners could solve the problem by perceptually aligning the representations by trial-and-error. This is not desired, as reflection is a key aspect of learning.

One way to counter this effect is to vary the feedback provided to learners depending on their performance: the better a learner gets, the less support he gets. This has been suggested before, for example by Price et al. (2009), who argued for the need to "specifically design learning activities that slow down interaction and promote opportunities for reflection". Do-Lenh (2012) also observed that the users of their tangible system tended to perform too many manipulations and not reflect enough. In a subsequent study (Do-Lenh, 2012) they introduced reflection tools in their system and showed that it had a positive impact on students' learning.

Feedback

The debate about immediate versus delayed feedback is as old as education. Novel interfaces have revived it, as a tight coupling of input and output (i.e. immediate or continuous feedback) contributes to their success in terms of usability and engagement. For instance, any short delay between input and output makes multi-touch applications difficult to use. Although the feedback presented in this experiment is given by a computer (as opposed to a human), the main findings on feedback are not specific to use cases with computers, and we therefore frame this work in the general literature on feedback.

Hattie (1999) found that feedback is in the top 10 highest influences on achievement in learning. Hattie and Timperley (2007) discovered through a meta-analysis that feedback is most powerful in the context of faulty interpretations, and not lack of information; when goals are specific and challenging but task complexity is low; when it provides information on correct rather than incorrect responses and when it builds on changes from previous trials.

Based on their findings, Hattie and Timperley (2007) developed a model for feedback where they distinguish four levels at which feedback can happen. It can be about a specific task. It can be directed at the process used to complete the task. It can focus on self-regulation. Finally, it can also be personal and directed at the person. In agreement with Kluger and DeNisi (1998), who found that feedback is most efficient when it is close to the task as opposed to close to the person, Hattie and Timperley (2007) found that feedback directed at the person is often unrelated to the task and does not impact learning positively. They also indicate that feedback on process is more effective for deeper understanding than feedback on the task.

The timing of feedback is an important factor with regards to efficiency. In a meta-analysis of 53 studies, Kulik and Kulik (1988) showed that immediate feedback is beneficial at the task level, but that at the process level, some delay is beneficial. Similarly, in a study involving high school students who completed a computer-based lesson, Clariana et al. (2000) found that delayed feedback was effective when greater degrees of processing were needed and for difficult problems.

However, the impact of feedback seems to depend on the learning task. For example, students using a microcomputer-based laboratory where a motion graph is updated in real-time as a consequence of physically moving a toy car understood the motion graph best when they got immediate rather than delayed feedback (Brasell, 1987). This contradicts Clariana et al.'s results. Thus, although it has been widely studied, the impact of feedback is not a closed topic, both because they are many possible types of feedback, and because feedback is highly dependent on the learning task.

Terminology. Different authors use different terminologies to describe the various levels of behavior regulation. Following the terms of Anderson (2002), the real-time coupling of physical manipulation and system augmentations corresponds to the level of operations. Users are informed every second (or even more often) about the effect of their actions. The provision of feedback about whether a question was solved correctly corresponds to the level of unit-task. Feedback about performance of a series of questions corresponds to the level of task. In the terms of Hattie and Timperley (2007), coupling the users' manipulations with augmentations in tangible environments is a form of process feedback and providing feedback about the correctness of the solutions is a form of task feedback. In this chapter, I will refer to the process and task level feedback.

Question

One of the features of tabletop TUIs is that they allow users to have external representations, and to have them co-located with their display (Price, 2008b). Co-located displays increase the potential for dyna-linking by making it possible to show to the user a direct feedback of his actions on top of the TUIs. Our specific question in this study is whether providing feedback at the process level is beneficial for learning. The goal of this work is not to investigate feedback in general. However, dyna-linking, which is similar to process level feedback, is one of the key features provided by augmented reality systems, and is therefore worth investigating.

6.3.2 Experiment and method

Participants

Fifty-six male carpentry apprentices in their first year of training participated in the experiment. The age of the participants ranged between 16 and 21 years, with a couple of participants in their late twenties. Participants had performed some exercises with orthographic projections prior to the experiment. They completed the activities in pairs. The pairs were formed freely by the students. Each pair was randomly assigned to one condition. In total, 14 pairs completed the experiment in the coupling condition and 15 in the no-coupling condition. As for the previous experiment, two students did the experiment twice to accommodate an uneven number.

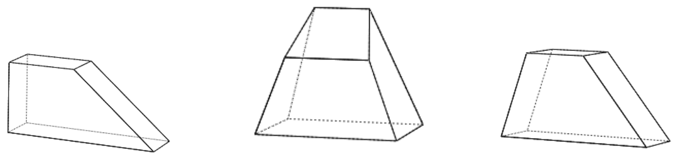
Environment

The environment and methods were the same as for the first experiment described in this chapter: two TapaCarp systems in the back of the classroom, to which apprentices came in pairs during a drawing class. The apprentices passed a pre-test and a post-test, each composed of 6 questions. There were three types of questions: (1) given the top view of an object, students had to choose the corresponding front and side views from 8 candidates (3 questions); (2) given the side view and front view of an object, students had to pick the corresponding top view from 6 candidates (2 questions); and (3), given the side view of an object, students had to draw the corresponding front view. The tests are shown in Section A.4. There was no limit of time for either the tests or the experimental task.

Task

The user interface was the standard TapaCarp one, with two exceptions. First, the top view was not shown, because it would have been too easy to match the position of the objects given the top view. Second, a grid pattern with squares of 2 centimeters was shown on the side and front views and on the block zone. The goal of the grid pattern was to give the user points of reference between the two views and the block zone.

The activities contained a total of 24 questions. Each question consisted of four steps, shown in Video 2. First, students had to place a block in the block zone at a position and orientation that the system indicated by projecting its top view. Second, once the block was placed correctly, a



(a) Block 1 (b) Block 2 (c) Block 3

Figure 6.10: The shape of the three blocks. Block 1 and 3 were made out of wood and block 2 was made out of folded cardboard.

movement was applied virtually to the object and the new position of the object was shown on the front and side views. Third, the students had to move the block to this new position. This transition from one position to the other had to be done with as few movements as possible. Fourth, when they were satisfied with their solution, students could check their solution by placing a specific token on the tabletop. If the solution was wrong, the correct solution was displayed as a projection of the top view in the block zone for 20 seconds, and students were asked to place the block as indicated. After this time, or immediately after if the solution was correct, the next question started from the last position.

The three blocks used are shown in Figure 6.10. Each of the blocks was used for 8 consecutive questions. Besides finding the correct position, the students were also asked to maximize a score that was displayed permanently on the top right of the projection zone. The score started at a 100 and was inversely proportional to the length of the path done by the block. Moving the block between the starting position and the correct solution could be done without losing any points, but not taking the shortest path led to a loss of points. The score was meant to encourage the use of mental rotation and movement planning over trial-and-error.

Question types. For each question, the block was moved either by a translation (on one or two axes), or by a translation combined with a rotation. Thus, the movement applied to the block could have from one to three dimensions. In the rest of this chapter, I will refer to questions that had a movement with a translation only on the y-axis as “y” questions, to questions with a translation on x and y as “xy” questions, etc. It is worth noticing that in a “y” question the side view changes and the front view stays the same, while it is the opposite for an “x” question, and that for “xy” questions, both the front and the side view change.

Conditions

There were two experimental conditions: *coupling* and *no-coupling*. The only difference between the two conditions was that in the coupling condition, the tangible representation of the block was dynamically linked to the virtual representations shown on the front and side views. In other words, students in the coupling condition could see in real-time on the two orthographic projections the effect of moving the tangible block. In terms of feedback, the students in the coupling condition were provided with immediate feedback at the process level, whereas students in the no-coupling condition did not receive any process feedback

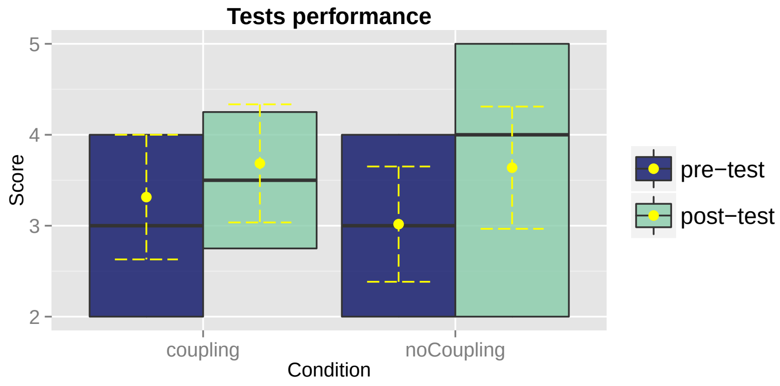


Figure 6.11: Comparison of the pre-test and post-test score by condition.

(except for the score). Both conditions offered immediate and delayed feedback at the task level. The immediate feedback was given to students as "correct", "almost correct", and "wrong" as soon as they validated their answer to a question. The delayed feedback appeared at the end of each series of 8 questions and summarized the result of each question (right/wrong).

6.3.3 Results

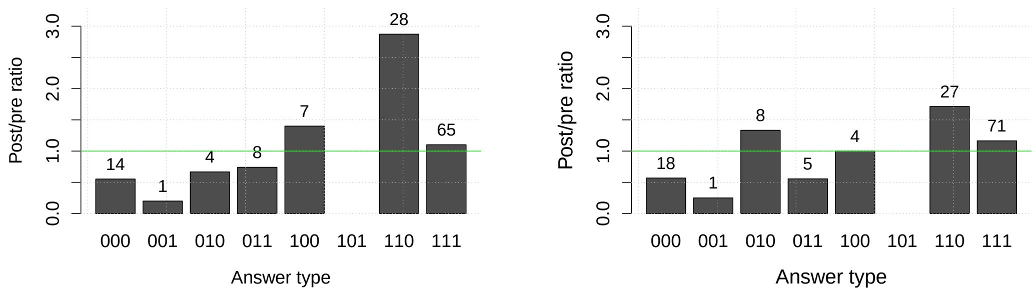
The presentation of the results follows the same order as in the first experiment presented in this chapter: first the learning performance measured by the tests, then the learning performance during the experimental task. An analysis of the answers given in both tests will help understand where the improvements were the largest. For the performance measured during the experimental task, the evolution of the performance will point out differences between the two conditions. Finally, we will look at other indicators, such as the type of paths chosen and the speed at which the block was moved, that will help explain how the type of feedback influences reflection.

Learning gain

There was no difference in the RLG between the two conditions (F[1,54]=.069, p=.79), as shown in Figure 6.11. However, there was an overall learning gain independently of the experimental condition (t[55]=2.28, p=.03), as the mean score went from 45.2% in the pre-test to 52.3% in the post-test. A separate t-test on each condition further shows that students in the no-coupling condition improved significantly between the pre-test and the post-test (43.1% versus 52.0%, t[28]=2.21, p=.035), but that those in the coupling condition did not (47.4% versus 52.6%, t[26]=1.07, p=.29). Students in the coupling condition performed slightly better in the pre-test, whereas in the post-test, the scores in both conditions were almost identical.

Answer types. Looking at the types of answers can help understand what caused the overall improvement. The coding scheme of a signature for an answer type is as follows: <orientation

is correct><front view is correct><side view is correct>. So for example, the 110 signature means that the orientation and the front view were correct, but the side view was not. Figure 6.12 shows the ratio of answers per answer type between the pre-test and post-test for each condition. There is no ratio for the 101 signature because no answer in the pre-test had this signature. The horizontal line indicates the limit between an increase and a decrease. The numbers on the bar is the number of occurrences of this answer type in the post-test.



(a) Coupling. (b) No-coupling.

Figure 6.12: Change of proportion in the type of answer between the pre-test and the post-test. Coding scheme: <orientation is correct><front view is correct><side view is correct>.

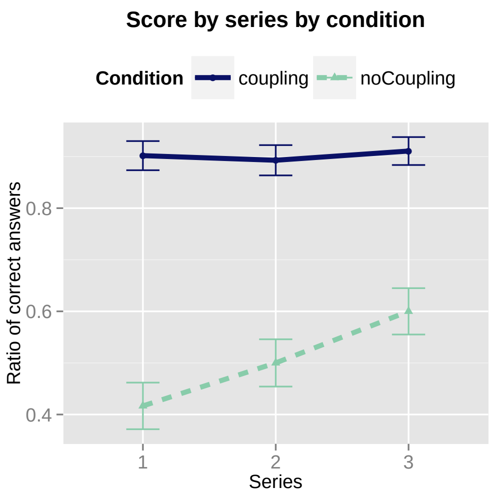
In both conditions, the number of correct answers (111) increased, with a slightly larger increase for the no-coupling condition. All answer types for which the front view was wrong ($*0*$) were equally or less frequent in both conditions, except for the 100 type that increased slightly in the coupling condition. Furthermore, 110 answers were the ones that increased the most, especially in the coupling condition. This indicates that students progressed at identifying correctly the orientation and the x-position (shown in the front view) of the block, but had more trouble finding the correct y-position (shown in the side view).

There were some differences between the two conditions. For example, while the 110 answers increased the most in both conditions, the increase was more than 1.5 times larger in the coupling case (287.2% versus 172.2%). In the coupling condition, the number of answers for which the rotation was correct ($1**$) increased more (+179% versus +129% on average). The number of 101 answers in the post-test was slightly more frequent in the no-coupling condition (38% versus 29% for the coupling condition, not shown in the graph). Again, this suggests that students in the coupling condition improved their ability to identify a correct orientation and positioning in the front view, but not in the side view.

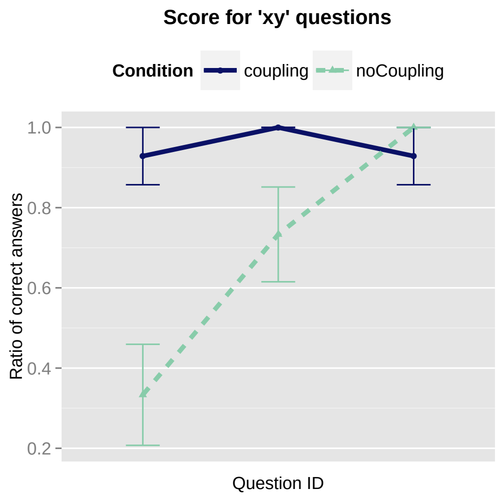
Score evolution during the experimental task

There was a significant difference in the experimental task performance between the two conditions ($F[1,27]=65.3$, $p=.000$). Pairs in the coupling condition completed 90.0% of the questions correctly versus 50.5% in the no-coupling condition. The ratio of correct answers increased on average significantly with each question in the no-coupling condition ($F[23,336]=4.24$, $p=.000$) but only marginally significantly in the coupling condition ($F[23,312]=1.5$, $p=.07$). This can also be seen when grouping the scores by series (Figure 6.13a), but must be taken with some care since the apparent ceiling effect in the coupling condition may explain part of it.

There were four main types of questions depending on the type of movement applied to the block. The movement varied along 2 dimensions with 2 modalities each: translation (on one or two axes), and rotation (with or without). The only question type for which the performance significantly improved with time was the questions that involved two translations, without any rotation (Figure 6.13b). There was also an interaction effect between the condition and the evolution over time for this type of questions ($F[2,54]=8.0, p=.001$), indicating that students in the no-coupling condition improved more over time on this type of questions. Again, this result should be taken with a grain of salt given the ceiling effect in the coupling condition, and the fact that there was only 3 questions of this type.



(a) Results by series, which each contained eight questions. All questions in one series used the same block.



(b) The interaction effect between condition over time for the “xy” type of questions, i.e. the questions that had one translation on each axis (x and y axes), but no rotation.

Figure 6.13

Behavioral differences during the experimental task

Waiting time and movement speed. Students in the coupling condition took on average 67 seconds to complete a question, compared with 89 seconds in the no-coupling condition, a marginally significant difference ($F[1,27]=3.45$, $p=.07$). In the coupling condition, students waited less before moving the block for the first time (8 seconds versus 14 seconds, $F[1,27]=16.88$, $p=.000$). From the moment a question started to the moment they checked their solution, students in the coupling condition moved the block for about 10% of the time, twice the percentage of the no-coupling condition ($F[1,27]=21.3$, $p=.000$). In the same vein, the average moving speed in the coupling condition was half the one in the no-coupling condition

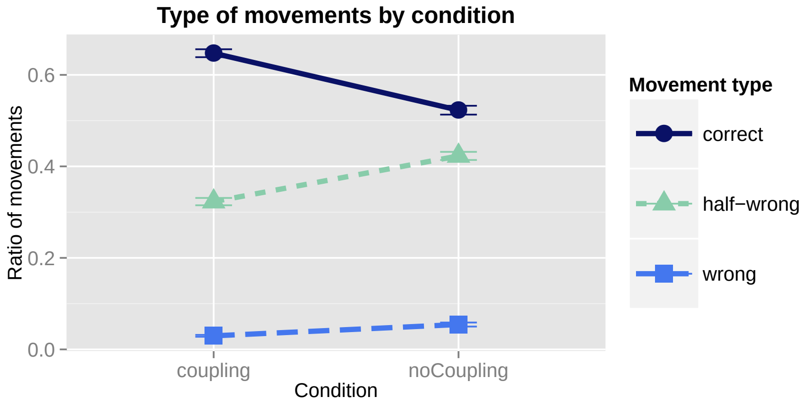


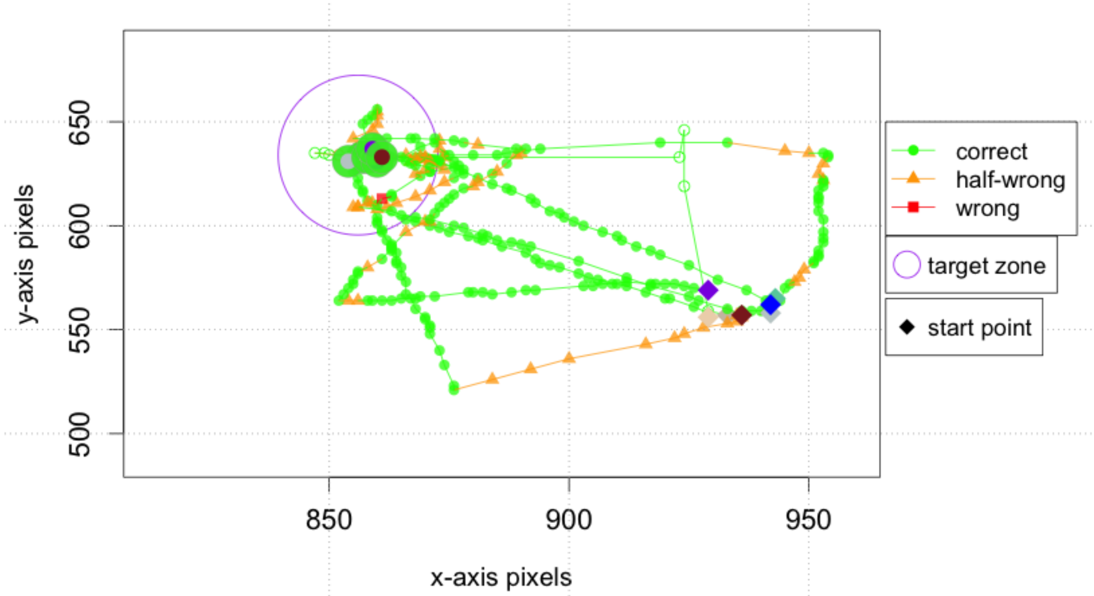
Figure 6.14: The ratio of correct, wrong, and half-wrong movements by condition.

(F[1,27]=16.88, p=.000). This indicates that students in the coupling condition thought less before moving the block, and moved it more slowly than those in the no-coupling condition.

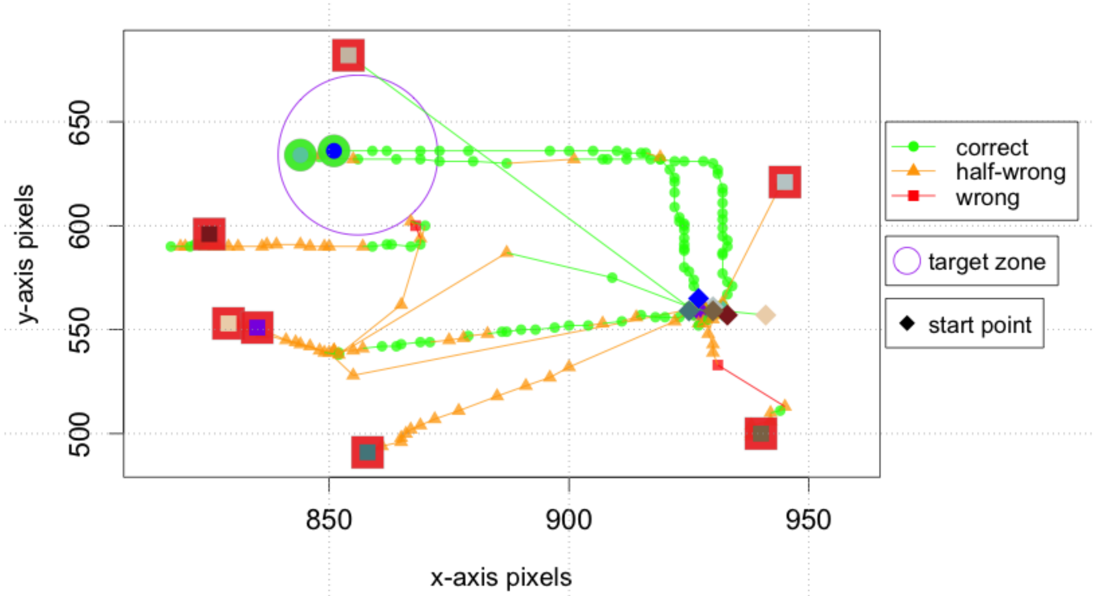
Path analysis. The position of the block is captured about 8 times per second (it can vary depending on the load on the system). The movements of the block can be computed based on two consecutive positions of the block, and are categorized into three types. A movement is *correct* if the block was moved closer to the correct solution on both x and y axes. A movement is *half-wrong* (or *half-correct*) if the block ended up being closer to the solution only on one dimension. Finally, a movement is *wrong* when the block was moved in the wrong direction along both dimensions. The relative amount of each of these categories of movements is shown in Figure 6.14. As can be observed in the Figure, the ratio of correct movements was significantly higher in the coupling condition than in the no-coupling condition (F[1,27]=26.5, p=.000). The difference between the two conditions was also significantly different for the half-wrong movements (F[1,27]=9.08, p=.005), but not for the wrong movements (F[1,27]=.93, p=.34).

Furthermore, the diversity of the paths among the groups was different across the two conditions. As can be observed in Figure 6.15, the paths chosen by students in the coupling condition were similar to each other. The purple circle indicates the correct solution. The full circles indicate the starting positions, and the squares the arrival points. The color of the square shows whether their solution was correct (green) or wrong (red). Each point represents a time lapse of 1 second. Red segments show wrong path; segments are orange when one of the two directions was wrong; the green segments correspond to correct moves.

In the no-coupling condition, the groups chose very different paths and often ended at an incorrect position. The strategies adopted also varied. In the no-coupling condition, students who found the correct position did so by decomposing the movement first on the y-axis, and then on the x-axis. In the coupling condition, there were two successful strategies: decomposing on both axes, or going straight in a diagonal line. The latter is the optimal



(a) Coupling.

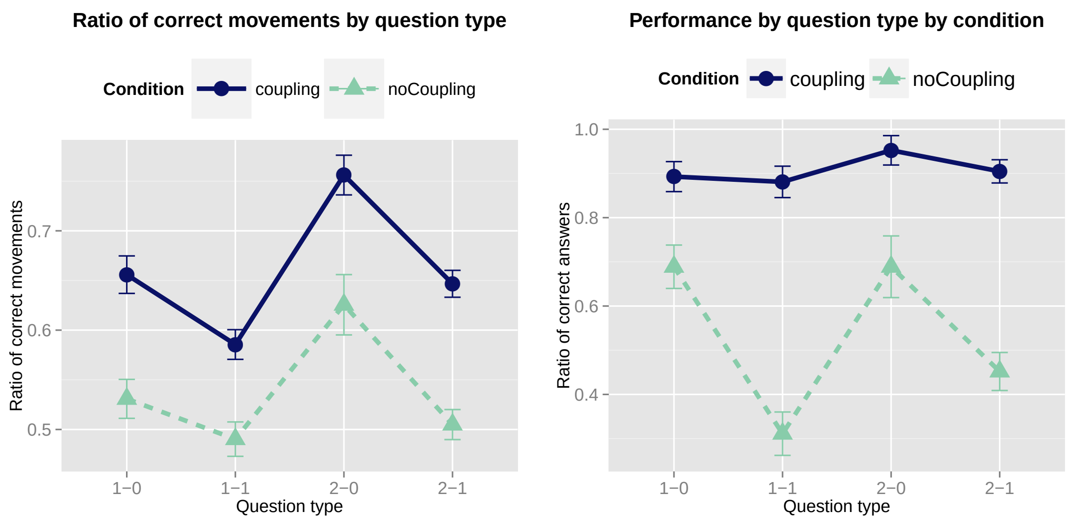


(b) No-coupling.

Figure 6.15: Detail of the path for one question in each condition.

strategy, but also the most difficult one, because it implies moving along the two axes at the same time and therefore controlling two factors simultaneously.

Variations among question types. The difference in the path pattern was more distinct for some types of questions. Grouping the questions by the number of translations (first number) and whether or not a rotation was involved (second number, 0 means no rotation, 1 means there was a rotation) helps highlight that. Figure 6.16a shows the average ratio of correct movements in the path chosen during the experimental task for four categories of questions. Figure 6.16b shows the percentage of correct answers given for the same four categories of questions. The coding scheme for the type of questions is <translation>-<rotation>. For translations, the number indicates along how many axes the translation occurred; for the rotation, 1 means there was a rotation, 0 means there was not. So for example, 2-0 means that there were translations on both x and y but that there was no rotation.



(a) Ratio of correct movement by type of questions and by condition.

(b) Ratio of correct answers (during the experimental task), by type of questions and by condition.

Figure 6.16: Details of block movements and performance by type of questions.

Regarding the path, there was no difference between the question types: students in the coupling condition had a consistently higher ratio of correct movements across all question types. On the other hand, the ratio of correct answers per question types (Figure 6.16b) shows that there was an interaction effect between the condition and the question type, indicating that some questions were harder to solve in one condition. In the coupling condition, all questions had a similarly high score, but in the no-coupling condition, the performance was significantly lower on questions involving a rotation than on those without one ($F[1,22]=8.5$, $p=.008$). This is not really surprising. Indeed, a question without rotation can be solved by simply decomposing between the x and y directions, whereas when a rotation is added, one must figure out not only the new position of the object, but also its new orientation. Finding the new orientation is a less constrained problem than performing a decomposition of a movement on two axes. This is confirmed by the observation that the number of translations did not affect the result.

6.3.4 Discussion

The type of feedback provided in the coupling condition was at the process level, but it had an impact on students' behavior and learning both at the process and the task levels.

Differences at the process level

At the process level, the path analysis revealed behavioral differences between the two conditions. Students in the coupling condition manipulated more and did not wait as long as no-coupling students before performing actions. Additionally, once they started moving the block, they moved it slower than no-coupling students. Since there is no reason to think that students in the coupling condition were naturally more prone to action, this suggests that the dyna-link between the tangible blocks and the virtual representations encouraged students to rapidly dive into action without much thinking beforehand, as if proceeding in a trial-and-error fashion. Students in the coupling condition should not be blamed for their lack of reflection: they were told to move the block to a new position using the shortest possible path and quickly figured out that they could use the feedback to infer the correct direction.

On the other hand, students in the no-coupling condition reflected more before acting. They also ended up with many different solutions, as indicated by the heterogeneity of the paths. This different solutions could be used in the future to generate a confrontation between the groups, which could improve learning (Doise et al., 1991). The path patterns in the no-coupling condition showed a higher rate of half-wrong segments, hinting that the students in this condition decomposed the movement into sequential movements along the two axes.

Finally, the effect of displaying the score was higher than expected. Students were serious about trying not to waste points. This was reflected by small initial movements in the coupling condition and by a longer thinking time in the no-coupling condition. Including the score as a performance metric was a deliberate design choice for this experiment, with the goal of restraining the coupling students. The same experiment without the score would probably have led to different results.

Differences at the task level

The differences at the process level between the two conditions were reflected at the task level in the learning outcome. In the no-coupling condition, in which students reflected more and manipulated less, the students' performance increased over time. Despite the lack of difference in the learning gain measured by the tests, this suggests that the higher level of manipulation involved in the coupling condition was not beneficial for learning. This is in agreement with Price et al. (2009) and Do-Lenh (2012) who argued that a slower manipulation could foster learning. It also confirms that simply manipulating an object does not automatically lead to learning, as a simplistic interpretation of the embodied cognition theory could lead to believe.

In both conditions, the type of answers given indicated that it was easier for students to find the right position on the x-axis than on the y-axis. This is in agreement with the literature on spatial skills that showed that the greater the mental rotation, the harder the representation (e.g Flusberg and Boroditsky, 2011). The movement of the y-axis appear on the side view and require a rotation of 90° to be interpreted correctly. There was a difference between the two conditions for questions involving a rotation, which were answered correctly more often by students in the coupling condition. The difference in the answers given between the pre-test and the post-test also showed that the no-coupling students improved more evenly than coupling students, who mostly improved at identifying the correct position on the x-axis and the correct orientation, but not the correct position on the y-axis.

This last result suggests that both conditions can confer advantages. For example, the coupling condition could be used for novice students who have none or little knowledge of orthographic projections. Using TapaCarp in the coupling condition would allow them to explore the relationships between the multiple representations of the block and to create their own set of rules, through induction, on how these representations are linked. It would also allow to break potential misconceptions (such as "the side view is seen from the right of the object"). Compared to their current drawing exercises, this type of exercises would be faster and provide a larger sample of cases from which to build rules. After some time, the coupling could be removed to force students to reflect more.

Summary

The purpose of this study was to examine the impact on students' behavior and learning when varying the type of feedback given to the user. The participants were split into two conditions, one with dyna-linking between the tangible and the virtual representations, and the second without. There was no significant differences between the two conditions in terms of learning gain, despite significant differences in strategies used to solve the activities. Indeed, the dyna-linking led to a higher degree of manipulation and a lower degree of reflection. It made the task easier, and students in this condition solved more questions correctly during the experimental task. However, their performance did not improve with time, as opposed to that of students who did not have the dyna-linking and whose performance significantly improved over time.

6.4 Discussion and Conclusions

The two studies presented in this chapter had a similar design but pursued different goals. I have discussed the specific results of each study above, and I will now discuss the limitations of these studies as well as the meaning of their results from a broader perspective.

6.4.1 Limitations

Both studies suffer from several limitations. One of them is the relatively small sample of participants. While a body of 50 participants is reasonable to observe some trends, a larger and more varied one (e.g. participants coming from different schools) would lead to more solid conclusions.

The semi-controlled setup of the studies can also be criticized: there was noise around the participants and distraction from the other apprentices going around, and they had to interrupt their other tasks to take part in the experiment. A more controlled environment, such as a laboratory, would have helped control these external factors, but would have been more distant from the target usage conditions of TapaCarp. As explained in Section 5.2, one of the principles of design-based research is to run studies in settings that are as close as possible to ecologically valid ones.

Another limitation of these studies is the type of activities that were considered. Indeed, while the tasks given to the participants are relevant for carpentry training, one cannot expect to learn carpentry by only doing small exercises lasting for a couple of minutes each. The activities targeted a precise skill in order to be able to measure the potential learning gain, and the duration of the activities was constrained by the environment: 1 hour of experiment (including the explanations and the tests) is a substantial amount of time for carpenter apprentices who have less than 4 weekly hours dedicated to carpentry classes.

A third limitation is the fact that apprentices completed the activities in pairs. The effects could have been larger if the study had been done individually. However, this choice in the design of the activity was made for two reasons: the apprentices usually work in pairs, and tabletop tangibles are often designed for a collaborative usage, and so is TapaCarp. Although we did not study the collaboration between the members of the pair, a future area of research could be to study the collaboration within pairs, especially the emergence of roles and how this affects the learning performance.

The last limitation concerns the measure of the learning gain. Testing the participants before and after the experimental task is a widely used method, but nonetheless one with limitations. The learning gain, as measured by the tests, highly depends on the design of the test, and designing good tests is an art in itself. Moreover, in the absence of a control group, it is not possible to distinguish the effect of learning from other effects, such as a difference in difficulty between the two tests, or the students getting used to test questions. To limit the risk of the latter, the task chosen for each study (and for the tests) was similar to another task of the regular curriculum.

6.4.2 Implications for the design of TUIs for learning

The two studies presented in this chapter answer the first research question presented in Section 5.5. Their results emphasize the impact that design decisions for TUIs may have on improving a particular skill. In the first study, we focused on the shape of the tangible object and saw that changing the physical correspondence between the tangible object and its virtual representations led to significant differences in the participants' performance during the experimental task, but did not yield a significant differences in terms of learning gain. The second study showed that the presence or absence of physical-virtual coupling (dyna-linking) can influence the difficulty of the task and lead to important differences in students' behavior and learning performance.

There are many design dimensions that have to be taken into consideration when designing a TUI for learning purposes. Some of them are domain-specific, whereas others are applicable to several learning domains. The first study shed light on the importance of the physical representation of the tangible interface in the context of spatial skills. Of course, the importance of the physical representation is tightly linked to the subject of spatial skills. In other contexts, other design dimensions might be of more importance. For example, in a TUI designed to learn about construction materials, the mass and the texture of the objects might be more meaningful dimensions to study than the object representation. My message is therefore not that every virtual object should be represented with a tangible of literal physical correspondence, but rather that for each learning domain, designers of TUIs reflect on what the key pedagogical aspects of the learning domain that they want to tackle are, and tailor the design of their interface to emphasize these aspects.

The result on feedback is more directly applicable to other domains. Because of its higher usability, physical-virtual coupling is often the default in TUI environments, possibly at the expense of learning. This coupling is not bad per se, and could even be of great help when students are exploring an unknown topic by helping them build their knowledge through induction. For example, in the beginning of their first year, many carpenter apprentices are completely new to orthographic projections and have difficulty making sense of them. Using TapaCarp with coupling could help them explore the link between the three views and the physical object. Depending on the stage in the learning process and the expertise of the students, the feedback could then be varied to foster learning.

For instance, an interesting future direction would be to explore how to use different types of feedback in the same activity, but with a different timing. Our results show that immediate feedback makes the task easier, and it could therefore be useful to provide it when exploring a new subject to lower the entry-level threshold. Because it does not promote reflection, this type of immediate feedback could be gradually diminished as students acquire expertise. TUIs could therefore be used to lower the barrier to entry to new domains, as already underlined by Shaer and Hornecker (2009). Orthographic projections are hard for novices, and the "easy" conditions ('block' and 'coupling') helped students complete activities correctly, in both studies. With well-designed activities of gradually increasing difficulty, tangibles could allow students to explore a new domain that would otherwise be too hard for them, and construct their own understanding of the new domain before being taught its theoretical foundations by the teacher.

These two studies contribute to the debate on TUIs and their potential for learning, by emphasizing the importance of the relationship between design aspects of the TUI and the learning outcomes that its users can gain. They are however only two small pieces in the puzzle that depicts the relationship between TUIs and learning, and much has still to be learned as to how TUIs can be used efficiently for learning, both in general and for specific learning domains.

7 Integrating TapaCarp in Classrooms

In the previous chapter, we saw that TapaCarp can be useful to train spatial skills. Although the system was tested in the classroom, it was not integrated in the lesson taught by the teacher. Rather, students used it on the side, in parallel to their other learning activities. This is not necessarily a problem: it is possible to envision the use of TapaCarp as a station where students go from time to time to do a set of exercises. Nevertheless, it is also interesting to study ways in which TapaCarp could be integrated in the classroom. This is the focus of this chapter, in which two ways to integrate TapaCarp in the classroom are presented.

The first one describes an attempt to integrate TapaCarp into existing pedagogical practices. The learning environment already exists and has many constraints. The constraints come from many sources, such as the curriculum, the habits of the teachers, the tradition of the profession, the limited amount of time that students get in the classroom, and all the physical constraints related to the classroom itself. To have a chance to survive in the learning environment, a learning technology must adapt to its constraints and mold into the existing environment. As the results will show, this is easier said than done.

A novel learning technology, such as TapaCarp, can also uncover new ways of teaching. Even if such ways cannot necessarily be used immediately, it is still worth thinking about what they could be if structural constraints were to shift. Showcasing these new ways could even help change those structural constraints. For instance, car designers went past the absence of paved road and gas stations (a structural constraint) when they designed the first motorized cars, and the structural constraint gradually lifted as paved roads and gas stations were built, allowing motorized cars to become mainstream. I am not claiming that TapaCarp will become mainstream in the future, but simply that it is worth thinking past current constraints. I will therefore describe, in the second part of this chapter, how TapaCarp can open up new perspectives on how to teach carpentry concepts.

The research question explored in this chapter is the second one presented in Chapter 5: what classroom pedagogical scenarios can TapaCarp support for the training of spatial skills. We will address it by splitting it into two parts, the first part considering the integration of TapaCarp in existing practices, and the second part in new practices.

7.1 Integration Into Existing Pedagogical Practices

As explained on page 59, the TapaCarp interface was first composed of the blocks, the cards, and a mouse, and the learning activities running on it were short activities, such as the two described in the previous chapter. After testing them with apprentices in a classroom, our next intent was to integrate TapaCarp in a regular teaching lesson. However, although he had participated in the design and testing of TapaCarp and the activities, the teacher did not want to use them "for real", i.e. as genuine classroom activities, because the activities were not part of the regular curriculum. This came as a surprise, since TapaCarp activities and the ones done routinely in the classroom had the same goal (improving the 2D-3D link), an observation that the teacher did not refute. After some more inquiry, we discovered that the major issue was that the professional and school environments of a carpenter are deeply embedded with paper and drawing, and that TapaCarp used neither paper nor drawing tools. Drawing-based practices are the "DNA" of these classrooms, and the absence of this "DNA" part in TapaCarp led to its rejection.

To the three components of the interface (blocks, cards, mouse) used in the previous experiments, we therefore added a new one: a paper activity booklet. This allowed students to draw on paper. They drew with their regular drawing tools, which further satisfied the teacher — learning to use these tools properly is a curriculum requirement. New activities were designed to make use of the paper and drawing tools. In the end, except for the presence of the block and the possibility to augment the paper with the projector, the activities were very similar to the ones done in the regular curriculum (see Video 1).

The new TapaCarp interface is now distributed over several components: the blocks, the cards, the mouse, the booklet, and the set of tools (see Figures 5.3, 5.4, and 5.5 for a reminder). While it met the teacher's requirements, introducing the activity booklet had a side effect: the use of TapaCarp became closer to the usual classroom pedagogical structure, and more scripted. Each activity is now designed as a step-by-step process through which the students are guided by a mix of printed instructions and dynamical instructions projected by the system. It is this final system that was tested in a classroom environment.

For the development of the new interface, I received help from Stéphane Testuz, Sara Alsudairy, and Yongsung Kim. Stéphane also helped in running the study described in the first part of this chapter as part of his Masters' thesis, as did my adviser, Pierre Dillenbourg.

Research questions

There were two distinct goals with this experiment: assessing the usability of TapaCarp, and measuring the impact of TapaCarp on the learning performance. As explained, many changes were made to the TapaCarp interface. In particular, the activity booklet and the drawing tools were added to the existing blocks and cards. This resulted in an interface made of many components, and thus arguably more complicated to use, and in a new interaction model, with the booklet activity (and drawing) becoming a central part of the activities. The first goal of this experiment was therefore to see if the system was usable by apprentices, in a classroom

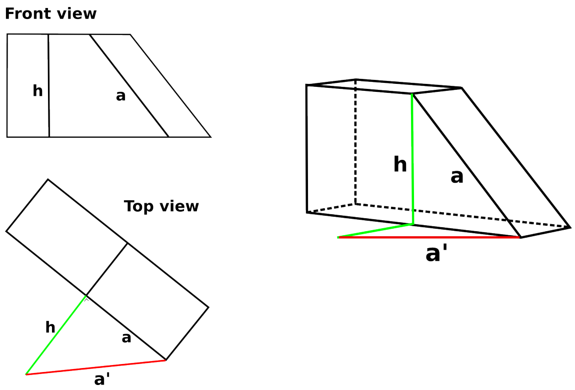


Figure 7.1: The rabattement technique.

context.

The second goal of this experiment was to assess whether TapaCarp improves learning. There were several reasons why we thought TapaCarp could help students learn better: a readily available help might encourage the students to be more autonomous and find the solution on their own; the integrative approach (drawing-object link) could help students understand better the rabattement technique (see below) by making it more concrete and tangible; finally, a step-by-step animation of the rabattement process was introduced with the hope that it could help students understand better, as has been shown in previously (e.g. García et al., 2007);

7.1.1 Experiment and method

Learning task: rabattement

The topic of the learning activity designed for the experiment was suggested by the teacher. The goal of the 90 minute lesson was to teach the apprentices how to find the true size of an object from its orthographic projections. This is one of the key tasks in the 2D-3D passage.

There are various techniques to find the true size of an object from its orthographic projections, but carpenters mainly use the rabattement technique, introduced by Monge (1798) and graphically explained in Figure 7.1. The core idea is that the true size of `a` is neither available on the front view, nor on the top view. To find it, one measures the height `h` on the front view, and reports it perpendicularly to the projection of `a` on the top view. The line binding the free endpoint of `h` and the other side of the projection of `a`, i.e. `a'`, is the true size of `a`. This corresponds to pivoting a vertical triangle until it is horizontal, as shown on the right side of Figure 7.1. Although it may seem easy on such a simple example, applying this technique in other contexts can soon become complicated.

Activities

Once again, apprentices used TapaCarp in pairs. The activities were designed together with the teacher. There were 11 activities presented in an increasing level of difficulty and grouped in three parts. All of them included exercises around the notion of rabattement: the first part (activities 1-3) was an introduction to the principle of rabattement; the second part (activities 4-6) dealt with finding the true size of an edge; the third one exercised finding the true size of a face (activities 7-11). The full booklet activity is shown in Section A.5. The blocks, the cards, the paper tools, and the activity booklet were provided to the apprentices. They were also asked to use their own regular tools: pencils, a ruler, a protractor, an eraser, and a compass.

Interface

There were 7 flippable cards and 3 plain cards (see Section 5.3.3 for a reminder). The four first flippable cards are “next step”, “previous step”, “help”, and “check solution”. “next” and “previous step” are used to navigate within an activity. The “help” card gives the apprentices step-by-step instructions to perform a rabattement and checks the correctness of their actions at each step. It therefore lightens the orchestration load on the teacher by giving a first help level to the students, but is only available in four of the 11 activities (5, 6, 8, 9), to encourage independence from the learner. The “check solution” card verifies the current solution and gives feedback accordingly. The three other flippable cards manage the animation: the first one launches the animation of the rabattement; the second one launches the same animation but with step-by-step textual explanations; and a third one quits the animation mode. The plain cards are “show axes”, “construction lines”, and “freeze”, all of which were described in Section 5.3.3.

None of the activities required the mouse, which was hence not included for the study. In total, six blocks were given to each pair. Each activity made use of one block. All the material was given to the apprentices at the very beginning of the class, except for three cards managing the animated feedback: these were given to them after they completed activities 3 and 4.

Participants and procedure

Three classes of second year apprentices took part in the study, in a classroom in a vocational school. Each class was split in two conditions (see below for more details on the conditions): half of the students attended the normal class with the teacher (paper condition), while the other half used TapaCarp under the supervision of one researcher acting as the teacher (TapaCarp condition). The 43 participants were assigned randomly to one of the two conditions. There was a total of 24 apprentices in the TapaCarp condition, and 19 in the paper one. All apprentices completed a pre-test before the experiment and a post-test after they finished the activities. Both the pre-test and the post-test assessed the understanding of the rabattement concept. Following the post-test, apprentices filled a questionnaire that asked about the perceived usefulness of the course (all apprentices) and about the usability and usefulness of TapaCarp (TapaCarp condition only). The tests and the questionnaire are available in Section A.5.

One frequent concern with learning technologies is how fast students learn to use them. Participants in the experimental condition were shown a short demonstration of TapaCarp (less than 5 minutes). To reduce the novelty effects that could distract them from the activities, the apprentices were allowed to play with the system for as long as they wanted before starting the activities. They typically tried the system for 2-3 minutes before starting the activities, and completed the activities for the remaining of the 90 allotted minutes.

Conditions

There were two conditions: TapaCarp and paper. Students in the paper condition completed the same activities as those in the TapaCarp one, except without TapaCarp. To compensate for the lack of animation, they received a sheet of paper (shown in Section A.5.5) that described the same steps as the animation given to the experimental group. They were also given the same tangibles as the participants in the control group, except without the fiducial marker. This was done so that the two conditions would be more comparable. At the same time, it means that the paper condition was further from the traditional classroom activities, where apprentices usually do not use tangibles in their drawing activities.

Data collection

The data used for the analysis of the results were collected through the log files of the application, video recordings, and the questionnaire given to the students at the end of the experiment. The three classes had an uneven number of students, and therefore 3 students (in the paper condition) had to complete the activity on their own. For the sake of consistency, these students were removed from the data.

Precision on some experimental design choices

Some of the experimental design choices might seem surprising. They are the results of running a study in a school, with real students, which exposed us to several unintended consequences.

The first and main one was that we would have liked the teacher, and not a researcher, to play the role of the teacher in the TapaCarp condition. However, practical constraints made it impossible: someone had to take care of the students in the control group, and the teacher decided to teach without TapaCarp rather than with it. To reduce the teacher effect, the learning material was designed to be mostly self-instructional, and the teacher was told not to give a lecture, but to simply answer questions, if there were some, but without giving away the answer. The hope was that this would limit the teacher effect, but this was not an ideal design.

Another limitation of the study was that participants were supposed to not have studied rabattement prior to the experiment. However, at the time that the study was run, it turned out that all of them already had had some prior exposure to the topic of rabattement.

The apprentices usually do not use tangibles when learning the rabattement technique. In this case they did receive tangibles. This was done so that the difference between the two conditions would be the use of TapaCarp, and not the use of tangibles. Ideally, with more participants and time, we would have added a third condition with just the traditional learning material.

7.1.2 Results

Learning gain

The performance improved between the pre-test and the post-test (33.4% to 43.2%, t[39]=2.68, p=.01). However, students in the *paper* condition improved more than those in the TapaCarp condition (33.3% to 51.5%, versus 33.5% to 37.7%), although the difference between the two conditions was only marginally significant (F[1,38]=3.7, p=.06).

This difference in the performance is rather disappointing, because it could mean that TapaCarp hampered learning. Although this could be the case, another more likely explanation is the teacher effect. During the experiment, we could not monitor the teacher at every moment. One researcher went to his classroom from time to time, and noticed that the teacher was not respecting the “only answer questions” instruction. Instead, he was taking advantage of the smaller class to spend more time tutoring groups of students. Furthermore, in one session, one of his colleagues who wanted to observe the experiment came in the classroom and spent more than half an hour doing one-on-one tutoring with one of the apprentices. Why the teacher did this is unknown. It is possible that his teacher instinct took over, that he forgot about the instructions, or that he wanted to show that he could teach better than a machine. Whatever the reasons, this must be kept in mind when analyzing the results.

Both group had a 100% completion rate up to activity 6. Then for activity 7 to 11, students in the paper condition finished 86.3% of the activities to compare with 75.8% for those in the TapaCarp condition. Surprisingly, though, the percentage of activities completed correctly was similar in both conditions (71.6% for paper versus 71.2% for TapaCarp). However, most of the mistakes in the paper condition happened in the first two questions, because of a misinterpretation of the instructions. Considering the remaining activities only, the percentage of questions answered correctly is higher in the paper condition (82.0) than in the TapaCarp condition (70.4). These numbers, combined with the teacher’s behavior described above, hint that students in the paper condition received more effective help than those in the TapaCarp condition, and this could explain the differences observed in the post-test.

Other minor factors may have played against students in the TapaCarp condition. For example, the tests were closer to the treatment for students in the paper condition, because students in the TapaCarp conditions passed the test on paper and without TapaCarp. The students in the TapaCarp condition also had slightly less time to do the activities because they needed to get used to the usage of TapaCarp. Thus, of the 19 groups in the paper condition, 15 were able to finish all the activities, to compare with only 8 in the TapaCarp condition.

Users feedback

We gathered both formal (written questionnaire) and informal (class discussion) feedback from the users. The questionnaire included 13 assertions to be assessed on a seven-point Likert scale (-3 to +3), and 5 more open questions (see Section A.5.3 for more details). A large majority of apprentices (18) had positive opinions on TapaCarp, 2 were neutral and 4 were negative. From the 4 negative ones, most of the criticisms came from the lack of accuracy of the projection due to some calibration inaccuracies between the camera and the projector.

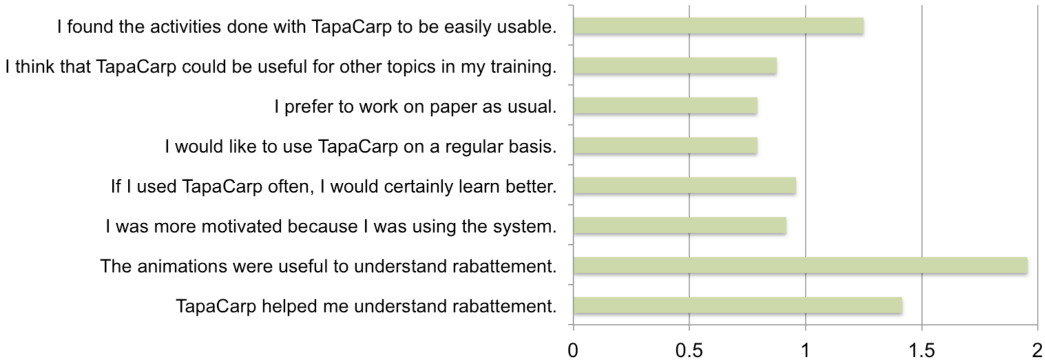


Figure 7.2: Mean scores of the questionnaire answers, on a seven-point Likert scale (-3 to +3).

The mean scores of the answers to the TapaCarp related questions are shown in Figure 7.2. The participants were generally enthusiastic about the system, both in terms of perceived usefulness of TapaCarp for their training as well as in terms of its usability. For instance, they were interested in using TapaCarp more often (0.96 on the Likert scale) and said TapaCarp helped them understand the rabattement better (1.25). Only three apprentices said TapaCarp did not improve their understanding of the rabattement, out of which two said that they had already understood it beforehand. The animation was deemed especially useful to better understand the rabattement technique (1.95).

Left-right differences

The manipulation of objects by the two users reflects the asymmetry and the modularity of the interface. Figure 7.3 shows the number of activities in which the user on the left or right of the workspace performed an action or manipulated an object. There was a significant difference in the usage of the modalities: the participant to the right manipulated the blocks more, while the left participant used the cards more. Part of this asymmetry can be explained by the physical placement of the manipulation zone of the blocks (on the right). However, the cards could be used anywhere, so the fact that they were used mostly by the left participant is more surprising. These differences in the usage of the interface did not lead to different learning outcomes ($F[1,22]=3.54$, $p=.07$), although in 9 groups out of 12, the post-test score of the apprentice sitting on the left was higher than his colleague's.

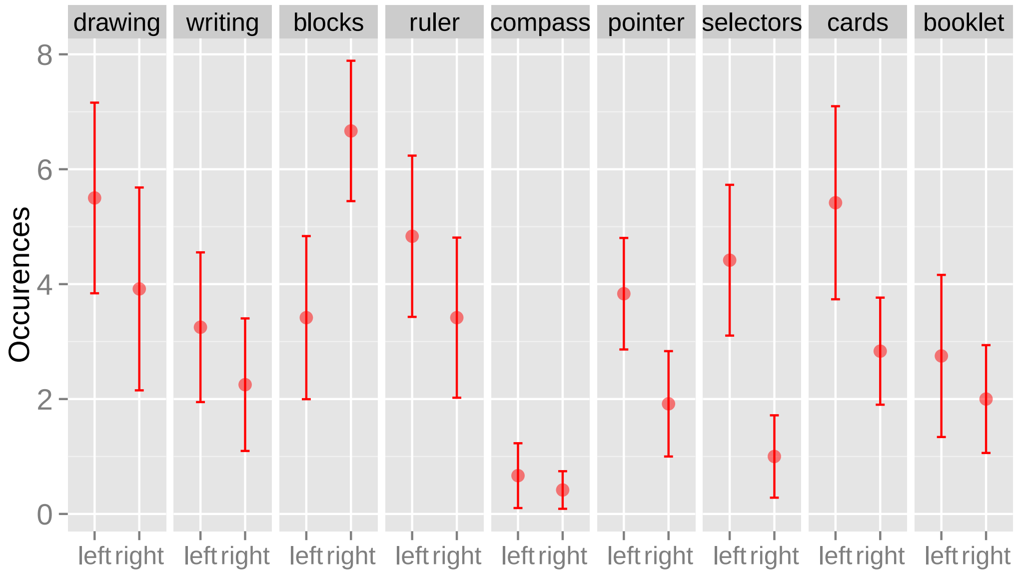
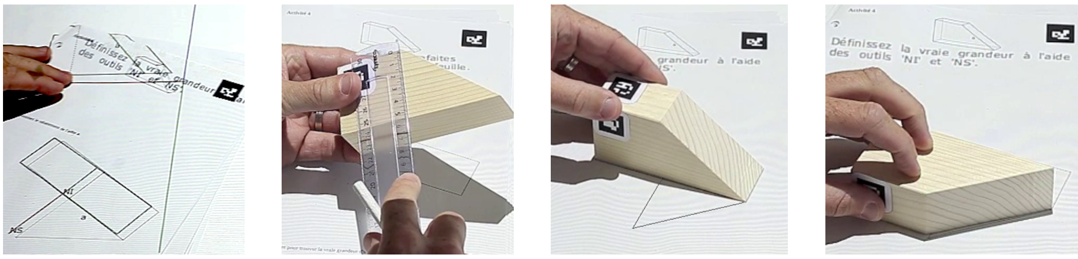


Figure 7.3: Average number of activities (with standard errors) in which each student on the left or right performed an action or manipulated the objects at least once.

Usage of the cards

In total the "freeze" card was used 22 times; the "show axes" 11; the "construction lines", 18. Every group used at least one of these cards. Two groups did not use the "freeze" card, four groups did not use the "show axes" card, and four groups did not use the "construction lines" card. The animation without the step-by-step explanations was watched 16 times, and the one with the step-by-step explanations, 11 times. This does not include uses that lasted for less than 4 seconds, which were numerous (22 in total). The reason for this is that apprentices expected the animation to be shown on the current block. Instead, in order to encourage the transfer of the learning, the animation was always displayed on a block that was not one of the six given to the apprentices. Still, 10 of the 12 groups played one animation to the end at least once. In many cases apprentices completed the drawing exercise using the step-by-step animation. There were a total of 24 uses of the help card out of 48 potential ones. Only one group used the help in all cases. Activity 6 was by far the one with the most help requests (10), most likely because it introduced a new difficulty (rabattement of a non-vertical edge).

Using the help was usually effective. Out of the 24 uses, only 2 uses resulted in a wrong drawn answer from the students. The help worked as expected in the first case, but the group did not manage to complete the step-by-step process completely. They did not call the teacher either, and just completed (wrongly) the exercise. In the second case, the group was trying to use a different method than the one shown in the help and they did not understand what the help was suggesting. Besides these two cases, the apprentices showed high meta-cognitive skills: only in two cases did they complete an activity wrongly without using the help.



(a) Paper with projection. (b) Tool with block. (c) Block with paper. (d) Block with paper.

Figure 7.4: Using the various modalities of the interface together.

Blocks and their interaction with the paper and tools

None of the groups tried to use the wrong block for an activity, most likely because a perspective view of the block was printed at the top of each activity page. As explained in Section 5.3.3, the blocks serve both as manipulation handles and as an external representation. The apprentices' behavior showed that they understood this, and that they made the link between the 3D block, their drawing, and the projected representations. They used the blocks extensively to take measurements, check their solutions, or change the angle of the displayed projection (see Figure 7.4). They measured dimensions both on the block (Figure 7.4b) and on the orthographic projections, and laid the block on their drawing to check that the length of an edge that they found by rabattement actually matched the real length (Figures 7.4c and 7.4d). One could say that some of them even understood all too well how to use the blocks, since they sometimes used it not only to check their solutions but to find the solution by measuring the true size of an edge or face directly on the block instead of finding it by rabattement.

Usually, carpenter apprentices do not dispose of the physical model of what they are asked to draw on paper. Our activities forced them to link the actual block to the drawing. This link between the block and the drawing was done either directly by laying the block on the drawing, or indirectly by taking measurements on the block and reporting them on the drawing. Noteworthy is the fact that some groups did not make a direct link between the blocks and the paper and always used an additional tool – ruler or compass – to make this link. Others, on the contrary, used the block directly by laying it on the drawing.

Activity booklet

The activity booklet had a good "orchestrability": easy to distribute and gather, no ordering problems or loss of activities. The apprentices are used to receive exercises this way and did not question it. The navigation between the activities was not programmatically enforced, allowing students to browse through all the activities without any mandatory checkpoints. This resulted in some students not calling the teacher when the written instructions asked them to (so that the teacher could check their solution and give them more cards). A minor issue was that after completing several activities and flipping the pages, the stapled corner of the page was higher than the other ones, leading to a less reliable detection of the fiducial marker. This could easily be fixed by placing additional fiducial markers on the page.

Other observations

The augmentation of the paper with dynamic instructions and feedback generated a split attention effect. Some basic instructions had been printed on the paper to help students complete the sheet in a structured way. The dynamic instructions were projected on the top of the page. However, despite the flashing of a bright color the projected instructions were sometimes ignored. When they were stuck and asked for help, the instructor simply pointing to the instructions often solved the problem.

Each group received a total of ten cards and six blocks. These came in addition to the drawing tools and the activity booklet, and resulted in a large number of objects to manage on the tabletop. It was sometimes complicated for the students to find the card they were looking for; in some instances, a card was activated by mistake. Although students did not complain about that in their feedback, this appears to be a usability issue. From the orchestration point of view, the instructor had difficulties distinguishing what activity the students were working on from a distance, because the table was so cluttered with objects.

7.1.3 Discussion

Limitations

As already mentioned, the ecological validity of this study was compromised by some parameters that we could not control. First, a researcher had to play the role of the teacher, while the teacher was teaching to the control group. Second, the apprentices had already been exposed to the learning topic, which was not planned. The former turned out to have another consequence on the results, because students in the paper condition received more help from their instructor than those in the TapaCarp one. The teacher effect might even have been increased by the unbalanced number of participants in each condition (24 with TapaCarp versus 19 with paper) with an average of 6.3 students per class (including the 3 students excluded because they worked alone).

Having students in both conditions do the tests on paper may have given an advantage to the paper condition, although such an advantage is hard to quantify. It is also difficult to imagine another solution that would not introduce a bias between the two conditions.

Another limitation concerns how realistic it is to have one TapaCarp system for every two students in a classroom. Carpenter classes are usually rather small (around 15 students), but even with this limited number of students, 7 or 8 systems would be needed. Not even mentioning the cost of such an installation, setting up the classroom as we did for the study would require a large effort from the teacher's part.

Positive outcomes

In terms of usability, the results were globally positive. With just a few minutes of introduction, the students were able to use TapaCarp to complete complex activities. Their feedback, although it must be taken with a grain of salt in light of the novelty effect, was positive both on the perceived usefulness and on the global usability of TapaCarp. The distributed interface allowed it to mold into the classroom ecosystem by using some of the media traditionally used in the classroom (paper and drawing tools). TapaCarp also allowed the students to work as usual on a tabletop and to keep their habit of working in pairs. In fact, I interpret the positive response from students to TapaCarp as the result of this molding into their habits and environment. Students who already mastered the subject of finding the true size of an object by rabattement could do the exercises without using the system much. On the other hand, those who were not so sure of how to proceed found help in the scaffolding provided by TapaCarp through animations and step-by-step help.

Another positive aspect of the distributed interface was that it fostered the emergence of roles within the pairs. Typically, while one student manipulated the cards, the other manipulated the blocks. On several occasions, one student prompted his partner with phrases such as “OK, we’re good, now you can use the card to go to the next step”. Role distribution in collaborative learning can be beneficial to learning (Burton et al., 1997) and has the advantage of engaging both students in the pair. In the case of tangible interfaces, where the “manipulation temptation” has been shown to be counter-productive to learning (Do-Lenh, 2012), paying attention and controlling who is manipulating what and when is especially important.

Things to improve

From its original orderly distributed configuration, the TapaCarp interface became “over-distributed” with its components being spread out all over the place without much order. This can be observed in Figure 7.5: on the right side of the workspace there is a ruler and a setsquare. On the left side, one can distinguish another setsquare, the six blocks, and an eraser. Close to the students’ arms are the cards, the compass, as well as some pencils and pens. With all the objects added across the four modalities, the final interface represented a total of 20 objects. We analyze the impact of the distributed interface on the usability by using the three levels of usability linked to the classroom orchestration theory (Dillenbourg et al., 2011).

- **Individual level**: as mentioned earlier, there were no major usability issues at the individual level.
- **Group level**: we observed that the large number of objects led to some usability issues at the group level, although the students did not explicitly complain about it. The issues arise mainly from the cards, which students either unintentionally activated or, in some case, had trouble locating. For example, the “next” card was unintentionally pushed under the projection surface – and therefore activated – which led to a rapid “completion” of the activity that was not planned by the students. On several occasions, the students wanted to ask the system for help and had to look for the “help” card because it was hidden under the activity booklet or another card, or behind a block.

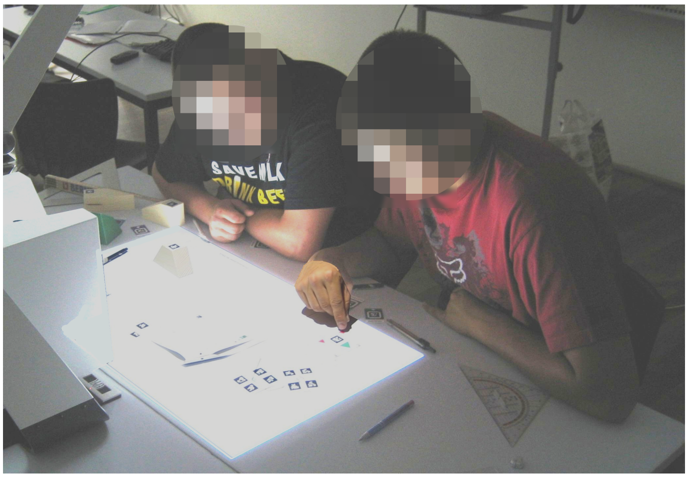


Figure 7.5: A group of students working with TapaCarp.

- **Class level**: the number of objects reduced the teacher's awareness of learners' work anbd progress by making it difficult for him to see what group was working on what activities from a distance. This increases the teacher's orchestration load.

"Hutchins threshold"

While it had some positive impact, the distribution of the interface also led to some usability issues. Two decades ago, Hutchins (1995) exposed how the distribution of information could help a cockpit remember its speed. He provided an explanation of why paper was surprisingly useful in an aircraft cockpit, which is a supposedly high-tech environment. According to him, the two pilots and the various artifacts spread over the physical cockpit environment form a distributed system in which information flows across different media. A classroom is a more diverse environment than a cockpit, but Hutchins' analysis is nonetheless relevant for the classroom. Similarly to Hutchins' approach, we developed a learning environment in which the interface is distributed across several modalities and media. Hutchins assessed the distribution of information as positive. In our study, we witnessed that the distribution of the interface has potential benefits, but also that over-distributing the interface can lead to reduce the usability on some of the three usability levels.

How many objects can a distributed interface have without hurting usability? In other words, could there be a point, which we could call the "Hutchins threshold", beyond which distribution of the interface could be harmful? In reality, Hutchins neither claimed that an interface should be distributed, nor how much it should be distributed. As a tribute to his work, we simply use his name to discriminate the point where the advantages of distributed interfaces may be counterbalanced by the shortcomings.

We do not have an answer to this question. This threshold is not simply a number of objects, it also depends on the characteristics of the objects: how much space they occupy on the interaction surface, how easily they can be stacked, sorted and put away, how often they are moved unintentionally (e.g. because they are too light), etc. It would take many carefully designed experiments to answer it in the case of TapaCarp. What we observed in the user study is that some distribution of the interface increased collaboration and decreased the orchestration load, but that a higher degree of distribution hurt the usability of TapaCarp and increased the orchestration load. While it is doubtful that there exists a general theory linking the number and type of objects in a distributed interface with the collaboration and orchestration loads, I believe that the degree of distribution of an interface is worth considering when designing a learning environment for the classroom.

For example, in our study, it is mainly the cards and the blocks that brought TapaCarp beyond the Hutchins threshold and caused the decreased usability. It may be that reducing the number of cards and blocks needed simultaneously would solve the usability issues. This can be achieved by placing only the objects needed for the current activity on the tabletop. The cards could, for example, be bundled with the corresponding activity sheet, while the blocks could be stored in a corner of the classroom where students would go pick up the block needed to their current activity.

7.1.4 Implications for future design

When creating a new learning technology, designers often have in mind usability at the individual and group levels. One reason is that it is easier to test usability at these levels than at the classroom level. Another is that the focus on the classroom is rather recent, and there is few existing guidelines on how to design to increase the usability at the classroom level (i.e. decrease the orchestration load), although some exist (see e.g. Dillenbourg and Evans, 2011; Dillenbourg and Jermann, 2010b). Based on the observation of this study and on others done with the Tinkerlamp by some of my colleagues (e.g. Bonnard et al., 2012; Do-Lenh, 2012), a set of 5 design principles emerged that help reduce the orchestration load of a tangible interface in a classroom. These 5 principles are: integration, empowerment, awareness, flexibility, and minimalism. The details about how we arrived to these principles, as well as examples on how they can be implemented in other tangible environments can be found in Cuendet et al. (2012a). Here, I simply present their definition in Table 7.1 and illustrate how they apply to TapaCarp below.

Integration. The integration of paper and drawing tools resulted from a curriculum and assessment constraint: carpenter apprentices need to master their drawing tools, and the teacher needs to evaluate how they use them. The usage of paper was decisive in the adoption of the technology by both the apprentices and the teacher. Apprentices simply solved the exercises on paper as usual, taking the augmentation as a nice add-on. The teacher saw that it did not change much in terms of preparation of exercises or explanations. TapaCarp also fit with the extra-classroom organization of the school that requires students who have not completed their drawing to finish them at home. Because they use the same means (paper) and tools as usual, these exercises can also be taken home. TapaCarp made use of wooden and

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 (Dillenbourg and Jermann, 2010b).
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.

Table 7.1: The 5 design principles.

cardboard blocks. The identification to wood and paper is strong in the carpenter community, and using those materials allowed TapaCarp to fit with the existing cultural environment. The shape of the blocks used was also similar to the objects drawn by apprentices in the regular class.

Flexibility. The affordance of paper lends great flexibility to quickly change activities, providing a better balance between routine and improvisation. The easy navigation between the activities (by turning the page) is an example of flexibility, making it easy for students to refer to a previous activity, or for the teacher to quickly access the history of what students have done. The distribution of cards and blocks opens up new possibilities of scaffolding. Indeed, it is possible for students to start off with a minimum amount of cards and blocks and gain additional ones as they progress in the activities. The activities worked well for students of various levels: the weaker students only completed the minimum drawing for each activity, while some stronger students drew the rabattement for all possible edges on the figure (even when they were not asked to). The help and the animation cards also helped the weaker students without preventing stronger students from going faster. The multi-component interface of TapaCarp fostered the emergence of role distribution between the two members of a group. For instance, in several groups, one student took control of the cards, while the other one controlled the block. This has the benefit of engaging both pair members and has been proven useful in terms of learning scenario (see e.g. Burton et al., 1997).

Empowerment. Giving the teacher special cards empowered him: on completion of two of the eleven activities, the apprentices had to ask the teacher to check their answer. If the answer was correct, the teacher gave them an additional card to play an animation. During the field study, we observed a tension between teacher empowerment and the flexibility principle. Indeed, some groups did not call the teacher when they should have and just went on with the next activity. They were able to do it because calling the teacher was not enforced for the sake of flexibility, but in this case, the higher flexibility removed power from the teacher. Teacher empowerment is also achieved in TapaCarp through the multi-component interface. The five types of components lower the barrier for teacher interventions, because the teacher does not have to take hold of the mouse or the keyboard and the large display makes it such that both apprentices and the teacher can interact with the system simultaneously.

Awareness. The awareness with TapaCarp is not optimal. During the design, the thought was that thanks to the tangible interface, the teacher could easily monitor the progress of each group. However, it turned out to be difficult for the teacher to estimate the progress of the groups from more than two meters away, mainly because of the small size of the blocks and the nature of the activity (a drawing is hard to observe from a distance). On the positive side, the teacher could nevertheless see which student was using what part of the interface and could therefore have an idea of who was playing what role in the group. The problem of following the students' progress could be partially fixed by centralizing the collections of blocks in one common place in the classroom, and the flow of students going to get the next block would give the teacher an indication on the students' progress. Another way could be to give the teacher a device that would communicate with the TapaCarp systems and indicate at what stage each group is.

Minimalism. The interface of TapaCarp is easily reduced to the minimal number of elements. In the field study, the apprentices often started the activity by pushing away what they did not need (for example the extra blocks and cards). The tangibility indeed helped sort out what was needed from what was not. The scaffolding of the cards is another illustration of the minimalism of TapaCarp: at the beginning the apprentices received only the strict minimum of cards needed to perform the first activities. As the activities went on they gained more cards that gave them access to additional features. Still, minimalism was not optimal in the experiment. The rather large number (more than 20) of objects sometimes caused problems, such as the loss of cards or the unwanted activation of a card. One solution could be to not only add cards and blocks, but also take some away, or design the activities so that they require fewer objects. The features of the system were kept to the minimum as well. For example, there is no artificial intelligence to correct the students work or give them suggestions on what they are doing wrong. There is no requirement to log into the system, and the only saving and history functions are the ones provided by the paper.

7.1.5 Summary and Conclusions

The focus of this first study was on the integration of TapaCarp in the classroom, using existing pedagogical activities. The teacher did not want to use the existing version of TapaCarp (with blocks, cards, and a mouse) because it was too far from the current pedagogical activities. Two components (a paper booklet and drawing tools) were therefore added to the interface and used to develop activities that the teacher would agree to use in the classroom. The outcome is a rather complex interface, distributed over several physical components. The two research questions for this study were concerned with the learning performance of students, and the usability of this complex system.

There was a marginally significant difference in the learning gain in favor of students in the paper condition. Several reasons were put forward to explain this, the main one being the teacher effect. Another way to see this is that the difference between the two conditions was help given from the teacher versus help given from a machine. This does not free TapaCarp from all blames, and it is for example possible that the step-by-step help was not reflective enough, but it is difficult for any technology to beat a motivated teacher.

The usability results were globally positive, although the many objects of the interface caused some problems at the group and classroom levels. We focused on the classroom level and presented 5 design principles to decrease the classroom orchestration load: integration, flexibility, empowerment, awareness, and minimalism. These principles were used to analyze where the interface of TapaCarp should be modified to improve classroom usability.

Distributed interfaces offer great potential for designing learning systems. However, over-distribution may hurt usability, and to this end we introduced the concept of "Hutchins threshold", a point at which the distribution of the interface becomes harmful. I do not claim that this threshold can be computed and will hold true invariant of the conditions. Rather, I see it as a design concept to keep in mind when designing tangible interfaces for learning in a classroom environment.

The integration of TapaCarp into existing practices and activities followed a request from the teacher and allowed TapaCarp to be deployed in the classroom. In this sense, it was a positive result. However, the activities that were done with TapaCarp were so close to the ones done in the regular curriculum that one can wonder if TapaCarp really brought a valuable contribution. If the activities are so similar, why could they not be done without TapaCarp? After all, students in the paper condition performed well, even if they received more help from the teacher. What is the added value of TapaCarp? Could it not be more valuable if new pedagogical scenarios were built around its unique features? These questions are explored in the next section.

7.2 Supporting a New Pedagogical Scenario

Integrating TapaCarp into existing pedagogical practices, as reported so far in this chapter, is needed to be able to deploy it in existing classroom environments rather than hypothetical ones. However, as our experience has shown, this integration also means that the freedom to design new learning activities is limited. Because the existing activities were not designed to take advantage of the capabilities of a system such as TapaCarp, merely integrating a system with the current pedagogical activities can limit the effectiveness of TapaCarp. This section reports on a classroom pedagogical scenario that was designed to involve TapaCarp from the start. The goal in designing this scenario was to use TapaCarp and its specific features to create a meaningful learning experience for carpenter apprentices. A further requirement was that the activity should be usable in the classroom, and consequently fit the classroom constraints.

7.2.1 Activity

Context and goal

A key element in carpentry is the plan. The plan is the only reference and communication means between all stakeholders involved in a construction. A plan can be global to a construction site, or specialized for a given trade. For carpenters, the plan is useful to know where and how to cut the beams at the workshop, and once on the construction site, to know how to put them together. It is therefore essential for carpenters to know how to read and – although to a lesser degree at the apprentices' level – to write plans. Reading and writing plans that can lead to the creation of an object is a difficult task that requires dedicated training, in the same way that reading and writing music score that can be transformed into music is hard and requires dedicated training.

Carpenter apprentices develop the skills of reading and writing plans in school, during the drawing classes. However, although they can seem similar because they use the same technique of descriptive geometry, the plans from the workspace and the drawing classes differ. The plans used in the drawing classes serve mainly as a technical exercise and a means of assessment for the teacher. The objects featured in these plans often do 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 geometry difficulty. Moreover, the plans done in the drawing classes are (usually) not used to build an object, as the focus is on the drawing of the plan itself. This is different from the workplace, where the plan is primarily a means to the creation of an object rather than an end in itself.

The goal of this activity is to make apprentices work on the plan as inspired by the use they make of it in the workplace: a means of communication and action. The activity includes the full workflow of carpentry, from the design of the object to the final cut out object. It does not require a full-fledged workshop, which is not available in a classroom. The activity navigates between the individual, the group, and the class levels of the classroom orchestration theory. It is decomposed in two phases: during the first one, apprentices work alternatively in pairs and individually; the second one is a class debriefing that involves all the students.

The first phase takes between 30 and 45 minutes. The collaboration between the two members of the pairs is designed according to the "Split Where Interaction Should Happen" (SWISH, Dillenbourg and Hong (2008)) pedagogical script. The idea of SWISH is to split a task between the members of a group of students, and thereby force students to interact with each other to complete the task. The learning results from the effort produced by the students to build a shared solution and understanding despite the initially different understanding, induced by the script, that they had of the problem. Of course, the split needs to be well designed, as

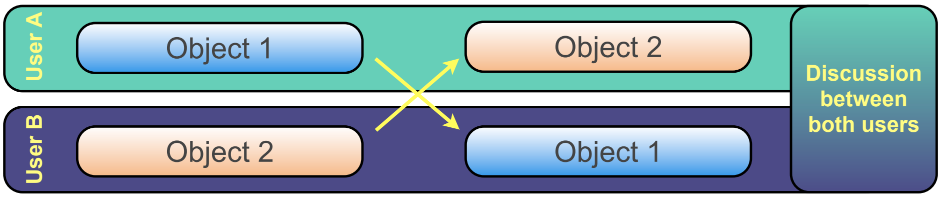


Figure 7.6: A graphical representation of the SWISH script.

increasing the collaborative effort too much may also damage collaboration.

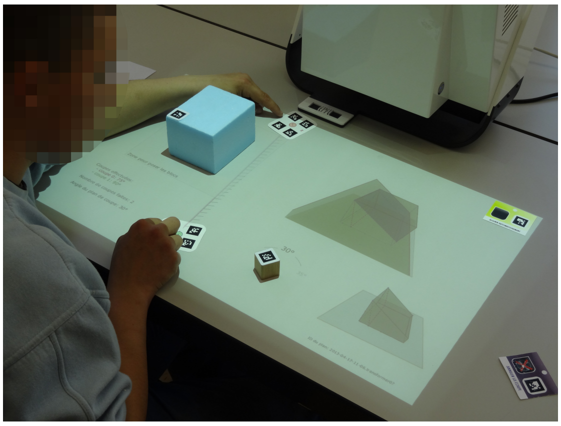
In the case of carpenters, there are two meaningful split options: the interpretation of a plan to mark out an object, and the interpretation of the marked out object to cut the object. Both are key activities for a carpenter and correspond to real-life scenarios. Indeed, in a company, the plan is made by a specialized carpenter. The other carpenters must then interpret the plan and mark out the wood to indicate how it must be cut. The cutting may or may not be done by the same carpenters and the information on how to cut the wood must therefore follow conventions and be clearly expressed in a non-ambiguous way. Since the focus of this experiment is on the plan, we chose to put the split on the transition from the plan to the marked out object. However, by only changing the position of the split to the marking out of the object, the focus of the activity could be changed to the marking out of the object.

Description

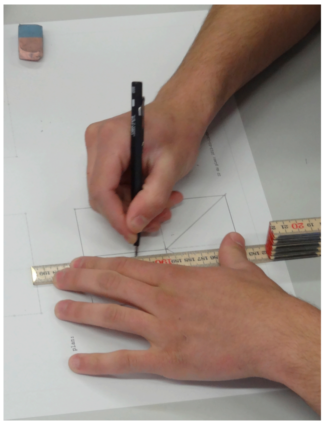
The activity is composed of two phases: a group-level activity and a class-level debriefing at the end. The first phase of the activity involves two students simultaneously and is composed of 5 steps in which the student is active: cutting the object (virtually) with TapaCarp, improving the plan, marking out the object, cutting the object, and comparing the objects. Each student is working on his own object and does not know the shape of the other student's object, as indicated in Figure 7.6. Thus, the only way for the apprentice to know the shape of the object that he must cut is to read the plan provided by his partner. The steps of the activity are described in more details below and summarized in Table 7.2; most of them are also illustrated in Figures 7.7 and 7.8, and in Video 3.

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 Figure 7.9). The view does not include any dimensions. The apprentice must therefore place the cutting planes so that the general shape of his object will resemble the object shown to him, but cannot be blamed for small inaccuracies.

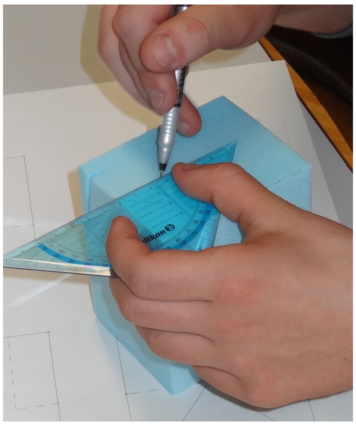
3. Improving the plan. On TapaCarp, the student inputs the position and orientation of the cuts required to create the object from the given solid. The printed plan displays the cuts made on TapaCarp, but in a "dumb" way: all the cuts are transverse to the object and are displayed with light stipple lines. For someone who does not have a prior knowledge of the final object,



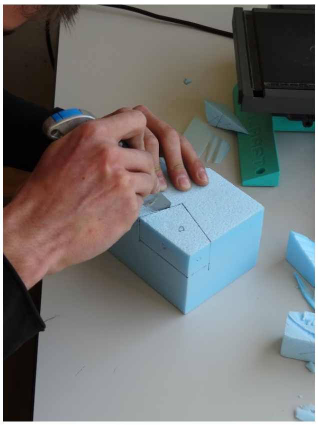
(a) Cut object virtually with TapaCarp.



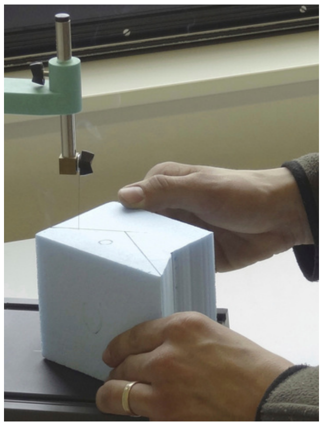
(b) Plan improvement.



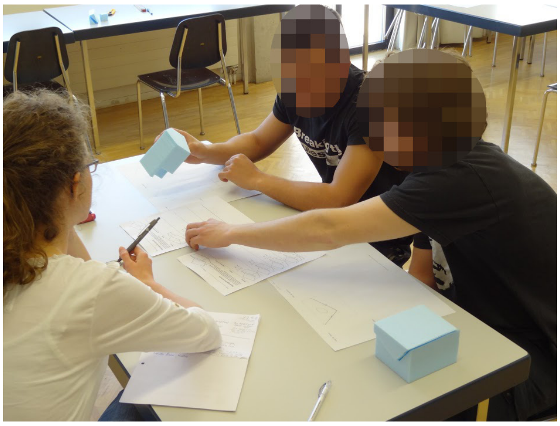
(c) Marking out the block.



(d) Knife cutting.



(e) Hotwire cutting.



(f) Group discussion.

(g) Debriefing preparation.

Figure 7.7: Steps of the first part of the activity.

Step	Name	Description	Timing
1	Cut object virtually	Using TapaCarp, define cutting planes.	5-15’
2	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’
3	Improve plans	The apprentice must improve his own plan so that it is understandable by his colleague.	5-15’
4	**⟷ Exchange plans**	**The plans of the students are exchanged.**	-
5	Mark out block	Based on the plan received from his colleague, the apprentice marks out a block.	5-10’
6	Cut block	Based on the marks he made, the apprentice cuts the block.	5’
7	Within group peer assessment	The two apprentices are brought together to compare the objects.	10’

Table 7.2: Description of the first phase of the activity. The bold line marks the split of SWISH.

just seeing these lines is not sufficient in order to create the actual object. The value of these lines is that they serve as construction lines for the student to produce the final plan that will be used by his partner. In other words, 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 what line should be solid, stipple, or simply non existing. This is a difficult task, as can be observed by looking at the raw plan in Figure 7.11a and the corresponding plan, once improved by a student (Figure 7.11b).

5. Marking out the object. When both students have deemed their respective plan ready, the plans are exchanged. Based on the plan received, the student marks out a polystyrene block so that he can cut it (Figure 7.7c). A key aspect of the activity is that the two apprentices in a same pair are working on two different objects. 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.

6. Cutting the block. Each student cuts out the object he has marked out, using a hot wire cutter and/or a utility knife (Figures 7.7e and 7.7d)).

7. Within group peer assessment. Once both objects are cut out, the two students meet with their object and their plan and assess the object created by comparing it with the printed perspective view that served as a prompt (Figure 7.7f). Their support to do that is a double-sided A4 paper sheet that shows a perspective view of the desired object as well as the correct three orthographic views (see Section A.6.3 for details). The evaluation is done mostly 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.

Figure 7.8: The second part of the activity: class debriefing.

Rationale for the activity

There are many reasons why this activity may benefit learning. It includes the entire workflow of carpentry in a short period of time, which can allow students to make a link between the different steps of their work (drawing, marking out the wood, cutting). In their current curriculum, students do not get to do that, either because the drawing activity is not followed by the cutting phase (e.g. at school), because they are not involved in the drawing part (e.g. at work), or because it takes them several days to draw the plan of the object before they can cut it (e.g. at the inter-company courses).

The hope is that doing both a school activity (drawing) and creating a final object, which they only do at work, will help them link the workplace and the school. More than the goal of this activity, this is in fact the goal of the research project in which this work takes place. The language used throughout the activity is also meant to increase the link with the workplace. Students are told that there is a client who wants the object that they are prompted with, and the peer assessment phase is presented as the delivery of the object to the client. It may also be that through this activity, students understand better the value of drawing, as well as the value and importance of a plan. As explained in Chapter 3, drawing is controversial in the carpentry community and the skepticism of practitioners is passed on to the apprentices. Through this activity, apprentices get a first-hand example of why drawing is important. It gives them an opportunity to make sense of what they know, which is commonly referred to as meaning-making (e.g. Suthers, 2006).

Another potential benefit of this activity is the confrontation between the students created by the SWISH scenario. This confrontation has been shown to benefit learning (e.g. Butterworth and Light, 1982; Vygotsky and Cole, 1978). It forces them to find the source and cause of a mistake and to negotiate until they agree. This is different than when they work individually, because a student who made a mistake might just correct it or cover it somehow (e.g. by copying an a neighbor) without understanding why it was wrong in the first place.

Having to generate the plans makes them reflect in a different way than usual. Be it at work or at school, they never “create” an object, but are rather given very precise instructions on how to make it. The activity puts them in the shoes of the plan designer, and although in this case they are not completely free (we prompt them with an object), one could imagine the same activity where students are given more freedom in the creation of the object. The advantage of prompting students with predefined objects is that the objects can be designed to make students notice information that they might otherwise oversee. This is a similar approach to the “contrasting cases” method that has proven effective for learning (Schwartz et al., 2011).

By making students go through the entire carpentry workflow, the activity can also trigger mistakes similar to the ones done in real work conditions. Moreover, if such mistakes happen, they are automatically documented through the artifacts (plan or object) and the teacher can use the artifacts to exemplify the mistake and make students reflect on them.

Finally, several aspects in the structure of the activity make it interesting. The activity is fast-paced, with several stations and frequent changes of station. The SWISH construction “gamifies” it and introduces a mix of collaboration and competition between the students.

All these are rationales explaining why the activity is potentially a good classroom learning activity. Some of the benefits are rooted in the design of the activity and could be claimed based on the existing literature (SWISH, contrasting cases, etc.). But most of them require an empirical validation, which is why we tested this activity in a classroom environment.

Research questions

There are three research objectives in this experiment.

1. Will this scenario, certainly promising on paper, work once deployed in the classroom? To measure that, we will first evaluate the usability of TapaCarp and that of the activity.
2. Second, we will also look at the learning gain that follows from such an activity, with the hypothesis that the activity brings a learning gain.
3. Finally, we also want to test whether working with plausible objects leads to a different learning gain than working with non plausible objects. Our hypothesis here is that working with plausible objects is more motivating and makes more sense to apprentices, but also easier, and could therefore lead to a lesser learning gain.

7.2.2 Experiment and method

Participants

Forty second-year carpenter apprentices, spread over three classes, took part in the study. The study took place at the practical school, where the inter-company classes are taught. The apprentices were engaged in a three week course of practical work. The apprentices were rewarded for their participation with a chocolate bunny. Fourteen other second-year apprentices served as a control group, and simply took the tests.

Flow of the experiment

The study was spread over a total of six days, two days for each of the three classes. The whole study happened in the theory room adjacent to the main workshop where the apprentices worked. The first day, the apprentices were brought all together to the theory room. The two experimenters explained the flow of the study (1 minute), demonstrated the use of TapaCarp (3 minutes), and administered a pre-test (15 minutes). All the apprentices were then left to go back to the workshop, except for the first pair of students who started with the experiment. As soon as the first pair was done with the TapaCarp station, the next pair was brought into the classroom. The pairs were the same as those already formed in the workshop.

All the pairs had done the first phase of the activity by the end of the morning or early afternoon. Using the material produced by the apprentices (plans, objects, and assessment sheets), the teacher and the two experimenters prepared the class debriefing session. This took 10 to 15 minutes. The apprentices were then brought back in for the debriefing session, which was led by the teacher, and lasted between 12 and 20 minutes. The procedure during the second day was similar to the first day, except that the objects were more difficult and that the debriefing session was followed by the post-test, and there was no pre-test in the morning.

Tests

The pre-test and the post-test can be found in Section A.6. The questions in both tests were the same, albeit with a different object. All questions were built around an object, and there were 4 questions per object. The first and second questions 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.

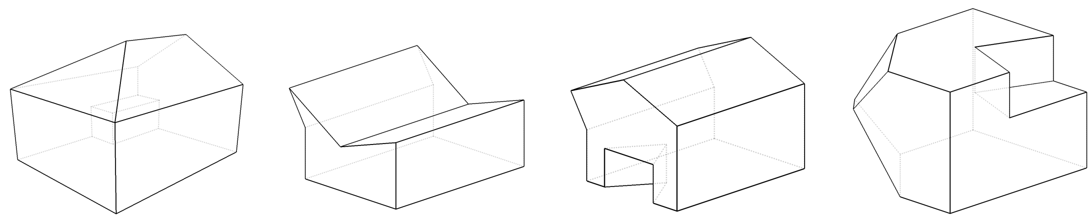
Variations in the material

The objects that the participants had to reproduce are shown in Figure 7.9. 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 objects that have a house-like shape. Non-plausible objects, on the other hand, are objects that are not necessarily more difficult to create, but whose shape would seem less natural in the construction world. In the rest of this chapter, 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 the objects, we asked 10 people in our laboratory to rate 13 objects on a plausibility scale from 1 (plausible) to 7 (non-plausible at all). The plausibility of each object was then defined as the median of all the ratings it received. Besides this plausibility score, a difficulty score was computed to reflect how hard it would be to draw and cut this object. The criteria to determine the difficulty score were: 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 selected according to their scores and the following constraints. Two objects with low difficulty were needed for the first day, and two objects with a higher difficulty for the second day. On each day, the two objects had to have a comparable difficulty, and one had to be plausible while the other had to be non plausible. Again, the objects chosen are shown in Figure 7.9 and their plausibility and difficulty scores are shown in parentheses in the caption of each figure.

The participants were randomly assigned to the objects on each day and we tried to balance them between the 4 possible combinations (plausible object on the first day, non-plausible on the next one; the other way around; plausible on both days; non-plausible on both days).



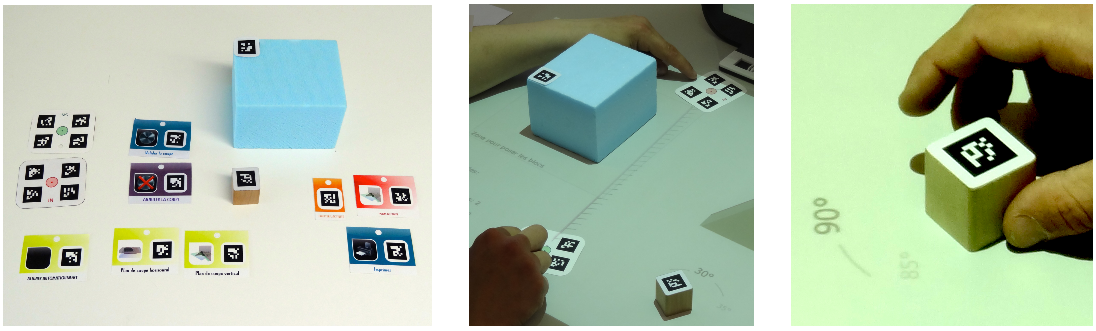
(a) P01 (1/2) (b) NP01 (5/1) (c) P02 (2/6) (d) NP02 (6/5)

Figure 7.9: 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.

Interface

For this activity, the interface given to the apprentices was composed of the blocks, two "selectors", one token, and five cards. These are shown in Figure 7.10a, on the left (the three cards on the right were not give to the students). The position of the cutting plane was determined by the two selectors (Figure 7.10b), and its orientation adjusted by turning a token (Figure 7.10c).

When the cutting plane was placed to their liking, participants used a card to validate the cut. Once a cut is validated, the intersection between the block and the cutting plane is shown in stipple lines on the virtual block. The last cut can be canceled using the "cancel cut" card. The "align automatically" card turns on a snapping of the cutting plane so that it is easier to cut in



(a) Components required for the activity.

(b) Using selectors to position the cutting plane.

(c) Orienting the cutting plane by turning the token.

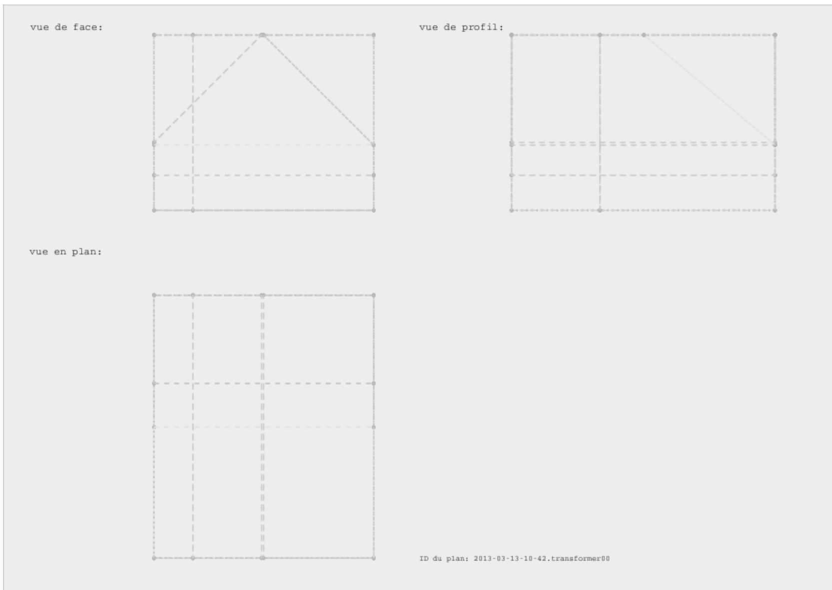
Figure 7.10: The components of the interface, and the manipulations of the cutting plane.

parallel or perpendicularly to the block, or to adjust the current cutting plane relatively to an existing one. The last two cards allow to change the cutting plane between a horizontal plane and a vertical one (default).

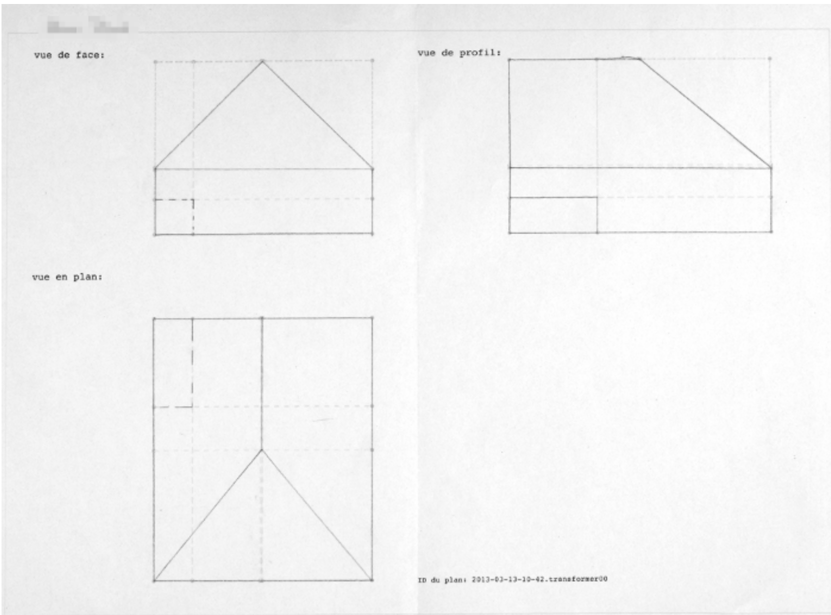
The experimenters kept some cards for themselves: the cards to quit the activity and launch the activity, and the card to print the plans (shown on the right in Figure 7.10a. The card to print the plans terminates the activity. Once this is card has been activated, a file is saved that contains the three orthographic views of the object cut, as shown in Figure 7.11a.

The typical display of the interface for this activity is shown in Figure 7.7a. 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.

Link with the 5 design principles. The design principles mentioned in the previous section were kept in mind while designing this new activity. Integration and flexibility were not fully respected. By design, although the activity makes use of paper (to draw plans), it does not integrate into an established pedagogical practice. The requirements for pairs of students and the SWISH script does not allow for much flexibility. On the other hand, empowerment, awareness, and minimalism, are well respected. The print card is left to the teacher, forcing students to call him before they can print their plan. The teacher is also empowered by his central role in the class debriefing. The awareness is high, because the physical organization of the classroom allows the teacher to see where students are in the workflow of the activity, as can be seen in Figure 7.12. Finally, the number of features and objects were kept as small as possible to enforce minimalism. TapaCarp was also used in a minimalist way, much of the activity being done without it.



(a) The plan saved.



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

Figure 7.11: The plan (a) saved by TapaCarp; (b) the same plan improved by an apprentice.

Figure 7.12: The physical organization of the classroom in stations increases awareness.

7.2.3 Results

Evaluation of the interface

Most of the apprentices were able to use TapaCarp in a satisfactory way, despite their lack of experience with TapaCarp and the rather high difficulty of the task.

Going past 90°. The most frequent question concerned how to make the cutting plane "go the other way". The cutting plane was indeed designed to be able to go from 0° (horizontal) to 90° (vertical), but not to more than 90°. The rationale for this design was that this is similar to real case situations in which a saw can only cut a piece of wood from one side. To circumvent this self-inflicted "limitation", the solution was to either turn the block or turn the cutting plane. This was explained to the users in the introduction, but some of them forgot. Once shown again how to do it, though, they remembered and the problem disappeared.

Height of horizontal plane. Another problem that users did not know how to solve was how to change the height of a horizontal cutting plane. The solution was to turn the same token used to adjust the vertical angle of the cutting plane. However, the users often tried to move the two ends of the cutting plane, which in this case had no effect. Again, once told how to do it, users remembered and the problem disappeared. In the future, an additional token could be used to adjust the height of the cutting plane. This would also simplify the interface by removing the cards to change from vertical to horizontal plane and make for a more consistent interface overall.

Automatic alignment. In some cases, the automatic alignment of the cutting plane with the object was a source of problems. This happened in specific situations where several points were simultaneously trying to snap the cutting plane to a different position or orientation. As a result, the cutting plane oscillated between several positions. This problem had been identified in pilot studies and addressed by the introduction of three cards instead of one for the snapping. Each card activated one kind of snapping only. Although they existed, we decided not to give these cards to the users as this would have overwhelmed the users, and the cases in which this problem appeared were rare. Instead, one of the experimenters kept them and used them when such a problem was reported by a participant.

Flickering. One of the systems had problems detecting the tags from time to time. This resulted in an unpleasant flickering effect of the entire workspace. The cause of the flickering was the constantly changing lighting conditions. Adjusting the capturing parameters of the camera solved the problem, but the application had to be stopped to do that, and it was therefore only done between students.

Variability among users. There was a dramatic variability among the participants, both in terms of attitude towards the system and in terms of efficiency in learning to use it. Most participants showed a positive attitude towards TapaCarp, and some of them even showed excitation and asked questions about how it worked. A handful of students, on the other hand, did not like the idea of working with such a machine, saying that they preferred working with their traditional pen and paper environment and were also more efficient with it.

Adaptation time. Getting used to such a new interface is not easy and participants needed an adaptation time. On the second day, they were in general faster at creating the objects despite the higher level of difficulty of the objects. The average time per cut fell from 3 minutes and 30 seconds (SD: 1 min 53 s) to 2 minutes and 58 seconds (SD: 1 min 1 s). This indicates that apprentices were were globally more effective with TapaCarp on the second day.

Dimensions. Beyond the novelty of the system, what unsettled some students was the activity itself. In particular, the absence of any dimensions on the prompt puzzled them, because they are not used to estimate ratios by simply looking at an object. This translated into a requirement to add a grid or some kind of reference system so that they could more easily reproduce the object.

Learning performance

There was no significant learning gain difference betweeen the control and the experimental group ($t[25]=.13$, $p=.89$). The control group performed slightly lower 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 in the next section.

Looking at each type of questions separately, there was no difference either. One marginally significant difference ($F[1,48]=3.6$, $p=.07$) appeared on the questions that asked whether all three orthographic views were necessary to be able to cut the final object. On these questions, participants in the control condition had a negative RLG (-13.7%) whereas the students in the experimental condition had a positive one (12.7%). The latter is not surprising since in each and every debriefing session there was a discussion on whether or not the 3 views were needed, and the students were made aware of the risk of adding errors by drawing unnecessary views.

Progression between day I and II

Besides the pre-test and post-test, another way to measure the learning gain is to look at the evolution of the performance between the two days of the experiment.

The main performance metric is the correctness of the object. The correctness of the object was scored according to four levels: (1) the object was not at all resembling the target object; (2) the object had one major mistake but making abstraction of this mistake, the object could be recognized; (3) the object was mostly correct, but contained some inaccuracies; (4) the object was correct. The scoring was done together by two researchers.

The details of the correctness score are shown in Table 7.3. 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.

	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

Table 7.3: Details of the correctness of the objects and of the plans.

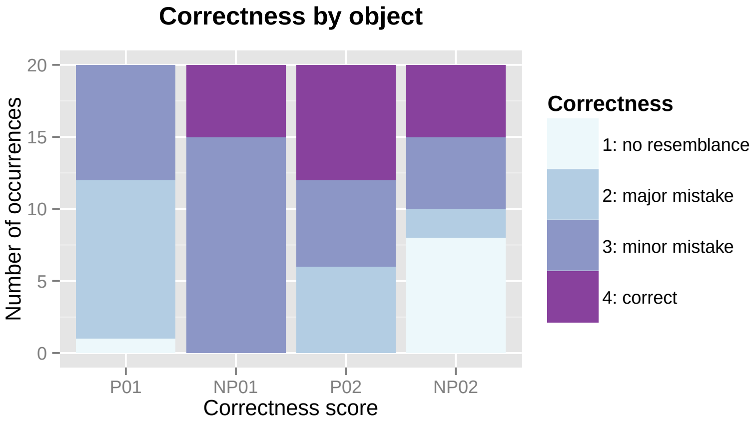


Figure 7.13: Correctness score by object.

While the average correctness hints that there was not a big difference between the two days, the distribution of the correctness scores tells a different story. As can be seen in Figure 7.13, 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 always almost correct.

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 ($F[1,78]=4.74$, $p=.03$).

The effect of plausibility

This 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 7.4 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).

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

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

Table 7.4: Source of errors for each type of object.

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.

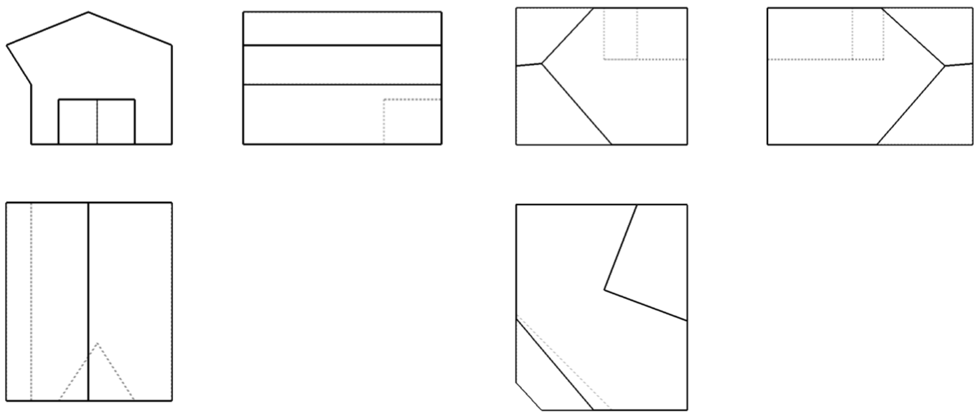
One explanation is that the criteria used to assess the difficulty level of an object do not capture everything related to the difficulty level. While this is possible, it is hard to find additional criteria that could reflect better this gap of difficulty. A more likely explanation is that there is 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, P02 and NP02 (Figure 7.14). Although there were fewer cuts on NP02 than on P02, the P02 plan (Figure 7.14a) is easier to interpret for a human brain that has been trained to read plans of houses, than the plan of NP02 (Figure 7.14b).

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, Figure 7.14b by itself is hard to decipher, but once one knows how 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.

"Carpenter instinct"

Out of the 18 correct objects, 9 came from wrong plans, and 9 from correct plans. While it is expected that correct plans lead to correct objects, it is surprising that 9 objects were correct although the plan provided to cut them contained, on average, 2.5 mistakes. When asked how they transformed plans in which there was an ambiguity, or a mistake, into a correct object, several apprentices said that they used their "carpenter instinct".

Correcting a potentially flawed plan is a key skill of carpenters. Even in the workplace, plans



(a) P02. (b) NP02.

Figure 7.14: Orthographic projections for the objects of the second day.

might be wrong and could need to be corrected. Assuming a plan on which the three views are drawn, and that at least two of them contain the correct information, many mistakes can be easily fixed by discarding the information that is the least frequent between the three views. One could expect that the plausibility mentioned above would play a role in correcting a wrong plan, too. However, out of the 9 objects fixed by the cutman, 4 were non-plausible (5 plausible), and it is therefore not clear that having plausible objects helped apprentices apply better their “carpenter instinct”.

The ability to correct plans could make us doubt the usefulness of a correct plan. However, the probability of having a correct object if the plan is wrong is 14.5%, whereas the probability of having a correct object given a correct plan is much higher (50%). This translates into a significant influence of the plan correctness onto the object correctness: a correct plan increases on average the object correctness score by 0.84 ($F[1,78]=12.0$, $p=.000$).

Types of mistakes

A total of 80 objects and 80 corresponding plans were created by the apprentices. Out of the 80 plans, 64 (80%) contained at least one mistake. Table 7.5 shows the distribution of the number of mistakes. About 50% of these 63 plans (30) had only one or 2 mistakes, while only 17% had 5 mistakes or more, the maximum being one plan with 8 mistakes.

We classified the mistakes into 4 categories: line, orientation, position, and shape. An example of each of these mistakes is shown in Figure 7.15. The line mistake is when lines are missing or are drawn with the wrong rendering (stipple instead of solid, for example). The position mistake is when a line has the right rendering, but that its position or orientation does not match the object, for example when the door is much bigger than it should be. The orientation mistake is when the view has been drawn by looking at the object from a wrong point of view. Finally, the shape mistake is when the shape itself is wrong, for example when a door is missing. When this is the case, the shape mistakes trumps the line and the position mistakes in the same view.

Table 7.6 shows the number of mistakes for each of the mistake categories. A plan can have several of the same mistakes. This is often the case, for example with the shape mistake: if a cut is missing on several views, the shape mistake will be noted several times.

number of mistakes	1	2	3	4	5	6	7	8
number of plans	7	23	12	10	7	3	0	1

Table 7.5: Count of plans with a given number of mistakes.

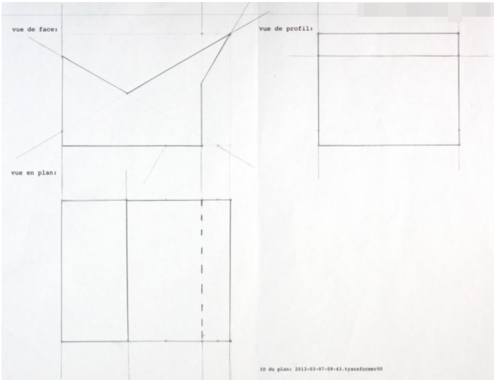
view	line	orientation	position	shape
front	19	2	26	24
side	17	10	12	25
top	13	0	19	22

Table 7.6: Type of plan mistakes.

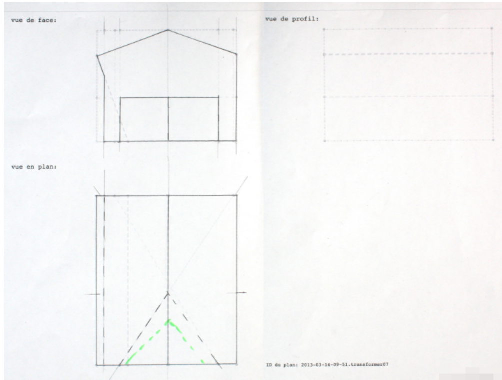
Orientation mistakes. The higher number of orientation mistakes on the side view confirms the confusion observed around in the previous experiments (see Chapter 6): the students do not seem to remember whether the object must be looked at from the left or from the right to draw the side view.

Position mistakes. The high number of position mistakes in the front and top views compared to the side view is surprising. One reason is that the side view was more often not drawn (11 plans) than the two others (2 plans for the front view, none on the top view). But even taking into account the actual number of views drawn, the frequency of these errors remains lower on the side view (17%) than on the front view (33%) and on the top view (24%). This may seem in contradiction with results of Chapter 6, where in both studies the side view was the most prone to mistakes. However, looking closer at these mistakes, it appears to be a consequence of the orientation of the object on the plan. Indeed, the salient features of the objects were more often linked to their x-coordinates, which are only observable on the front and top views. For example, one common mistake was that the eaves ("avant-toit") was too short on object P02, which could only be seen in the front and top views.

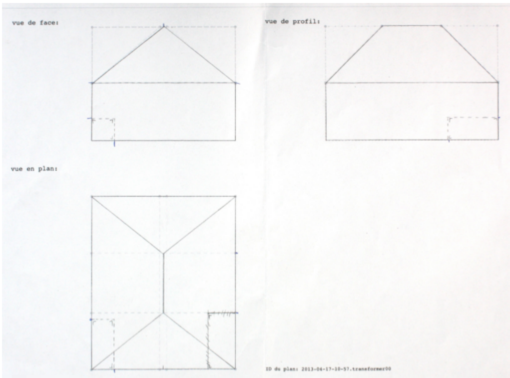
Line mistakes. For the number of line mistakes, taking the ratio of the number of mistakes to the number of views effectively drawn shows that the frequency of such mistakes on the side and front views is similar (about 25%), but lower on the top view (16%). This is not easily explained, because the orientation and topology of the objects does not seem to favor the top view. One explanation could be that it is simply easier to imagine an object from the top than from the side or the front views.



(a) Line: on the side view, a horizontal line is missing.



(b) Position: the eaves is too short and the door too wide.



(c) Shape and orientation. Shape: there are two hip roofs instead of one. Orientation: the side view was drawn looking at the object from the wrong side (right instead of left).

Figure 7.15: The four categories of mistakes found on the plans.

Shape mistakes. The high number of shape mistakes is also surprising at first sight. It is explained by the fact that many of these instances were mistakes where the contour was missing (more than half of these mistakes for the front and side view, and one third in the case on the top view (13, 13, 7)). Students forgot to draw the contour or thought it was obvious that it would be there, and did not draw it. This leads to unfinished shapes that, if the plan if followed exactly, are uncuttable.

Inconsistency between the views

A concept whose importance students often fail to grasp is *consistency* between the views that they draw. Double checking that their drawing is consistent can help fix mistakes and make sure the plan is not ambiguous. If a plan does have an ambiguity, i.e. if the views are inconsistent, then it is up to the cutman to choose what information is the correct one. In some cases, as mentioned above, the cutman is able to recover the right object by cross-checking information between the views. But this is not the case in general, and plans should therefore not have any ambiguity, or in other words, all views must be consistent. However, while consistency is necessary, it is not sufficient: a plan can be consistent but wrong.

We recorded inconsistencies pairwise between the views of a plan, and the number of inconsistencies was equally distributed between pairs of views. In total, 44 plans had at least two inconsistent views and were therefore wrong. That represents 55% of all plans and 69% of the plans that were wrong. It also means that 36 plans were perfectly consistent: the 16 correct plans, but also 20 (31%) of the wrong plans.

These consistent, yet wrong, plans, are interesting because they imply that the same mistake was reproduced on the 3 views. This highlights a problem in reading a 3D perspective view of an object rather than in transcribing a 3D object into 2D views. Transcribing an object from 3D to 2D is an unusual task for apprentices and they had trouble doing it accurately. For example, on P01, a frequent error was for students to add another triangle hip facing the one that actually exists (as shown in Figure 7.15c). But if this hip existed, there would be a stipple horizontal line in the background of the 3D view (Figure 7.9a) to mark the base of the triangle. The good news is that the number of wrong, and yet, consistent plans, decreased by half between the first and the second day (13 to 7). At the same time, the number of wrong and inconsistent plans did not decrease similarly (21 to 22). This indicates that the reduction of wrong plans from day one to day two (from 34 to 29) was mostly due to a better consistency check.

In rare cases, an inconsistency was corrected by adding an annotation (e.g. an arrow to indicate a point of view for a given view) to the plan. Annotations were also used to add information to the plan in order to improve its readability, such as crossing out lines, adding hatching, naming the cuts with letters, or adding short written comments. However, annotations were rare: out of the 80 plans, only 15 contained some. Most of these were lines that were crossed out, and only two included short textual information (e.g. “not needed” and “only pencil lines”). This is surprising and could be explained by the didactic contract (Brousseau, 1997) between the apprentices and the teacher: the teacher does not expect any annotation, and therefore the apprentices do not provide any. Another explanation is linked to the identity of carpentry, as will be discussed later.

Class debriefings

The class debriefings were prepared together with the teacher by examining the objects and the plans produced by the apprentices (Figure 7.7g). All the objects – each object together with its plan and the group discussion sheet – were placed next to each other. They were grouped by object type, and then by correctness. The researchers and the teacher made a list of points to discuss based on the objects produced by the apprentices. This 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. 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, as a tailored feedback is what helps the students understand their mistakes and build their knowledge.

In total, there were six debriefing sessions. Each of them lasted for 12 to 20 minutes 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 7.7.

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

Table 7.7: 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.

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 a faulty interpretation of the prompt object, as described above. 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 work of a carpenter in each session, and very frequently. 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." Many of the links to the workplace came from the value of the plan. The design of a plan was by far the subject that was mentioned the most by the teacher. 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.

These four aspects are tightly linked together. The need for only the necessary information came mostly from the fact that many apprentices drew three views when only two where necessary. Adding the third view often introduced inconsistencies between the views. One kind of inconsistency often observed was that a view drawn from a wrong viewpoint (typically the side view from the right side, instead of the left side). All this allowed the teacher to remind the students that the plan is the only means of communication: "Everything is based on the plan. If the plan is wrong, you cannot cut the desired object. So you have to pay a lot of attention to the plan you are drawing". This is typically not something that apprentices learn at school, where they work individually on an object, from drawing the plan to the final object. However, in companies, the various steps are often done by different people, as the teacher reminded them.

Plan reading. Still linked to the plan was the notion of reading a plan. 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 followed in school and recommended that they should use this knowledge to apply a step-by-step approach to mark out the object and cut it. The teacher's message can seem contradictory: on the one hand, he asked his students to observe the prompts and reproduce them in an intuitive way with TapaCarp, and on the other hand he also asked them to use the drawing algorithms that they know. This sums up a recurring tension in carpentry training: carpenters need to be quick at going back and forth between the 2D and 3D representations of an object. Using their intuition, they should be able to find the important features of an object rapidly. At the same time, carpentry is also a craft of accuracy and precision, and no approximation can be allowed when converting a plan into a final object. One benefit of this activity is that it involves both aspects of the profession, and it allowed the teacher to highlight them with examples.

7.2.4 Discussion

The goal of this experiment was to try TapaCarp in a new pedagogical scenario, and to study (1) if it was usable, (2) if students could learn with it, and (3) if varying the plausibility of the objects had an influence on learning. The novelty of the scenario was twofold. First, the activity itself was new in that it required the apprentices to go through all the main steps of designing and creating an object. Second, it was also the first time that TapaCarp was integrated in the classroom as one piece of a classroom activity.

Learning gain

The experimental group and the control group did not display any learning gain differences. While this is disappointing, several factors can explain it. The first is that 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. This could be improved in the future based on the current results.

There was a 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. 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.

Activity

The activity “worked well”, i.e. its global usability seemed good. Its fast pace and the novelty effect helped create a positive atmosphere around it. The simple interface made getting started with TapaCarp smooth for all but a few apprentices, and they were all able to produce their plan. In some cases, however, the difference of speed between students forced one of them to wait idling. This could be avoided by pairing students by level of expertise, but doing so could reduce the pedagogical richness of the discussion (discussing perfect objects is uninteresting, and discussing completely wrong object is pointless).

Tangibility. Tangibles were used in the interface of TapaCarp, but also served for the entire activity, and were integral in easing the orchestration process. Apprentices moved along the stations with their plan and their block. During the debriefing, it was easy for them to find their blocks and plans, and to compare their performance with others’. Tangibility was also helpful for the teacher, who could easily get an overview in preparing the debriefing, sort objects according some criteria, and compare them with their plan. During the debriefing session, he could place the objects next to each other to compare them and point out some particularity on it. In summary, these artifacts contributed greatly to reducing the orchestration load and easing the transition between the three levels of orchestration (individual, group, classroom). It also eased another transition: that between the digital and the non digital part of the activity. Using the design principles defined in the previous section, we can say that the artifacts allowed a greater awareness and flexibility, while being minimalist (they are just objects). It would probably have been hard to achieve the same reduction of orchestration load for a similar activity done entirely in a digital space (although a comparison is needed to assess that for sure).

Meaning-making. A side effect of the tangibility and cutting actual objects was that it allowed the apprentices to link this classroom learning activity to their job. As explained in the introduction, to many apprentices, drawing is disconnected from producing objects. This activity aimed at giving an action value to the plan by making it an object of communication through which the object is built. The focus of the activity was on the final object rather than the plan. This led to many references to practical work situations and forced the apprentices to think about the value of the plan. For example, many did not even think of not drawing the three views, and when asked why they had done it, they said "this is how we have learned to do it in the drawing class". The teacher then pointed out that when they draw a hip rafter at work, they only draw two views. This, hopefully, helped them realize that the drawing is not an end in itself, but a means to create an object, and that it should be used with judgment.

Changing the habits. While still in the design process of the activity, the teacher had pointed out the advantage of cutting through a solid rather than folding flat sheets of paper into a solid. The latter is often done during drawing classes at school, and means that the apprentices only end up with a 3D object at the very end of the process. In this activity, to generate the 2D plan, one apprentice starts by (virtually) cutting an object in 3D, and the other one does the contrary. The 3D-2D transition is not done by drawing 3 views, but by cutting, which is much closer to their future practice. The 3 orthographic views come only after the cutting phase, as a byproduct. Some apprentices were thrown off by this approach, as reflected by their complaints that the prompts did not have any dimensions, but it forced them to observe the objects with more acuity and to create an accurate mental model of them.

Need for a mediator. The success of the group discussion was mixed: 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 the mediator's intervention. Whatever the reason for this (uneasiness in blaming their partner, lack of attention to details, etc.) it is one of the weak points of the activity and would need to be addressed so that all students discuss. One solution could be to explicitly reward the finding of mistakes. A recurrent problem was that once they had discussed, they needed to write down the problems they had spotted. Although the writing was done partly in a graphical way (see the debriefing sheets in Section A.6.3), they were unsure of what to write.

Common language and music analogy. The scarcity of annotations (written instructions or graphical indications other than the drawing itself) on the plan surprised us. As mentioned above, one explanation is the didactic contract. The teacher, however, had a slightly different take on his. He explained that "many carpenters, including me, have an inferiority complex when it comes to writing because we make lots of spelling mistakes. Therefore, we keep to the plans and to drawing, which is our way of communicating". Plan literacy is where carpenters excel, and part of their professional identity is rooted in their ability to read a plan and cut a piece of wood based on it. Therefore, however complicated the plan is, it does not need annotations because the plan is self-explanatory to another carpenter who "speaks" the same language. The development of a common language to communicate information is key for carpenters. Knowing how to communicate through plans (orthographic projections) is what makes them belong in their community of practice. Indeed, the perspective view of an object makes sense to everyone, but only experts (including carpenters) can imagine an object based on its orthographic projections, and only them, too, can transcribe an object into a plan. To make an analogy with music, novices are able to sing a song or recognize one when it is played;

but only expert musicians can imagine how the song goes when just seeing the music score, and only them, too, can transcribe a melody into a music sheet. The auditory perception that everyone possesses is similar to the visual perception that allows everyone to interpret a perspective view into a 3D shape, and music scores are the language of musicians as plans are the language of carpenters.

TapaCarp

There were no major issues in the usability of TapaCarp. The minor problems encountered did not prevent any of the apprentices from completing his plan and the very limited interface allowed them to quickly get started and be efficient. There is nevertheless room for improvement, such as allowing the cutting plane to rotate more than 90° and by a finer step than 5° or revisiting the way the horizontal plan is moved up and down.

The teacher's feedback was that this was good for novices, but that a more advanced version of TapaCarp could be designed to allow students to finalize the plan with the machine and remove the need to manually improve the plan. The goal of the current activity was precisely to have students improve a plan by drawing, but moving forward, the design suggested by the teacher could serve as an introduction to computer-aided design (CAD), certainly an important topic for future carpenters.

This would also answer the criticism of a couple of apprentices who did not see any added value in using TapaCarp. In the experiment, we prompted them with an object because we wanted to be able to compare the objects between them for the class debriefing. However, students could be free to design any – and potentially more complicated – objects, in which case they might see more use to TapaCarp.

Other improvements suggested were to display the orthographic projections at the time of cutting, and to allow the user what views to print. Displaying the orthographic projections would allow the user to see with a greater precision how they are cutting. Moving up the decision of what view(s) to include before the impression of the plan would also go towards CAD. Every future change should be considered in terms of usability, but also in terms of its pedagogical impact. For example, giving an indication of dimensions through a grid, or displaying the intersection of two cutting planes on the plan, as wished by the apprentices, may please the students and make TapaCarp look more "usable" to them, but it would make some parts of the activity easier and limit a desired cognitive effort. It is also possible to consider that some functions be (de)activated depending on the focus that the teacher chooses for the activity.

7.2.5 Conclusions of the study

In this section, a rich and novel learning activity was presented. TapaCarp served as a launchpad for the activity, but was only used for a part of it. 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.

The usability of both the activity and TapaCarp were good. The tangible artifacts provided a natural workflow and reduced the orchestration load. The class debriefings, based on these artifacts, were extremely rich and allowed the teacher to touch upon many practical points, making frequent references to the workplace, and thus participating to a meaning-making process for the apprentices.

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.

There is still room for improvement. For example, one open question is how to ensure that a rich discussion takes place in a pair, without having to resort to a mediator. There are also various minor improvements that could be done on TapaCarp, and one major one that would transform TapaCarp into a tangible CAD learning tool. Finally, it would be interesting to explore whether other objects might trigger different discussion points than those noted with the current four objects.

7.3 Conclusions and Discussions

This chapter presented two activities that attempted to integrate TapaCarp in the classroom in two different ways. The first activity used TapaCarp in existing practices, whereas for the second one, a new pedagogical scenario was created that used TapaCarp in part of it only. Both activities were tested in a classroom with carpenter apprentices.

From the first activity, we derived 5 design principles (integration, flexibility, awareness, empowerment, and minimalism) that can help decrease the classroom orchestration load. The deployment in the classroom showed that a distributed interface could mold into the classroom, but that too much distribution could decrease its usability. To that end we introduced the concept of Hutchins threshold, a point after which the distribution of an interface becomes harmful.

The second activity showcased how a new pedagogical scenario could take advantage of TapaCarp. The apprentices used TapaCarp only at the beginning of the activity. We observed that the use of tangibles throughout the entire activity, i.e. beyond the interface, participated to the success of the activity by reducing the orchestration load and helping the students remember their mistakes. The activity also allowed the teacher to make frequent references to the workplace, a key aspect to bridge the workplace and the school.

7.3.1 Limitations

One limitation of both studies is their failure to show a positive learning outcome. As explained in Section 7.2.4, there are various reasons that make it hard to show the impact of the activities on learning. Both 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 activities would work in different contexts. For both activities, we only worked in one school, with one teacher, and 3 classes of apprentices.

Both activities were run by researchers. We tried to involve the teacher as much as possible, but for logistics reasons, researchers still needed to be involved. Having the activities be run by the teacher would require a training on the usage of TapaCarp, and could also potentially alter the findings.

7.3.2 How should TUIs be used in classrooms?

Each activity integrated TapaCarp in the classroom in its own way: in existing pedagogical practices, and by defining new ones. It is hard to judge if one is better than the other. Designing a new scenario with TapaCarp in mind allowed to build a scenario that fostered rich interactions and discussions. The teacher and the apprentices did not reject TapaCarp; on the contrary, they asked for more features (CAD) that would increase its usage. When integrated in existing learning practices, TapaCarp served mainly to guide students through the activities and to help those that had more difficulties. The augmented reality capabilities of TapaCarp were used more heavily than for the second activity, and the students' feedback was positive. Yet, the added value of TapaCarp for learning seemed less compelling in the first activity than in the second one, even though in both cases there was no clear learning gain difference between the experimental and control conditions.

Two things must be distinguished: the pedagogical scenario (new versus existing), and the usage of TapaCarp in the classroom (one TapaCarp per group of students versus one or two TapaCarp for the entire class). The two are not completely independent, as it is harder to integrate TapaCarp as part of an existing activity. This is why we started by fully replicating an existing activity in TapaCarp and deployed five systems in a classroom. However, having five TapaCarp systems in the classroom seemed to scare the teacher, is expensive, and requires more logistics. Therefore, the more sustainable way to use TapaCarp (or for that matter any TUI) in a classroom, is probably to integrate it as a part of a learning activity, as was done in the second activity presented in this chapter.

When doing so, tangibles should be used heavily, as they can significantly decrease the orchestration load. Blending with the rest of the activity is a unique advantage of TUIs, compared to GUIs or multi-touch environments, that has not been mentioned in the literature before. In the future, this integration could even be increased by closing the cycle and bringing some modified object back to the system.

8 Scaling up

The studies reported in the previous chapters have shown that TapaCarp can be useful to develop spatial skills and how it can be integrated in the classroom. There is a true potential for TapaCarp, and for tabletop TUIs in general, to improve the learning and teaching experience. Despite this potential, TUIs are still used scarcely in schools. For instance, a private company (Simpliquity Sàrl) has been offering a version of the Tinkerlamp with many activities for logistics apprentices for 3 years, but only 5 schools (25 lamps) currently use it. Reasons for this go beyond TUIs and include the doubts about the impact of technology on learning as well as the usually slow adoption of technology in schools.

However, the slow spread of TUIs may also have been caused by some of their features. Thus, most of the TUIs require a dedicated software and hardware, as well as objects that form the interface. The specific nature of this hardware and the setup it requires makes it hard to reuse existing hardware that schools already own. For example, the Tinkerlamp includes a projector and a computer, two items that are present in most Swiss classrooms, but reusing those of the classroom for the augmented reality environment of the Tinkerlamp is not easily feasible. This means that the ratio of cost to usefulness of tabletop TUIs for schools is rather high. Moreover, there are other logistics constraints, such as storing those often bulky and heavy items, and attending to their maintenance.

If the Tinkerlamp (or other tabletop TUIs) are to be used more widely, these issues need to be tackled. The cost of hardware is likely to go down in the future, but it will never be free, and adding the design and manufacturing costs is likely to keep the Tinkerlamp at a price too high for most schools. In this chapter, we will investigate how recent advances in web technologies can be used to achieve a higher scalability of tangible tabletop interfaces, mainly through a lower cost of the technology.

Surprisingly, this issue has not been mentioned much in the literature on TUIs. Shaer and Hornecker (2009) wrote about the issue of scalability for TUIs, but focused more on the difficulty to solve more complex problems (i.e. with more objects) than on increasing the user base. Costanza et al. (2010) built d-touch, a visual marker recognition system that allows users to build low-cost TUIs. Based on d-touch, Costanza et al. advertised two TUIs for music that were freely downloadable from a website, as well as tutorials on how to build the various components of the interface to get started. They registered 25'000 visits over 9 weeks, but only 273 visitors (1.09%) interacted with the website for more than a minute. They did not provide an analysis of why so few visitors ended up using their system.

One project at a large scale was the AR-Jam book project run by the BBC in 2006. AR-Jam book is an augmented reality application for storybook aimed at children of 5 to 7 years old. Children interact with a scene of the book by manipulating markers that control the main characters of the story. The markers are recognized by a webcam placed above the screen. On the screen, children see the camera's view, except that the markers are replaced by elements of the scene in a traditional augmented reality way. Unfortunately, the project was stopped and the webpage is not accessible any more. There has been no report focusing on the scale of the on the project, and the scalability issues that we are interested in studying in this chapter. The on existing reports on this project being small-scale qualitative studies (Hornecker and Dünser, 2009).

The scalable solution that we designed and that is presented in this chapter, e-TapaCarp (Figure 8.1), differs from that of Costanza et al. in three ways. First, it runs online and does not require any download and installation of software. The second difference is the application domain: whereas their interface was for musical composition and performance, ours targets the training of spatial skills. As a consequence, and this is the third main difference, our interface uses visual instead of audio output.

Compared to TapaCarp, in e-TapaCarp, the projector is replaced by a screen, and the camera by a webcam. We assumed that users of e-TapaCarp would have access to a computer connected to the Internet and the application is therefore running in a browser. This is a reasonable assumption for so-called developed countries; for example, in 2012, 73% of households in the European Union had an Internet access, according to Eurostat, the statistical office of the European Union (Seybert, 2012). Another assumption is that users have access to a printer to create the parts of the tangible interface. Despite the need to have an access to a printer and to the Internet, the potential user base for e-TapaCarp is much larger than that of TapaCarp, for which users need the Tinkerlamp.

A larger pool of potential users does not guarantee a larger number of actual users. Besides cost, other factors such as interest in using the technology, or the effort and time to invest before being able to use it, can slow down adoption. To get first-hand data on these aspects, we did a first test deployment of 5 weeks with e-TapaCarp. The work on e-TapaCarp is just starting, and what is reported in this chapter is a first feasibility study. E-TapaCarp and its design is described in the next section. Section 8.2 reports on a first 5 week deployment, and the findings of this first work with e-TapaCarp are described in Section 8.3.

8.1 Designing the Scalable Interface

8.1.1 Hardware components and setup

The idea of e-TapaCarp is similar to TapaCarp's: the system connects 3D objects (equipped with fiducial markers) and their digital 2D orthographic representations (Figure 8.1). However, the absence of a projector implies some changes. The input and output spaces are not co-located and the input space cannot be augmented; this means that the input space (also referred to as the workspace) needs to be delimited in another way than with the projection of light. E-TapaCarp is also less controlled than TapaCarp, as the resolution of the screen and the webcam are unknown, and so are the position and orientation of the webcam relative to the workspace.

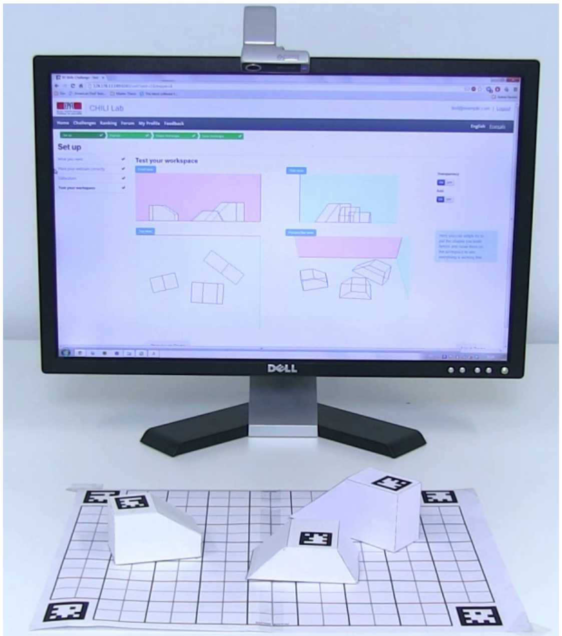


Figure 8.1: E-TapaCarp setup.

In TapaCarp, the camera and projector positions are fixed. This means that the system can be calibrated once for all before it is handed to the user, and the user is therefore oblivious to any calibration process. With e-TapaCarp, users place the webcam in an unknown place and the system therefore needs reference points on the tabletop to know where the webcam is placed relatively to the workspace. This is the purpose of the grid with the 4 markers (Figure 8.1), which allows the system to determine the scale of the workspace and to transform the location of the objects in the user's perspective.

The calibration is fully automatized, and the user's task is only to set up the workspace and webcam so that all 4 tags of the workspace can be seen by the webcam. The details on how to place the webcam optimally are explained to the users on the website through textual, graphical, and video information[1]. The challenge for the user is to find a position for the webcam that is neither too high (otherwise the markers are not detected), nor too low (otherwise the 4 markers of the workspace cannot all be detected). The range of distances depends on several parameters (field-of-view of the webcam, size of the tags, etc.), but a typical well-placed webcam is at a distance of 30 to 60 cm from the workspace.

8.1.2 Software components

E-TapaCarp is made of three main components: a tracking system, a rendering engine, and a learning management system, all of which run online. Each of them is described in some more detail below. Once again, more detail on these components is provided in Bossy (2013).

Tracking. Similarly to TapaCarp, e-TapaCarp uses fiducial markers to track the position and orientation of objects. Besides accuracy and robustness to changing light conditions, one important feature of the tracking library for e-TapaCarp is the maximal distance and angle at which a marker can be detected. This is important because it gives more flexibility to the user in positioning the webcam. After several comparisons, we chose to use `JSARToolKit` (Heikkinen, 2013).

Rendering. Like TapaCarp, e-TapaCarp needs to do non trivial rendering of 3D scenes. For this as well, an existing library is used (Gardner, 2013). There are two display options that the user can turn on or off. The first option allows the user to switch the rendering of the objects from transparent to opaque. The default is transparent, but an opaque rendering can be helpful when they are many objects in the scene, as described for TapaCarp in Section 5.3.4. The second option allows the user to show or hide the grid. By default, the grid is shown, as it can serve as a helpful reference between the three views and the workspace.

Learning management system. One of the goals for e-TapaCarp is that it should be self-explanatory and offer learners a personalized experience. There are many available web platforms, and we chose to use Google Course Builder (GCB) (Google, 2012), an open-source project that already includes features such as login, HTTPs, user progress, forums, and announcements, and has a strong support community. GCB allows to create basic courses divided in lessons. We integrated the rendering part of e-TapaCarp into GCB. The final interface is demonstrated in Video 4.

8.1.3 Comparison with TapaCarp

For the user, the main differences between TapaCarp and e-TapaCarp are the requirements to build the interface, to position the webcam and to run the calibration. The separation of the input and output spaces also changes the interaction, and so does the absence of augmentation on the objects, at least for some activities.

[1] As of August 7, 2013, the website is still up and running at http://tangiblecourse.appspot.com.

Foreseeing e-TapaCarp in a classroom, we can evaluate it according to the 5 design principles presented in the last chapter. Awareness and flexibility do not change significantly. Integration can be seen as better: assuming there is a computer in the classroom that can be used to run e-TapaCarp, the logistics of setting-up the system is easier. The connection to the web can also make for a better integration with the rest of the apprentice's world, for instance by allowing him to upload pictures or plans from the workplace. In the same vein, using the computer can also facilitate the integration with the existing material that the teacher may have on the computer, such as CAD models.

TapaCarp can be used "out of the box", as opposed to e-TapaCarp, which requires a set-up from the user (in the classroom, that would probably be the teacher). Even if a physical support for the webcam was provided (or built by the teacher), the teacher would still need to ensure that the system is calibrated and ready for use. This reflects a trade-off between integration and minimalism: on the one hand the system is better integrated in the classroom because it uses hardware that is already present in it, but on the other hand the effort needed from the teacher to set it up is larger.

Thanks to the keyboard and mouse, there is no need for the cards, and the number of components of the interface is therefore smaller. This is good for minimalism, as the interface has fewer components, but damages empowerment, because part of the interface cannot be given to the teacher as easily. One way to increase empowerment would be to reintroduce cards for some actions, or to have the same actions be protected by a password.

8.1.4 Usability study

Goal and setup

The goal in scaling up TapaCarp is to implement all the activities that were featured in TapaCarp, but for the purpose of testing e-TapaCarp, we first implemented the matching model activity described in page 61. As a reminder, in this activity, one or several blocks form a "model", and the user's task is to find the position of each block of the model based on the front and side views of this model. We refer to each model that has to be found as a "challenge". The challenges can be generated either by the system, or by a human.

We conducted a first study with four colleagues. Some were more familiar with TapaCarp than others. This study was done in two phases. The three first users used the same version of the system. We then implemented a second version based on the comments of the three first users, and had the fourth user test this new version. The goal of the study was to test the usability of the website and the clarity of the instructions on how to solve the challenges, and we therefore gave the tangible interface to the participants instead of asking them to build it.

Before they could try and solve challenges, users needed to set up the workspace and webcam. A tutorial explained how to do that and how to complete the activity, since users were not expected to be familiar with orthographic projections. The tutorial included a practice challenge, and once they solved it, users were asked to create 3 challenges, with 1, 2, and

3 blocks, respectively. The three blocks that could be used for the challenges are shown in Figure 8.2. When done with the creation of the challenges, users could start solving challenges available in the system. The available challenges were those created by other users, as well as 6 challenges that were pre-loaded in the system so that even the first users had some challenges to solve. The time taken to solve a challenge was used to rank the performance of the users, and an overall ranking was computed based on the rank on each challenge. This ranking adds a gamification aspect to the system, which can increase motivation.

All four users were given the task to complete the tutorial, including to try to solve a challenge. We told them that the goal was to improve the usability of the website, not to test them; we asked them to think aloud and told them that every suggestion would be welcome.

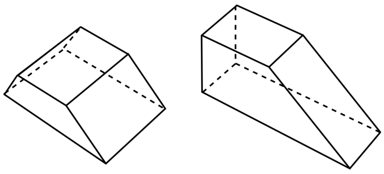


Figure 8.2: The blocks of the interface.

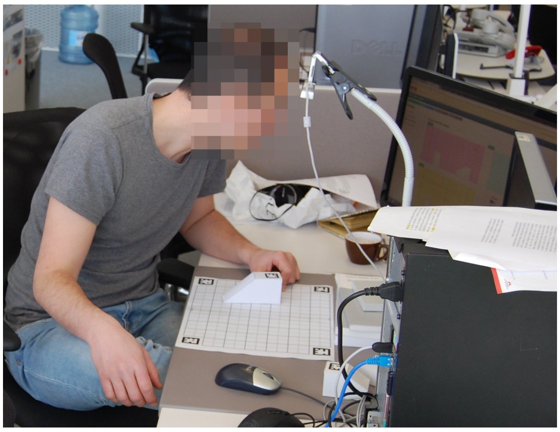


Figure 8.3: Having a support for the webcam helped this user position it.

Results

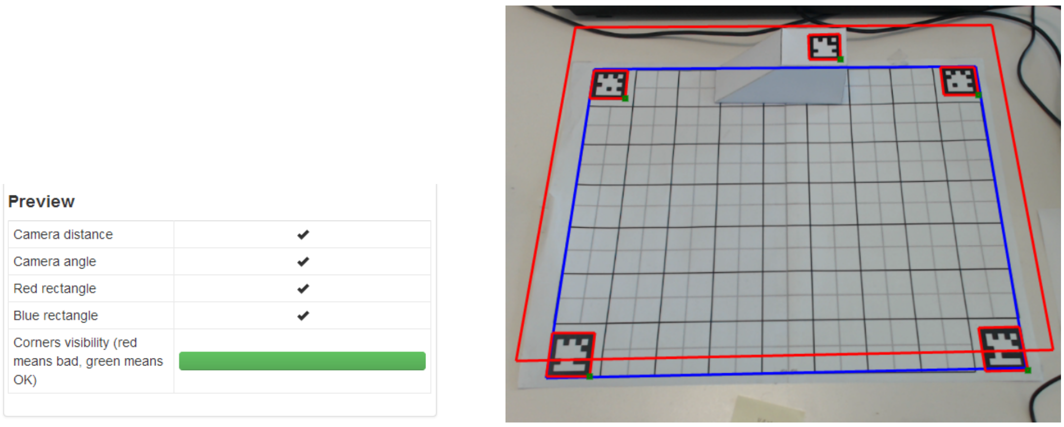
The first participant struggled mainly on positioning the webcam. He first tried to hold the webcam in his hand, thinking that he could use the website while holding the webcam in one hand. When it became clear that this was not realistic, he struggled to find a place for it and ended up putting it on the top of his screen. He suggested that we add pictures to illustrate possible webcam positions. This participant was never really sure what to do, he recurrently wanted more detailed explanations. He also struggled with the challenges, and could not solve a single one correctly, because he did not understand why his solutions were not correct.

The second participant faced different problems. He had a tripod with a clip on it (Figure 8.3), so it was easier for him to place the webcam. His main concern was how to navigate through the tutorial. He often asked where he should click. Unlike the first participant, he was familiar with the orthographic projections and could solve challenges in the practice phase of the tutorial.

Like the first participant, the third one had trouble placing the webcam and tried to keep it in his hand, then found a stable position for the webcam that was too far from the workspace to be able to detect the markers. He then tried to put the workspace on the ground and the webcam on the edge of the tabletop, but soon realized this was unpractical because he could not see his screen anymore. This participant skipped over the textual explanations, and only read them when he did not know what to do next. He was able to solve some challenges in the practice phase.

Improvements

The usability study revealed three issues: the placement of the webcam, the navigation in the tutorial, and the lack of feedback in the practice of challenges. These issues were addressed in a second version of the system. We added pictures of how to place the webcam and dynamic feedback that showed the view that the webcam has of the workspace and indicates how well it detects the markers. Feedback was also added to the practice phase to facilitate solving challenges, and the navigation was made more intuitive by adding a pop-up window at the end of each challenge, with a button allowing to directly go to the next step. A fourth participant tested this new version and did not encounter any of the problems met by the other participants. With so few participants, it is hard to judge if the last participant struggled less because of the improvements that were made to the website, or if she would have succeeded as well with the previous version. However, after testing with four participants, we had a good sense of what was most challenging for the users, and that was the placement of the webcam. To improve this, we added dynamic feedback to help users make sure that the position of their webcam was optimal.



(a) The 5 types of feedback given to the user.

(b) Illustration of the red and blue rectangle.

Figure 8.4: Feedback given to the user on the webcam positioning: the blue and red rectangle, graphically, as well as the 5 indicators.

This feedback came in the form of blue and red rectangles superposed to the webcam image, and displayed on the screen, and of three additional metrics: distance and angle of the webcam, and detection of the 4 markers (Figure 8.4). The user therefore gets a feedback on whether or not the webcam is placed correctly, but also on how to improve its placement. If all 5 metrics are validated, the user is sure that the webcam position will not cause problem. The reason for limiting the distance and the angle of the webcam have been explained before. Detecting the 4 markers simultaneously is needed to compute the homography between the workspace and the webcam referential. The blue rectangle shows the workspace contour and should be completely visible. It gives the user an easy way to estimate how accurate the calibration is and how centered and visible the workspace is. The red rectangle gives a visual feedback to the user of how large a margin they should leave between the workspace and the edges of the webcam image. The margin depends mainly on the height of the tangibles used, as shown in Figure 8.4b: the higher the marker, the larger the margin.

Finally, to allow the users to get a rapid global understanding of what the project was about and what the interface allowed them to do, we added a 3 minute introductory video, in both French and English (Video 4). The video shows the users how to place the webcam such that it works and helps them get started without spending too much time reading.

8.2 Testing the E-TapaCarp on a Larger Scale

8.2.1 Goal of the study

The goal of e-TapaCarp is to demonstrate that it is possible to build a low-cost and scalable TUI. The design of e-TapaCarp makes it a cheap system for someone who already owns a computer. The results of the usability study highlighted some usability issues that were addressed in an additional design iteration. However, the focus of this first study was on the usability of the website and it did not require the users to build the tangible interface, which is a significant part of the scale up. We were curious to see if people outside our laboratory could successfully complete the whole process of building the interface and use it to solve spatial skills exercises.

8.2.2 Participants and diffusion

Participants who are asked to test an interface need to have an incentive to do so. The incentive can be extrinsic, such as when participants are paid or, for students, get bonus points or course credits. It can also be intrinsic when people have a genuine interest in the topic of the study. In the present case, we created an extrinsic motivation by offering Amazon vouchers of 100, 50, and 20 US dollars to the participants who ranked in the first 3 places of our contest. Ideal participants would have been carpenter apprentices, because they have an intrinsic motivation to get better at orthographic projections. However, the timing of the study (fixed because of Claude's own deadline) clashed with the end of year obligations of the apprentices, and it was not possible to enroll them.

Despite the risk of not getting intrinsically motivated participants, we decided to advertise the website with the hope that we would nonetheless learn something from a first deployment. We sent e-mails to one hundred friends and colleagues, advertised through social networks, and posted a link on Hacker News[2], a news aggregator where people can post a URL to "anything that gratifies one's intellectual curiosity". This obviously makes for a very biased technology savvy sample of participants, and we were aware of this limitation, but again, we thought that we could still learn a lot from this first deployment. A second study with a more balanced sample could be done later.

8.2.3 Task

When connecting to the homepage, users could see the introductory video or start the tutorial by logging in. The tutorial explained how to build the interface, set up the webcam, and how to solve challenges and gain points. We distinguish three kinds of participants: contestants, who participated in the contest; subscribers, who logged into the website but did not complete a single challenge; and visitors, who came to the website but did not subscribe. Because participants were volunteers, we did not require from the participants to complete a specific task, or ask them to spend a minimal amount of time interacting with the system. Instead, we simply encouraged them to try the interface and highlighted the possibility to win prizes.

8.2.4 Data gathering

During the deployment, data was gathered through the logs of the application, through Google Analytics, a web navigation analysis tool, and through questionnaires that were sent to the subscribers (Section A.7.1) and to the contestants (Section A.7.2) towards the end of the study.

8.2.5 Results

Participation and navigation pattern

The study lasted 25 days, from May 21 through June 14, 2013. During this time, Google Analytics recorded 339 unique visitors (615 visits in total), out of which 48 (14.2%) subscribed to the course, and 6 (1.8%) completed the tutorial. At first sight, this appears as a small subscription and usage rate, but the ratio of users to the number of visits (6 for 615, or 0.98%) is actually similar to the one reported by Costanza et al. (273 active users for 25'000 visits, or 1.09%), except on a smaller scale. More than half of the visitors (57%) left without having seen anything but the homepage. This could indicate a lack of interest for spatial skills training. When clicking on the "Start" button (the only choice on the homepage, besides watching the video as shown in Figure 8.5), participants were taken to a login page, at which another 35% left the website. In the end, only 8% of all visitors reached the first step of the tutorial, and less than 3% pursued past the first page describing how to build the TUI.

[2] http://news.ycombinator.com/

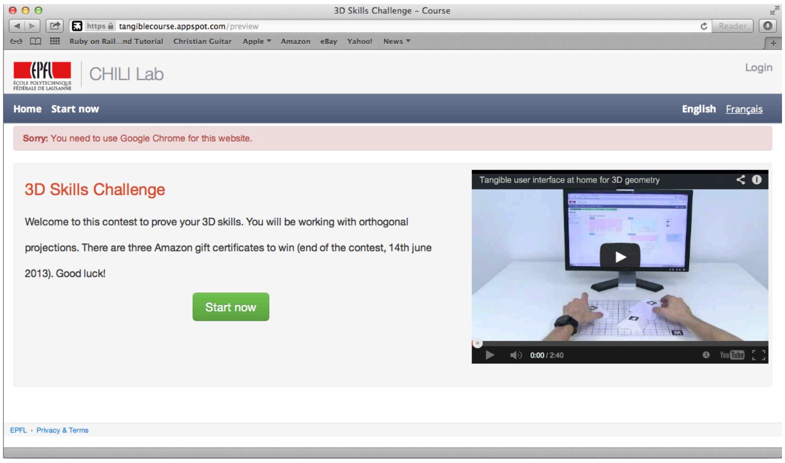


Figure 8.5: Screenshot of the homepage of the website.

Costanza et al. observed a similar pattern in the visits of their website and blamed the lack of participation on the incompatibility of hardware. Our results show that this was not the case here: while there was some hardware requirements, 92% of the visitors did not even reach the page explaining these requirements. A more likely explanation is that participants were interested neither in spatial skills training, nor in tangible user interfaces, and did thus not see the benefit of spending more time on this website.

During the usability study, navigating the tutorial was sometimes problematic. During this deployment, an analysis of the navigation pattern revealed that participants went back and forth mostly in the first phase of the tutorial, i.e. when building the interface and setting up the webcam. Once the interface was built, the users navigated as expected. This again indicates that setting up the webcam was the most difficult step.

Contestants results

A total of 21 challenges was available to each contestant (15 from other participants, 6 from us). Table 8.1 gives basic information on the challenge completion. On average, each participant tried 13 challenges (60.3% of the available ones), and only one contestant tried all the available challenges. On average, participants spent about 20 minutes solving challenges (SD: 10 minutes), with the most persistent user spending 36 minutes, and the least persistent one 7 minutes and 30 seconds.

Most participants worked on the challenges on a single day. This hints that they were more interested in trying out the interface than really winning the contest, or developing their spatial skills. However, there was one exception: a 13 years old boy, who worked on the challenges over three days, and tried solving all of them. This participant was also the only one to do practice challenges between the contest challenges and to report that he wanted to improve his spatial skills.

	Number of objects			
	1	2	3	Total
challenges	8	8	8	24
challenges tried	29	24	23	76
participation	69.0%	57.1%	54.8%	60.3%
mean time in seconds	40	132	185	112
abandonment	0.0%	12.5%	17.4%	9.2%

Table 8.1: Data on challenge trials.

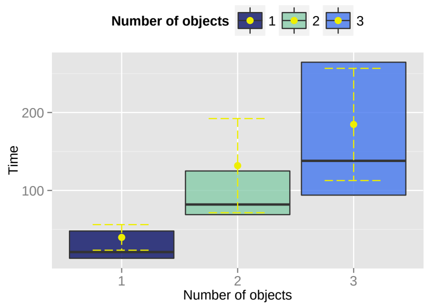
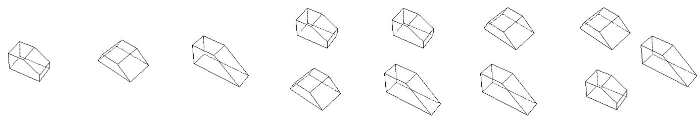


Table 8.2: Time taken (in seconds) for challenges with 1, 2, and 3 objects.

Influence of the shape. The difficulty of the challenges, estimated by the time taken to solve a challenge, varied according to the shape of the object, as shown in Table 8.3. For challenges with one object, the object listed in the second position in Table 8.3 gave significantly more difficulties to the users than other objects (F[1,27]=5.24, p=.03). Challenges with two objects, that included this object, also took more time, although not significantly (F[1,19]=.2, p=.65). The only challenges on which the participants abandoned (i.e. quit the page without having solved the challenge) were those that included this object. This echoes the result of Section 6.2.2 on page 76, which showed that difficulty increased with the number of symmetry planes of an object. Although this object is not perfectly symmetrical in its width, it is very close to be, making it easy to mix up the orientation of the object by 180°.

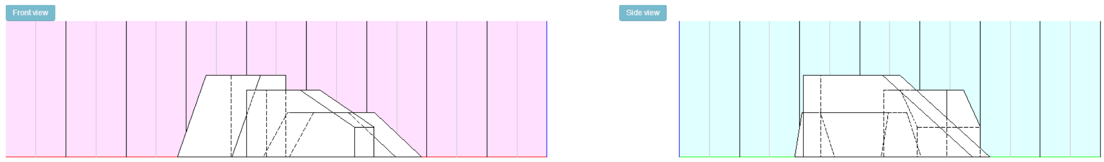


Object							
challenges	3	4	1	3	2	3	8
results	9	19	1	13	2	9	23
abandonment	0	0	0	2	0	1	4
Avg. time (in sec.)	15	52	26	126	90	151	185

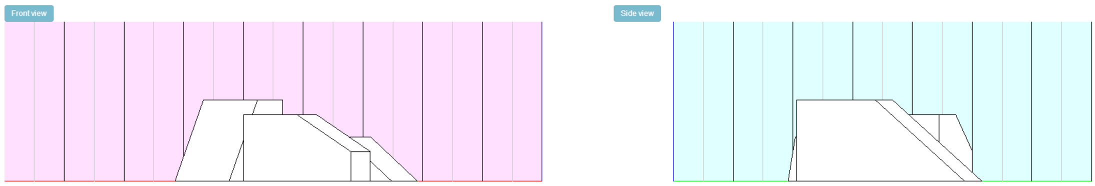
Table 8.3: Data for the challenges, by object.

Number of objects. Not surprisingly, the average time taken to complete a challenge increased with the number of objects (F[1,66]=10.6, p=.001, Figure 8.2). Solving a challenge with one object took on average 40 seconds versus 2 minutes and 12 seconds for challenges with two objects, and close to an additional minute for challenges with 3 objects (3 minutes and 5 seconds). The smaller duration and the lower rate of abandonment for challenges with a single object indicate that these were easier to solve. The difference between challenges with 2 and 3 objects is less evident, as shown by the non significance of the difference in the time needed to solve them (F[1,38]=1.4, p=.24).

Using transparency. As explained for TapaCarp in Section 5.3.4, changing the rendering of objects from solid to wireframe (i.e. transparent) can help when trying to interpret orthographic projections. The analysis of the log files revealed that switching the rendering type increased with the number of objects in a challenge. The transparency was turned off 11 times when there was 1 object, 25 times when there were 2, and 65 times when there were 3 ($F[1,64]=13.8$, $p<.001$). On some challenges, users alternated several times between the transparent and opaque mode. For example, out of the 65 utterances for challenges with 3 objects, 41 came from a single challenge. The front and side views of this challenge are shown in Figure 8.6, once with a transparent rendering, once with an opaque one. Because the objects are hiding each other on both views, the transparent rendering (Figure 8.6a) features a mix of stipple and full lines around the same area, and it is hard to guess the position of the objects. Switching to the opaque rendering (Figure 8.6b) makes it easier.



(a) Transparent rendering.



(b) Opaque rendering.

Figure 8.6: Example of when transparency can help.

The other option to customize the display was to hide the grid, but it was used only three times. This was expected since the grid is the only common reference between the three views and the workspace.

Questionnaire

Five of the six constestants filled the questionnaire (shown in Section A.7.2). The questionnaire data confirmed the bias in the sample: 3 participants had a strong IT background. As in the usability study, the main problem reported by the participants was the placement of the webcam. They also signaled that making the objects required additional material (cutter or scissors, glue, and tape) and that building shapes was difficult, but only one reported being frustrated with the whole process. Once the objects built, however, all were able to use the interface without problem. All the participants said that the challenges were fun and interesting. They could see the usefulness of such an activity, although only one participant (the 13 years old boy) said he was interested in developing his spatial skills.

8.2.6 Summary

This is only a first feasibility study, and it suffers from several limitations. One is the biased sample of participants, who were highly qualified and with a particular interest in IT and research on novel interfaces. Another limitation is the short duration of the deployment and the small number of participants, which resulted into a limited amount of data overall. A longer deployment with more participants would allow us to gather more data, such as the different strategies used in solving the challenges, the difference between high and low achievers, or the performance progression over time.

Less than 2% of all visitors took the time required to build the interface. This may be due to a lack of interest for spatial skills, as the participants did not fall into the target population of the system (e.g. carpenter apprentices). Still, 8% of all visitors reached the first page of the tutorial, but less than half of those went further, which is a rather large dropout rate. Two reasons can explain why these participants left the website despite an apparent interest: either they got discouraged by the effort and time that it would take to build the interface, or they did not have all the necessary material (especially the external webcam).

Despite the low participation, we could glimpse at some of the benefits that having an online tangible can bring to research on learning and on human factors. The data collected through online interaction might not be as rich as with a field study, but it is dramatically cheaper and easier to gather. Moreover, it can still bring insights on the learning task and the usage of the interface, as we saw with the analysis of the rendering options, and the influence of the shape and number of objects.

A positive outcome of the study was that the participants who built and used the interface (i.e. the contestants) gave a very positive feedback on its usability. All were able to create and solve challenges, and said that they did not have any problems following the tutorial. The main problem was the placement of the webcam, most probably because of a lack of options in the physical positions available on a desk.

8.3 Discussion and Conclusions

Despite its limitations, this study showed that “bringing Ubicomp to the masses”, to paraphrase Costanza et al., is not only a matter of lowering the cost and providing an interface for download. It also involves getting people’s interest, and convincing them to dedicate the effort and time necessary to build the interface and get the required hardware. There are ways to reduce the effort of building the interface, such as using a flat paper interface, objects that can be purchased in stores, or object generated by 3D printers. But the burden of getting the webcam and setting it up is not as easily lightened. One solution would be to provide users with a package containing the webcam and a physical support for it at an affordable price, but this would be a setback to the idea that anyone could use a TUI anywhere. Another way to “solve” that would be to go to a different application domain in which accuracy is not as important as in our case. What would then matter is whether an object is present and its relative position to other objects present in the scene, as was the case in the AR-Jam book, for example.

Perhaps the biggest obstacle into building this setup is the impossibility for the user to get a working demonstration of the system quickly. The demonstration video was a step in this direction, but nothing replaces being able to try the interface. The time that a user is willing to invest before being able to use the system remains to be evaluated, and will depend on the motivation of the user and on the context in which the system is used. But in any case, the setup cost of a TUI is likely to remain higher than that of a regular GUI application.

We started developing e-TapaCarp with the vision that apprentices could use it outside the classroom. The practical obstacles mentioned above suggest that this might not be a realistic scenario. Still, coming back to learning in the classroom, e-TapaCarp is promising and opens up new perspectives. If the interface was used regularly through the year and shared between students, the effort and time to build it would appear less significant. The teacher could create the interface and a support for the webcam at the beginning of the year, together with the rest of the school material. Thus, schools could get a working TUI to train apprentices at a fraction of the cost of TapaCarp.

Of course, studies still need to be conducted to assess how the differences between TapaCarp and e-TapaCarp impact the learning performance. The pecuniary gain brought by e-TapaCarp is obvious, but research is needed to evaluate whether this results in some loss in the learning experience and outcome. Several aspects should be considered, such as the split of attention that can result from the separation of the input and the output spaces, and the lack of augmentation on objects in the input space. E-TapaCarp might also bring advantages in terms of learning, for example by allowing new types of activities thanks to the use of the keyboard, or by creating activities with a stronger social component thanks to the web connection.

9 Discussion and Conclusions

9.1 Summary and Contributions

9.1.1 Summary

As explained in Section 2.4, there are claims that TUIs could benefit learning in many ways, but little empirical demonstration to support those claims. There is also a lack of studies on how TUIs can be integrated in classrooms, or used on a larger scale. The goal of this thesis is to contribute to solving these two issues. It does so in the specific context of training carpenter apprentices' spatial skills, and by answering the following research questions:

Q_F Which features of a TUI support the acquisition of spatial skills?

Q_{PS} What classroom pedagogical scenarios do TUIs support for the training of spatial skills?

After a review of TUIs (Chapter 2), we described in more details the context of carpenter apprentices (Chapter 3). We summarized the work to date on spatial skills, and did a first study that showed that carpenters indeed have well-developed spatial skills (Chapter 4). A TUI, TapaCarp, was developed, that aimed at improving carpenters' spatial skills, and especially the link between the 2D and 3D representations of an object (Chapter 5). Four empirical studies conducted in classrooms (Chapters 6 and 7) and one online study (Chapter 8) yielded the following results, which address the challenges and open questions described in Section 2.4:

1. TUIs can help train spatial skills. However, design changes, even minor ones, can significantly impact the learning experience.
   (a) **The shape of a tangible plays a key role for spatial skills tasks.** When solving spatial tasks involving a virtual object, the shape of the tangible that allows the manipulation of the virtual object matters. A tangible with the same shape as the virtual object (literal physical correspondence) can help solve spatial tasks, especially difficult ones (Section 6.2).
   (b) **Dyna-linking is not always good for learning.** The presence of a dynamic link (dyna-linking) between the tangible object and its virtual representation(s) encourages students to explore more and reflect less. Removing the dyna-linking makes

the task harder, but forces students to think more (Section 6.3).

2. **Pedagogical scenarios.** We showed 3 ways in which a TUI can be used in a classroom: (1) as a small task that can typically be done in parallel to the main learning activity; (2) to support the main classroom learning activity, done only partially on the TUI; and (3) to support the main classroom learning activity, done exclusively on the TUI. The most promising approach was the second one, because the benefits of the tangibles carried over beyond the usage of the TUI, helping the TUI integrate with the rest of the activity, and the entire activity to integrate better in the classroom (Section 7.2).
3. The two previous points can help create a learning-effective technology and incorporate it in the classroom. However, introducing technology in a classroom can lead to an undesirable increase of the classroom orchestration load. We introduced **conceptual tools that can help reduce orchestration load**.
   (a) **Design principles.** Integration, flexibility, empowerment, awareness, and minimalism, are 5 design principles that can help reduce the orchestration load and, consequently, facilitate the adoption of TUIs in the classroom (Section 7.1).
   (b) **Hutchins threshold.** Distributing the interface of a TUI over several tangible components can help integrate a TUI in a classroom and reduce the classroom orchestration load. However, after a certain point, the distribution becomes harmful. This point, which we called the Hutchins threshold, cannot be computed, but can help to conceptualize the design of TUIs (Section 7.1).
4. **Low-cost TUIs**. We built a prototype that runs online and for which the only hardware required is an external webcam and a printer. The results of a first user study indicated that low-cost TUIs could be a potential solution for schools, but that making TUIs available easily to anyone ("personal TUIs") would be harder to achieve. However, more research is needed to evaluate the impact that using such a low-cost TUI, instead of a regular one, can have on learning (Chapter 8).

9.1.2 Contributions

The results presented in this thesis contribute to three areas of research: training spatial skills, learning with TUIs, and classroom orchestration and technologies.

Training spatial skills

As we saw in Chapter 4, the interest for spatial skills is increasing because of their influence on the development of expertise in STEM fields. Spatial ability varies greatly between individuals, and the view that it needs to be specifically trained is gaining traction among educators. The results of this thesis can contribute to designing technology to develop spatial skills. Spatial tasks can be difficult, especially for novices. Our results showed that tangibles could help learners when solving difficult tasks, such as mapping a 3D view of an object with a rotated 2D projection of the same object.

The benefits of tangibles for spatial skills are linked to perception and the ability of the human brain to quickly capture physical distances and shapes. Our results indicated that tangibles help the learner build a mental representation of an object, which is the cornerstone of many spatial tasks. Tangibles can also serve as an anchor when exploring the links between the multiple representations of an object. This is useful for spatial tasks, in which many representations of the same object are often involved (e.g. orthographic projections).

Another set of benefits of tangibles is the easier manipulation of objects. This is neither a new finding, nor one specific to spatial skills, but it can be of particular importance for spatial tasks. Indeed, when the learning task is difficult, an easier manipulation can help because the learner's cognitive effort can be focused on the task rather than on the manipulation of the objects.

Learning with TUIs

As explained in Chapter 2 there are doubts about the impact of TUIs on learning. Although we were not able to show it with pre-test and post-test, our studies indicated that the physicality of tangibles can help in developing one's spatial ability, for the reasons mentioned in the previous paragraph.

Of course, this does not mean that TUIs can improve learning in any topic, or for any kind of learning. What a TUI can bring to a given topic is highly dependent on the topic itself. For instance, one of the questions for this work was to see whether, in the continuity of the TinkerLogistics (Zufferey et al., 2009), TUIs could help train apprentices in another profession than logistics. Although our approach was similar to that of Zufferey et al., the resulting technology differs in many ways and the benefits for learning are not the same. This underlines that there is potential for TUIs to support learning, but that TUIs need to be adapted to each learning domain that they target.

As pointed out by Shaer and Hornecker (2009), TUIs for learning often address the basics of a domain. They are generally seen as useful for beginners, and for a limited time only. The results of this thesis shed a new light on this in two ways.

The first one is that TUIs can be used for more than just learning the basics. For example, the activity used in Chapter 8 showed that it was possible, with a simple activity, to create difficult tasks that required a thorough understanding of orthographic projections. In this case, the complexity is linked to the task and the object rather than the number of objects. One could argue that orthographic projections themselves are the basics of spatial skills, and that any level of understanding of spatial skills is basic. However, we showed through the activity presented in Section 7.2 that tangibles could also be used for activities that built upon the knowledge of orthographic projections.

Second, a careful design of the learning activities can help TUIs address the need of more advanced students. The two design variations on dyna-linking and the shape of the tangibles (Chapter 6) showed that it was possible to dramatically change the difficulty of a learning task by slightly altering the design of the interface. This could be used to build activities of increasing difficulties, which would have the benefit of keeping the learners interested for a longer period of time, and potentially for them learn more.

Another contribution was the confirmation, in the experiment on dyna-linking presented in Section 6.3, that an increased usability does not necessarily benefit learning. Others have made the point before that the increased usability of TUIs does not necessarily benefit learning (Do-Lenh, 2012; Do-Lenh et al., 2010; Price et al., 2009), and that the design of a TUI can significantly impact the learning processes and outcome (Manches et al., 2010). The presence or absence of dyna-linking led to different strategies, hinting that TUIs could be used for different phases of learning. For example, dyna-linking could be switched on during an exploration phase, and switched off later to foster more reflection.

A final contribution towards the field of TUIs for learning is the low-cost TUIs. Many existing TUIs require costly dedicated hardware, and this cost hampers the deployment of TUIs in schools. New standards in web technologies make it possible to develop a new kind of low-cost TUIs that reuses the computers already present in most classrooms. To show this, we used cutting edge web technologies to build a working low-cost version of our TUI. Preliminary tests indicated that this could be a viable option for schools. However, research is still needed to evaluate the cost of the pecuniary gain in terms of learning performance (see below).

Classroom orchestration and classroom technologies

Research on learning with TUIs has not focused much on the integration of TUIs in classrooms. This is a pity, because the best place for a TUI for learning to be used is probably the classroom. Integrating technology into a classroom is no trivial task, as the classroom is a complex environment with many constraints. The recent theory of classroom orchestration (Dillenbourg et al., 2012) acknowledges the need to satisfy these practical constraints when introducing technology into the classroom, and that failing to do so can increase the classroom orchestration load (Dillenbourg et al., 2011).

The TUI presented in this thesis was used in a classroom in three different scenarios: (1) as a small task that can typically be done in parallel to the main learning activity; (2) to support the main classroom learning activity, done only partially on the TUI; and (3) to support the main classroom learning activity, done exclusively on the TUI. To the best of my knowledge, it is the first time that a TUI is used in such different classroom scenarios.

We did not study the first scenario in terms of classroom orchestration. For the third scenario, the results indicated that entirely supporting the main activity (with 5 instances of the TUI in a classroom) can result in an increased classroom orchestration load. The design of the TUI should therefore be adapted to limit the orchestration load. To this end, we introduced 5 design principles, and for the case of distributed interfaces (TUIs often *do* have a distributed interface), we also introduced the concept of Hutchins threshold to help find the right balance between a sufficient distribution of the interface, and too much distribution.

Based on our experiments, the most promising scenario to integrate TUIs in the classroom is to use them as part of an encompassing learning activity (scenario 2). Using a TUI allowed us to remove the menial tasks of the activity (drawing construction lines) and have students focus on the added-value part of the task (connecting these construction lines). Using the TUI only at the beginning allowed us to stay close to their practice (drafting a plan on paper and building an object based on this plan) and therefore create a meaningful activity. Using an object (the polystyrene block) in both the TUI and outside of it helped students link the various steps of the activity together, and so did the plan that they created with the TUI but kept all along the activity. These two objects also increased the teacher's empowerment and awareness by helping him to give contextualized feedback to the students at any time of the activity .

9.2 Limitations

I was lucky to be able to run many studies in classrooms. Although this offers greater ecological validity, it also implies that many parameters (e.g. students' prior knowledge, mood, or motivation) cannot be controlled. One of the consequences is that the learning gain due to the intervention is hard to assess. Another is that, sometimes, the experimental design needs to be adapted to fit practical constraints, or that an experiment does not go as planned. The study presented in Chapter 5, where a researcher had to play the role of the teacher in the experimental condition while the teacher was teaching in the control condition, is an illustration of this.

Evaluating the learning gain is difficult for other reasons as well. I used pre-test and post-test, an approach that suffers from several limitations. It is hard to make sure that the tests actually measure the skills (or set of skills) that was supposedly developed during the treatment. Moreover, the questions of a tests usually need to be fine-tuned so that they complement each other and allow a fine-grain assessment of the knowledge and understanding of students. To be most informative, tests should be long, but when run in classrooms, they cannot be longer than 10 to 15 minutes. Longer tests are especially useful when the learning task is complex and there can be many types of improvement, such as in Chapter 7. The tests also highly rely on the motivation and state of mind of students, which we cannot control. When testing several conditions, it can be hard to design a test that does not give an advantage to one or the other condition. In fact, in our experiments, all tests were done on paper, forcing the students who used the TUI during the treatment to perform a transfer, and possibly hurting their test scores in the process. An interesting future area of research would be to use TUIs to develop new kinds of assessment, as has for example been done by Sharlin et al. (2002).

Although running studies in the classroom increased the ecological validity, it could have been increased more. Indeed, in every study, one or several researchers were involved, which would not be the case if TapaCarp were to be used further in classrooms.

There are many ways to design studies to evaluate a learning technology. I did only short interventions and did not, for example, perform any longitudinal study. This was a deliberate choice made to promote the design of new learning scenarios instead of a single, longer one. However, it means that I could not evaluate the long-term impact of the usage of TapaCarp in depth for one given scenario. Similarly, there are many design aspects that I chose not to explore. For example, more studies could be done to understand better the impact of the shape of the block on the learning performance.

It is an open question how much the findings of this work can be generalized to other contexts. This work focuses on carpenter apprentices, and all participants to the studies came from this population. A greater variety in the population may have brought a greater variety in the results. In the same vein, what I referred to as “spatial skills” was in fact always linked to orthographic projections, and although these are an important topic in spatial skills, there are other subjects in spatial skills that could be explored.

In all the classroom studies, the apprentices worked in pairs, as is common practice at school. However, we did not focus on the collaboration between the two members of a pair. Collaboration can serve as a mediator to explain the individual learning outcome measured through the pre-test and post-test, and would therefore be worth studying in more detail in the future.

9.3 Design Implications

One lesson from this research is that when designing a learning technology for the classroom, one should integrate the classroom constraints from the start. In this work, although the teacher had been involved from the start, he refused to use our technology during one of his lessons, because our activities did not fully match the curriculum and existing practices. The five design principles that were introduced (integration, flexibility, empowerment, awareness, minimalism) should help prevent such disappointments in the future.

As researchers, designers, or computer scientists, our focus tends to be on the technology, and we use some learning activities to show off the features of the technology. But in a real classroom, the most efficient way to use the technology might be to use it just sparingly. Of all the activities presented in this thesis, the most successful one used the TUI for about one fourth of the time only. In the same way, some “dumb” design mechanisms might be more powerful and easier to use than fancy features, such as giving a paper card to the teacher instead of implementing a password-protected set of administration functions.

Our results showed that TUIs can be used for advanced learning domains. There are no doubt some constraints to TUIs, such as the limited space available, which in turn limits the number of objects that can be used simultaneously. But this does not mean that TUIs cannot be used for learning in more complex domains or for more complex tasks, it just means that we need to be creative when designing learning activities with them.

Design variations can also help in modifying the level of difficulty. TUIs are a great way to lower the barrier to entry for novices, but it does not mean that this is all they have to be. To encourage learning, we should make sure that TUIs allow the level of difficulty of the activities to rise with the learner's understanding. This is possible through slight design variations. For example, having a literal physical correspondence between the tangible object and its virtual representations can help a novice understand orthographic projections. However, what we really want to do in this case is to encourage the learner to build a mental image of the block. Having the tangible of literal physical correspondence at disposal may help do this at the beginning, but the activity should be designed so that students learn to do that even once the tangible is removed or replaced with one of a symbolic representation.

9.4 Future Research Perspectives

This thesis is a small drop in the ocean of research on learning with TUIs, and much work remains to be done. One future research direction is to study in more detail what exactly about tangibles helps train spatial skills. There is already much literature on spatial imagery (see e.g. Schwartz and Heiser, 2006). However, tools such as eye-tracking could help discover in finer ways the difference between various physical representations of a tangible object. For example, what would be the impact of using a wireframe version of the block instead of a full wooden block? It would also be interesting to observe in more details how the various tangible representations impact the collaboration between groups of users.

This thesis focused on the training of spatial skills, and we found that for this particular domain, TUIs could be useful. This builds upon the work of Zufferey (2010) and Do-Lenh et al. (2010) who showed how TUIs could be used to train logistics apprentices. In the future, it would be interesting to investigate further what other topics can benefit from the usage of TUIs. Would TUIs show the same kinds of benefit in other domains? Would the benefits brought by TUIs vanish for topics that do not have a strong spatial components? Would other benefits surface?

I believe that there is great potential for developing further the training of spatial skills in the domain of carpentry. One skill that has been added to the new curriculum of carpenter training is the use of CAD software. Using CAD software is hard because it requires prior knowledge (especially in geometry) and because of the complex interfaces that CAD software feature. A TUI to help learn to use CAD software would be of great use. The advantage of using a TUI in this case would be to simplify greatly the interface and allow the learner to focus on the procedures to, say, create a cut and generate the plan.

One area in which there is still many things to uncover is classroom orchestration. In particular, it would be interesting to find out what other variables affect the success of a learning activity with a TUI in the classroom. For example, Schneider et al. (2013) have recently shown that using a TUI before receiving formal instruction was more effective for learning than using a TUI after having received the formal instruction. Such findings are key to inform the design of pedagogical activities so that they can make good use of TUIs.

We showcased how building low-cost TUIs and making them largely available is easily done with the current web standards. With 3D printers becoming more affordable by the day, it is likely that the use of tangible interfaces will broaden. In our case, we replaced the camera and projector by a webcam, losing the possibility to augment the physical object, which can impact the learning experience. However, pico projectors are also becoming more mainstream, and future low-cost TUIs could use them too. In any case, research needs to be done to assess the impact of the differences between full-scale and low-cost TUIs in terms of learning.

Bibliography

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education.*

Ainsworth, S. (2006). DeFT: a conceptual framework for considering learning with multiple representations. *Learning and Instruction.*

Alavi, H., Dillenbourg, P., and Kaplan, F. (2009). Distributed awareness for class orchestration. In *EC-TEL 2009, Synergy of Disciplines.*

Alcoholado, C., Nussbaum, M., Tagle, A., Gomez, F., Denardin, F., Susaeta, H., Villalta, M., and 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., and Yedidia, J. S. (2000). Tangible interaction + graphical interpretation: a new approach to 3d modeling. In *Proc. SIGGRAPH 2000.*

Anderson, J. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science.*

Antle, A. N. (2012). Knowledge gaps in hands-on tangible interaction research. In *Proc. ICMI 2012.*

Antle, A. N. and Wise, A. F. (2013). Getting down to details: Using theories of cognition and learning to inform tangible user interface design. *Interacting with Computers.*

Antle, A. N., Wise, A. F., and Nielsen, K. (2011). Towards utopia: designing tangibles for learning. In *Proc. IDC 2011.*

Arnheim, R. (1980). A plea for visual thinking. *Critical Inquiry.*

Bachour, K., Kaplan, F., and Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies.*

Bishop, J. (1978). Developing students' spatial ability. *Science Teacher.*

Blaye, A. and Light, P. (1990). Computer-based learning: The social dimensions. In Foot, H. C., Morgan, M. J., and Shute, R. H., editors, *Children helping children.*

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

Bonnard, Q., Legge, A., Kaplan, F., and Dillenbourg, P. (2012). Hands-on Symmetry with

Augmented Reality on Paper. In *Proc. HSCI 2012*.

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

Bossy, C. (2013). Scaling up a Tangible User Interface for Education. Technical report, EPFL.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*.

Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A., and de Ridder, H. (2005). pOwerball: the design of a novel mixed-reality game for children with mixed abilities. In *Proc. IDC 2005*.

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

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*.

Buechley, L., Eisenberg, M., Catchen, J., and Crockett, A. (2008). The LilyPad arduino: using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. In *Proc. CHI 2008*.

Burton, M., Brna, P., and Treasure-Jones, T. (1997). Splitting the collaborative atom: How to support learning about collaboration. In *AI in Education*.

Butterworth, G. and Light, P. (1982). *Social Cognition*. University of Chicago Press, first edition edition.

Carter, C., Larussa, M., and Bodner, G. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *Journal of Research in Science Teaching*.

Clariana, R., Wagner, D., and Roher Murphy, L. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development*.

Contero, M., Naya, F., Company, P., Saorin, J. L., and Conesa, J. (2005). Improving visualization skills in engineering education. *IEEE Computer Graphics and Applications*.

Costanza, E., Giaccone, M., Kueng, O., Shelley, S., and Huang, J. (2010). Ubicomp to the masses: a large-scale study of two tangible interfaces for download. In *Proc. Ubicomp 2010*.

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

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

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

Cuendet, S., Jermann, P., and Dillenbourg, P. (2012c). Tangible interfaces: when physical-virtual coupling may be detrimental to learning. In *Proc. British HCI 2012*.

Dillenbourg, P. (1999). What do you mean by collaborative learning? *Collaborative-learning: Cognitive and computational approaches.*

Dillenbourg, P. (2002). Over-scripting CSCL: the risks of blending collaborative learning with instructional design. In Kirschner, P., editor, *Three worlds of CSCL. Can we support CSCL.*

Dillenbourg, P. (2008). Integrating technologies into educational ecosystems. *Distance Education.*

Dillenbourg, P., Baker, M. J., Blaye, A., and O'Malley, C. (1995). The evolution of research on collaborative learning. *Learning in Humans and Machine: Towards an interdisciplinary learning science.*

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

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

Dillenbourg, P. and Jermann, P. (2010a). Technology for classroom orchestration. In *New Science of Learning.*

Dillenbourg, P. and Jermann, P. (2010b). Technology for Classroom Orchestration. In Khine, M. S. and Saleh, I. M., editors, *New Science of Learning.*

Dillenbourg, P., Järvelä, S., and Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In Balacheff, N., Ludvigsen, S., Jong, T., Lazonder, A., and Barnes, S., editors, *Technology-Enhanced Learning.*

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

Dillenbourg, P., Zufferey, G., Alavi, H., Jermann, P., Do-Lenh, S., Bonnard, Q., Cuendet, S., and Kaplan, F. (2011). Classroom orchestration : The third circle of usability. In *Proc. CSCL 2011.*

Dimitriadis, Y., Asensio-pérez, J. I., Hernández-leo, D., Roschelle, J., Brecht, J., Chaudhury, S. R., Digiano, C., and 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.

Do-Lenh, S., Jermann, P., Cuendet, S., Zufferey, G., and Dillenbourg, P. (2010). Task performance vs. learning outcomes: a study of a tangible user interface in the classroom. In *Proc. EC-TEL 2010.*

Do-Lenh, S., Jermann, P., Cuendet, S., Zufferey, G., and Dillenbourg, P. (2010). Task performance vs. learning outcomes: a study of a tangible user interface in the classroom. In *Proc. EC-TEL 2010.*

Doise, W., Deschamps, J., and Mugny, G. (1991). *Le conflit socio-cognitif.* Armand-Colin.

Dourish, P. (2004). *Where The Action Is: The Foundations Of Embodied Interaction.* MIT Press.

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests.* Educational Testing Service.

Fiala, M. (2005). Artag, a fiducial marker system using digital techniques. In *Proc. CVPR 2005.*

Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing.*

Fitzmaurice, G. W. (1996). *Graspable User Interfaces.* PhD thesis, MIT.

Fitzmaurice, G. W. and Buxton, W. (1997). An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In *Proc. CHI 1997.*

Flusberg, S. and Boroditsky, L. (2011). Are things that are hard to physically move also hard to imagine moving? *Psychonomic bulletin & review.*

Fraisse, J. (1987). Etude du rôle perturbateur du partenaire dans la découverte d'une stratégie cognitive chez des enfants de 11 ans en situation d'interaction sociale. *Bulletin de Psychologie.*

Gaillard, L. (2011). Perspectives de la formation - scénarios 2011–2020 pour le degré secondaire ii. Technical report, Office fédéral de la statistique.

Gaillard, L. (2012). Perspectives de la formation - scénarios 2011–2020 pour le degré secondaire ii. Technical report, Office fédéral de la formation professionnelle et de la technologie OFFT.

García, R., Quirós, J., Santos, R., González, S., and Fernanz, S. (2007). Interactive multimedia animation with Macromedia Flash in Descriptive Geometry teaching. *Computers & Education.*

Gardner, H. (2006). *Multiple intelligences: new horizons.* Basic Books.

Gardner, R. (2013). three.js.

Gerson, H. B. P., Sorby, S. A., Wysocki, A., and Baartmans, B. J. (2001). The development and assesment of multimedia software for improving 3D spatial visualization skills. *Computer Applications in Engineering Education.*

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

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., and Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology.*

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

Google (2012). Google course builder.

Gouchie, C. and Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology.*

Ha, V., Inkpen, K., Mandryk, R., and 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.*

Hambrick, D., Libarkin, J., Petcovic, H., Baker, K., Elkins, J., Callahan, C., Turner, S., Rench, T., and Ladue, N. (2011). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *J Exp Psychol Gen.*

Hamlin, A., Boermsa, N., and Sorby, S. (2006). Do spatial abilities impact the learning of 3D solid modeling software? *American Society for Engineering Education.*

Harel, I. E. and Papert, S. E. (1991). *Constructionism.* Ablex Publishing.

Hattie, J. (1999). Influences on student learning. *Inaugural lecture given on August.*

Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research.*

Hegarty, M., Keehner, M., Khooshabeh, P., and Montello, D. R. (2009). How spatial abilities enhance, and are enhanced by, dental education. *Learning and Individual Differences.*

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., and Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence.*

Heikkinen, I. (2013). JSARToolKit.

Holmquist, L., Redström, J., and Ljungstrand, P. (1999). Token-based access to digital information. In *Proc. HUC 1999.*

Horn, M. S., Solovey, E. T., Crouser, R. J., and Jacob, R. J. (2009). Comparing the use of tangible and graphical programming languages for informal science education. In *Proc. CHI 2009.*

Hornecker, E. (2012). Beyond affordance: tangibles' hybrid nature. In *Proc. TEI 2012.*

Hornecker, E. and Buur, J. (2006). Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proc. CHI 2006.*

Hornecker, E. and Dünser, A. (2009). Of pages and paddles: Children's expectations and mistaken interactions with physical–digital tools. *Interacting with Computers.*

Hornecker, E., Marshall, P., Dalton, N. S., and Rogers, Y. (2008). Collaboration and interference: awareness with mice or touch input. In *Proc. CSCW 2008.*

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science.*

Ishii, H. and Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI '97.*

Jacob, R. J. K., Ishii, H., Pangaro, G., and Patten, J. (2002). A tangible interface for organizing information using a grid. In *Proc. CHI 2002.*

Jermann, P. and Dillenbourg, P. (1999). An analysis of learner arguments in a collective learning environment. In *Proc. CSCL 1999.*

Kim, M. J. and Maher, M. L. (2008). The impact of tangible user interfaces on designers' spatial cognition. *Human–Computer Interaction.*

Kimura, D. (1994). Body asymmetry and intellectual pattern. *Personality and Individual Differences*.

Klahr, D., Triona, L. M., and Williams, C. (2007). Hands on what? the relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*.

Kluger, A. N. and DeNisi, A. (1998). Feedback interventions: Toward the understanding of a Double-Edged sword. *Current Directions in Psychological Science*.

Koleva, B., Benford, S., Ng, K. H., and Rodden, T. (2003). A framework for tangible user interfaces. In *Workshop Proc. Mobile HCI 2003*.

Kulik, J. A. and Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*.

Lafue, G. (1976). Recognition of three-dimensional objects from orthographic views. In *Proc. SIGGRAPH 1976*.

Larkin, J. H. and Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*.

Linn, M. C. and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*.

Maier, P. H. (1994). *Räumliches Vorstellungsvermögen*. Peter Lang Frankfurt.

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

Manches, A., O'Malley, C., and Benford, S. (2010). The role of physical representations in solving number problems: A comparison of young children's use of physical and virtual materials. *Comput. Educ.*

Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proc. TEI 2007*.

Marshall, P., Fleck, R., Harris, A., Rick, J., Hornecker, E., Rogers, Y., Yuill, N., and Dalton, N. S. (2009). Fighting for control: children's embodied interactions when using physical and digital representations. In *Proc. CHI 2009*.

Martín-Gutiérrez, J., Luís Saorín, J., Contero, M., Alcaņiz, M., Pérez-López, D., and Ortega, M. (2010). Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics*.

McGlone, M. S. and Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology*.

McGrath, J. E., Arrow, H., and Berdahl, J. L. (2000). The study of groups: Past, present, and future. *Personality and Social Psychology Review*.

McKim, R. H. (1980). *Experiences in visual thinking*. PWS Publishers.

McNeil, N. and Jarvin, L. (2007). When theories don't add up: Disentangling he manipulatives debate. *Theory Into Practice.*

Monge, G. (1798). *Géométrie Descriptive.* J. Klostermann fils, Paris.

Montessori, M. (1912). *The Montessori Method.* Frederick Stokes Co.

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

Newcombe, N. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator.*

Norman, D. A. (2011). *Living with Complexity.* MIT Press.

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

OPET (2012). Vocational and professional education and training in switzerland: Facts and figures. Technical report, Federal Office for Professional Education and Technology.

O'Malley, C. (1992). Designing computer systems to support peer learning. *European Journal of Psychology of Education.*

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* Basic Books.

Parkes, A. J., Raffle, H. S., and Ishii, H. (2008). Topobo in the wild: longitudinal evaluations of educators appropriating a tangible interface. In *Proc. CHI 2008.*

Patten, J. and Ishii, H. (2000). A comparison of spatial organization strategies in graphical and tangible user interfaces. In *Proc. DARE 2000.*

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

Pestalozzi, J. H. (1803). *ABC der Anschauung, oder, Anschauungs-Lehre der Massverhältnisse.* in Commission bey Heinrich Gessner, Buchhändler ; In der J. G. Cotta'schen Buchhandlung.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., and Richardson, C. (1995). A redrawn vandenberg and kuse mental rotations test: different versions and factors that affect performance. *Brain and cognition.*

Piaget, J. (1972). *The principles of genetic epistemology.* Routledge and K. Paul.

Piaget, J. (1974). The future of developmental child psychology. *Journal of Youth and Adolescence.*

Piper, A. M. and Hollan, J. D. (2009). Tabletop displays for small group study: affordances of paper and digital materials. *Proc. CHI 2009.*

Price, S. (2008a). A representation approach to conceptualizing tangible learning environments. In *Proc. TEI 2008.*

Price, S. (2008b). Towards a framework for investigating tangible environments for learning. *International Journal of Arts and Technology.*

Price, S., Falcão, T. P., Sheridan, J. G., and Roussos, G. (2009). The effect of representation location on interaction in a tangible learning environment. In *Proc. TEI 2009*.

Price, S., Rogers, Y., Scaife, M., Stanton, D., and Neale, H. (2003). Using 'tangibles' to promote novel forms of playful learning. *Interacting with Computers*.

Prieto, L. P., Villagra-Sobrino, S., Jorrin-Abellan, I. M., Martinez-Mones, A., and Dimitriadis, Y. (2011). Recurrent routines: Analyzing and supporting orchestration in technology-enhanced primary classrooms. *Computers & Education*.

Quarles, J., Lampotang, S., Fischler, I., Fishwick, P., and Lok, B. (2008). Tangible user interfaces compensate for low spatial cognition. In *Proc. 3DUI 2008*.

Quazzo, D., Cohn, M., Horn, J., and Moe, M. (2012). Fall of the wall - capital flows to education innovation. Technical report, GSV Advisors.

Rankowski, C. A. and Galey, M. (1979). Effectiveness of multimedia in teaching descriptive geometry. *Educational Communication and Technology*.

Reeves, T. C. (2000). Enhancing the Worth of Instructional Technology Research through 'Design Experiments' and Other Development Research Strategies. In *Proc. of Annual Meeting of AERA 2000*.

Resnick, M. (1993). Behavior construction kits. *Commun. ACM*.

Resnick, M., Martin, F., Berg, R., Borovoy, R., Colella, V., Kramer, K., and Silverman, B. (1998). Digital manipulatives: new toys to think with. In *Proc. CHI 1998*.

Robert, M. and Chevrier, E. (2003). Does men's advantage in mental rotation persist when real three-dimensional objects are either felt or seen? *Memory & Cognition*.

Rogers, Y. and Ellis, J. (1994). Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*.

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

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

Roschelle, J. and Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. *Knowledge Creation Diffusion Utilization*.

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

Scaife, M. and Rogers, Y. (1996). External cognition: how do graphical representations work? *Int. J. Hum.-Comput. Stud.*

Schneider, B., Jermann, P., Zufferey, G., and Dillenbourg, P. (2011). Benefits of a tangible interface for collaborative learning and interaction. *IEEE Transactions on Learning Technologies*.

Schneider, B., Wallace, J., Blikstein, P., and Pea, R. (2013). Preparing for future learning with a tangible user interface: The case of neuroscience. *IEEE Transactions on Learning Technologies.*

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

Schwartz, D. L. and Heiser, J. (2006). Spatial representations and imagery in learning. *The Cambridge handbook of the learning sciences.*

Seybert, H. (2012). Eurostat report.

Shaer, O. and Hornecker, E. (2009). Tangible user interfaces: Past, present and future directions (preprint). *Found. Trends Hum.-Comput. Interact.*

Sharlin, E., Itoh, Y., Watson, B., Kitamura, Y., Sutphen, S., and Liu, L. (2002). Cognitive cubes: a tangible user interface for cognitive assessment. In *Proc. CHI 2002.*

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

Shea, D. L., Lubinski, D., and Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology.*

Shepard, R. N. and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science.*

Sims, V. and Mayer, R. (2002). Domain specificity of spatial expertise: the case of video game players. *Applied Cognitive Psychology.*

Smith, I. M. (1964). *Spatial ability: its educational and social significance.* University of London Press.

Sommer, R. (1978). *The Mind's Eye: Imagery in Everyday Life.* Delacorte Press.

Song, H., Guimbretière, F., Hu, C., and Lipson, H. (2006). ModelCraft: capturing freehand annotations and edits on physical 3D models. In *Proc. UIST 2006.*

Sorby, S. (2009). Educational research in developing 3-d spatial skills for engineering students. *International Journal of Science Education.*

Sorby, S., Casey, B., Veurink, N., and Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences.*

Speelpenning, T., Antle, A. N., Doering, T., and van den Hoven, E. (2011). Exploring how tangible tools enable collaboration in a multi-touch tabletop game. In *Proc. Interact 2011.*

Stalder, B. (2011). Le niveau d'exigences intellectuelles des formations professionnelles initiales en suisse. classement des années 1999 à 2005. Technical report, Université de Bâle / TREE.

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

Sylla, C., Branco, P., Coutinho, C., and Coquet, E. (2012). TUIs vs. GUIs: comparing the learning potential with preschoolers. *Personal Ubiquitous Computing*.

Terlecki, M. S., Newcombe, N. S., and Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Applied Cognitive Psychology*.

Triona, L. and Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction*.

Tuddenham, P., Kirk, D., and Izadi, S. (2010). Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In *Proc. CHI 2010*.

Turkle, S. and Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs*.

Ullmer, B., Ishii, H., and Jacob, R. J. K. (2005). Token+constraint systems for tangible interaction with digital information. *ACM Trans. Comput.-Hum. Interact.*

Underkoffler, J. and Ishii, H. (1998). Illuminating light. *Proc. CHI 1998*.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., and Newcombe, N. S. (2012). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*.

Vetter, T., Poggio, T., and Bülthoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. *Current biology: CB*.

Voyer, D. and Saunders, K. A. (2004). Gender differences on the mental rotations test: a factor analysis. *Acta Psychologica*.

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

Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*.

Wang, F. and Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*.

Wang, H.-C., Chang, C.-Y., and Li, T.-Y. (2007). The comparative efficacy of 2D- versus 3D-based media design for influencing spatial visualization skills. *Computers in Human Behavior*.

Watanabe, R., Itoh, Y., Asai, M., Kitamura, Y., and Kishino, F. (2004). The soul of ActiveCube: implementing a flexible, multimodal, three-dimensional spatial tangible interface. In *Computers in Entertainment*.

Wellner, P. (1991). The DigitalDesk calculator: tangible manipulation on a desk top display. In

Proc. UIST 1991.

Xie, L., Antle, A. N., and Motamedi, N. (2008). Are tangibles more fun?: comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In *Proc. TEI 2008*.

Zaman, B., Vanden Abeele, V., Markopoulos, P., and Marshall, P. (2012). Editorial: the evolving field of tangible interaction for children: the challenge of empirical validation. *Personal Ubiquitous Computing*.

Zhang, J. and Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*.

Zuckerman, O., Arida, S., and Resnick, M. (2005). Extending tangible interfaces for education: Digital Montessori-Inspired manipulatives. In *Proc. CHI 2005*.

Zuckerman, O. and Gal-Oz, A. (2013). To TUI or not to TUI: evaluating performance and preference in tangible vs. graphical user interfaces. *International Journal of Human-Computer Studies*.

Zufferey, G. (2010). *The Complementarity of Tangible and Paper Interfaces in Tabletop Environments for Collaborative Learning*. PhD thesis, EPFL.

Zufferey, G., Jermann, P., Lucchi, A., and Dillenbourg, P. (2009). TinkerSheets: Using Paper Forms to Control andVisualize Tangible Simulations. In *Proc. TEI*.

A Appendix

This appendix includes the documents used in the studies presented in this thesis. One important document, the mental rotation test used in Chapter 4, could not be included, because it is not allowed to publish it. The readers interested in getting this test should contact the authors of the test directly.

A.1 Movies

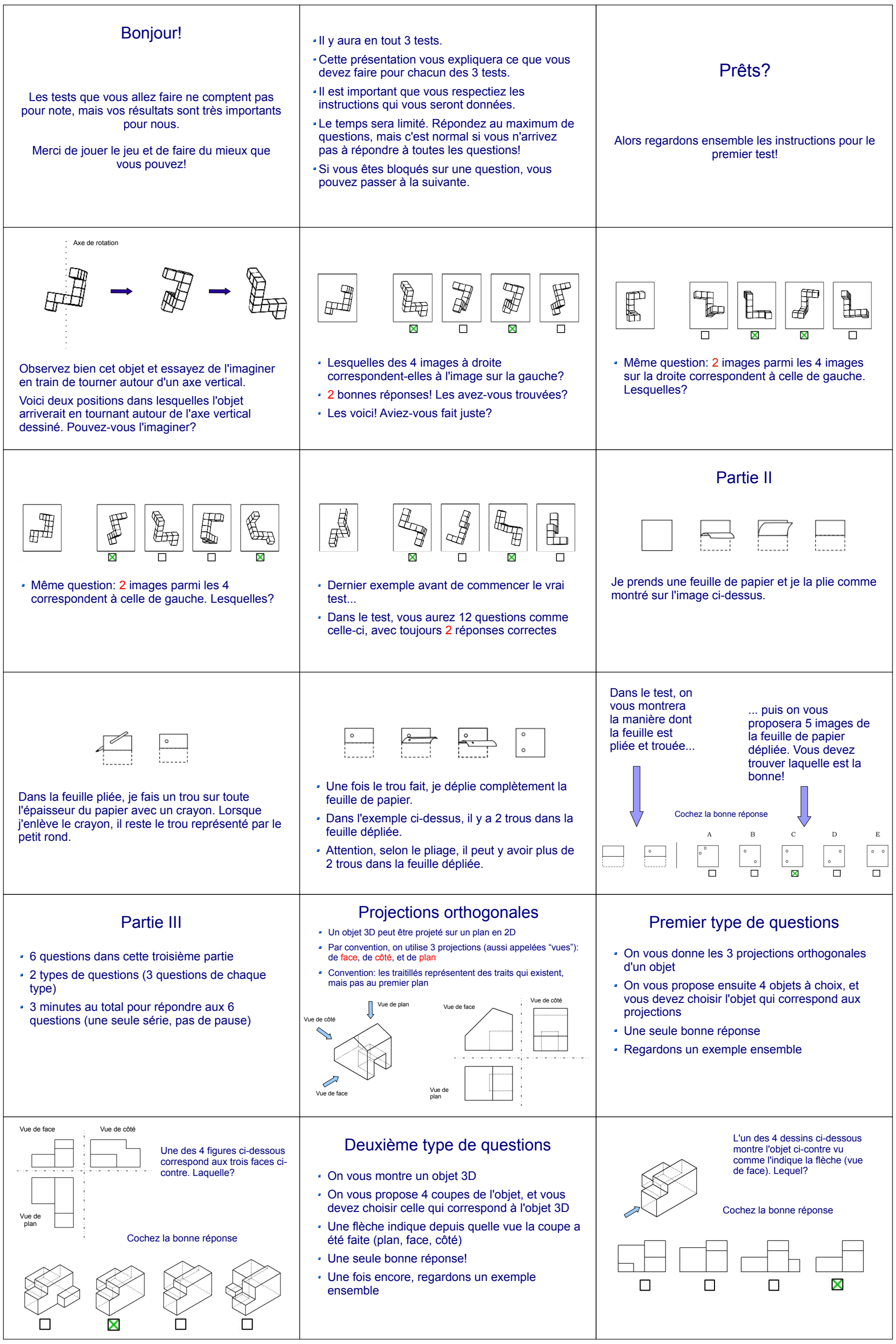
As discussed in this thesis, multiple representations of the same material can lead to a better understanding. With this in mind, I created the videos referenced in Table A.1.

ID	Description	URL
Video 1	Presentation of TapaCarp (Section 5.3) and demonstration of the rabattement learning activity (Section 7.1)	http://www.youtube.com/watch?v=vnlLeCYxmCs
Video 2	The impact of feedback (Section 6.3)	http://www.youtube.com/watch?v=l8-g5bXNWIk
Video 3	Integration in new Pedagogical Scenarios (Section 7.2)	http://www.youtube.com/watch?v=1S3gURKUjSA
Video 4	Scaling up (Chapter 8)	https://www.youtube.com/watch?v=KvBWMvONiRA

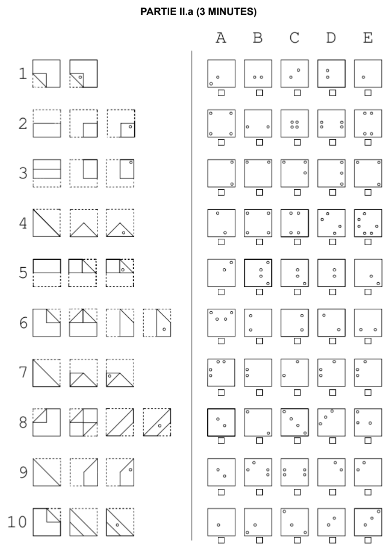
Table A.1: Videos.

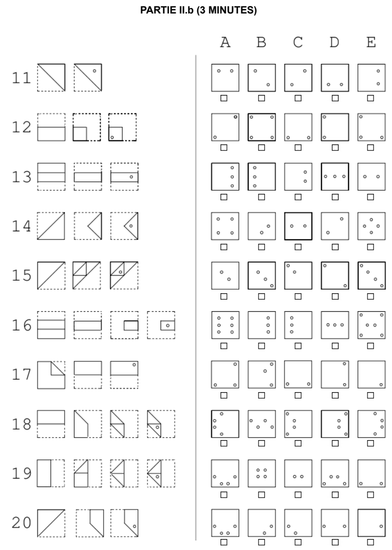
A.2 Documents for Study of Chapter 4

A.2.1 Slides of the PowerPoint presentation

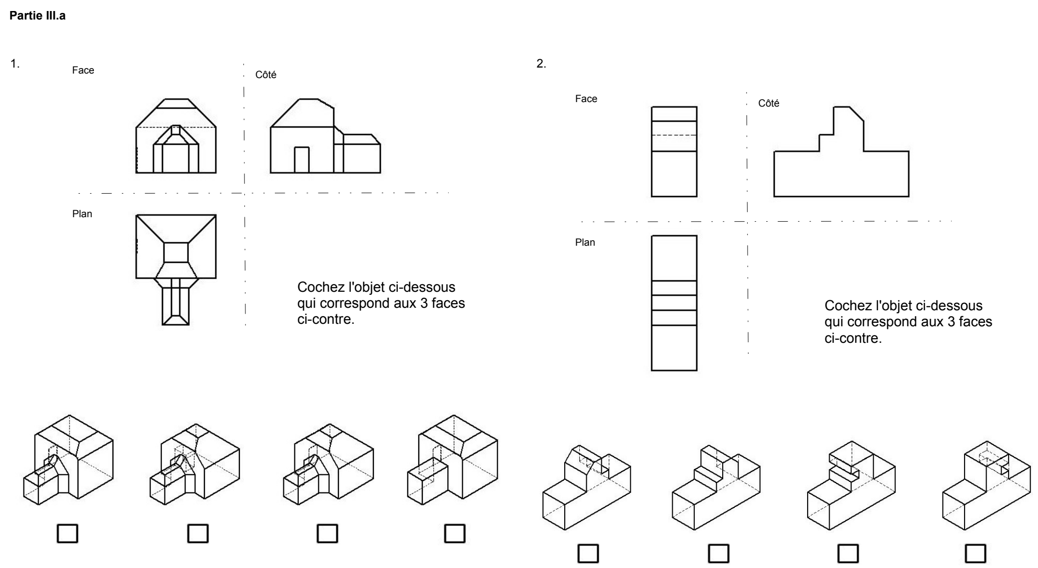


A.2.2 Paper-folding test



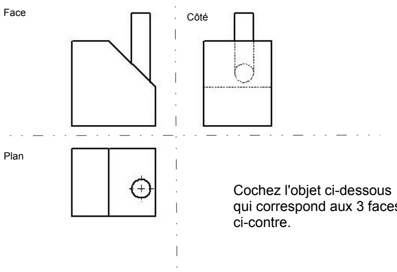


A.2.3 Orthographic projections part

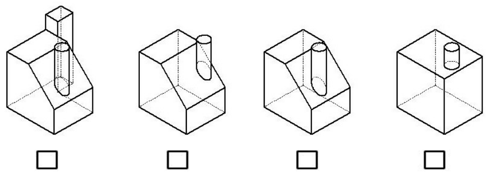


Partie III.b

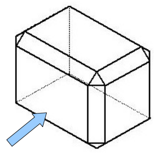
3.



Cochez l'objet ci-dessous qui correspond aux 3 faces ci-contre.

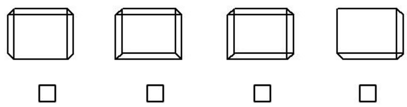


4.



Vue de **face**

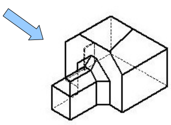
Cochez la face parmi celles ci-dessous qui corresond à l'objet ci-contre vu comme montré par la flèche



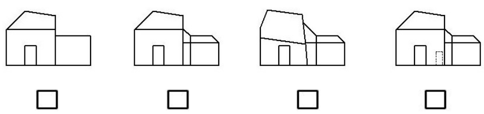
18

5.

Vue de **côté**



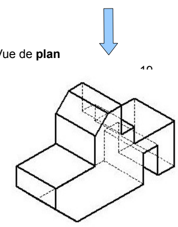
Cochez la face parmi celles ci-dessous qui corresond à l'objet ci-contre vu comme montré par la flèche



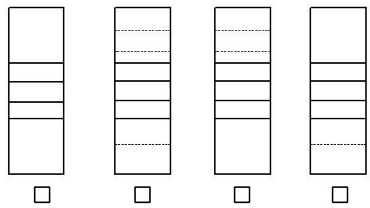
20

6.

Vue de **plan**



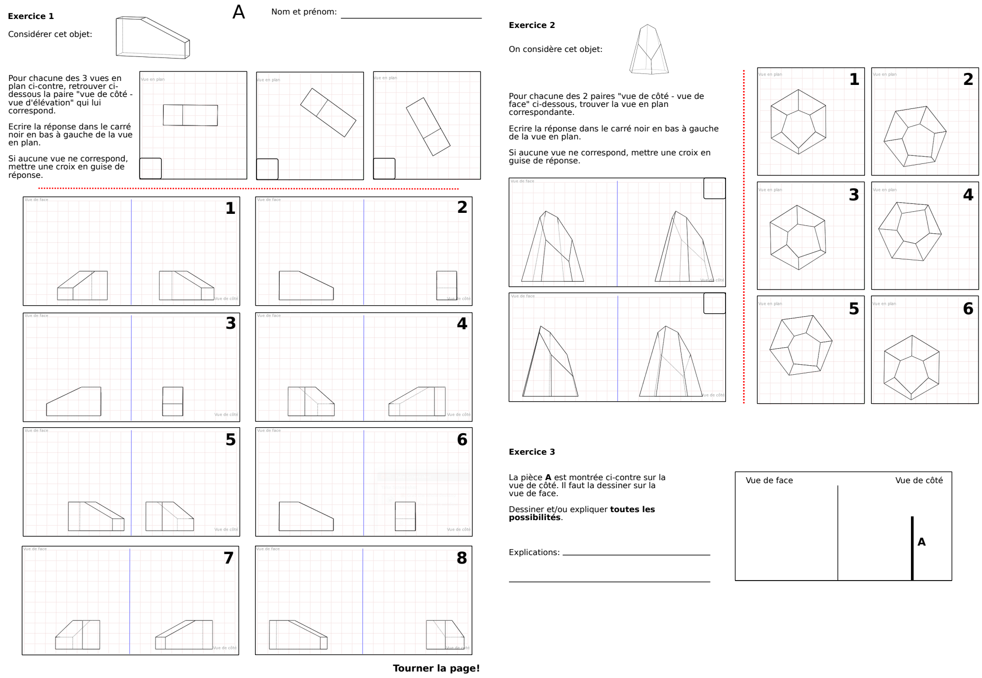
Cochez la face parmi celles ci-dessous qui corresond à l'objet ci-contre vu comme montré par la flèche



21

A.3 Documents for Study of Section 6.2

A.3.1 Pre-test



Exercice 1 A Nom et prénom: ________________

Considérer cet objet:

Pour chacune des 3 vues en plan ci-contre, retrouver ci-dessous la paire "vue de côté - vue d'élévation" qui lui correspond.

Ecrire la réponse dans le carré noir en bas à gauche de la vue en plan.

Si aucune vue ne correspond, mettre une croix en guise de réponse.

1 2 3 4 5 6 7 8

Tourner la page!

Exercice 2

On considère cet objet:

Pour chacune des 2 paires "vue de côté - vue de face" ci-dessous, trouver la vue en plan correspondante.

Ecrire la réponse dans le carré noir en bas à gauche de la vue en plan.

Si aucune vue ne correspond, mettre une croix en guise de réponse.

1 2 3 4 5 6

Exercice 3

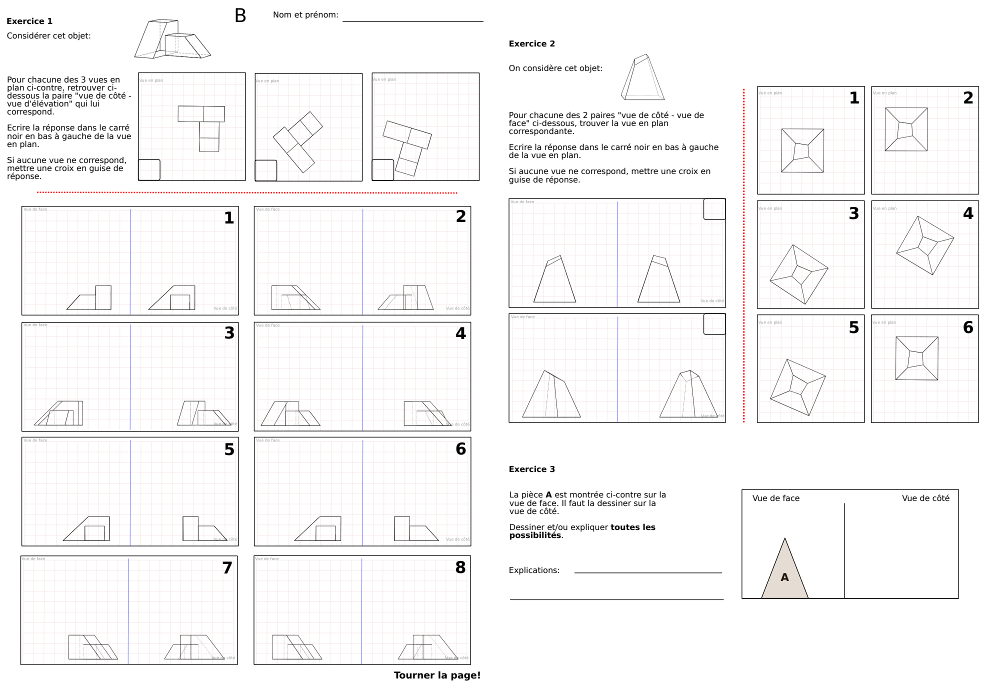
La pièce **A** est montrée ci-contre sur la vue de côté. Il faut la dessiner sur la vue de face.

Dessiner et/ou expliquer **toutes les possibilités**.

Explications: ________________

Vue de face Vue de côté A

A.3.2 Post-test



Exercice 1 B Nom et prénom: ________________

Considérer cet objet:

Pour chacune des 3 vues en plan ci-contre, retrouver ci-dessous la paire "vue de côté - vue d'élévation" qui lui correspond.

Ecrire la réponse dans le carré noir en bas à gauche de la vue en plan.

Si aucune vue ne correspond, mettre une croix en guise de réponse.

1 2 3 4 5 6 7 8

Tourner la page!

Exercice 2

On considère cet objet:

Pour chacune des 2 paires "vue de côté - vue de face" ci-dessous, trouver la vue en plan correspondante.

Ecrire la réponse dans le carré noir en bas à gauche de la vue en plan.

Si aucune vue ne correspond, mettre une croix en guise de réponse.

1 2 3 4 5 6

Exercice 3

La pièce **A** est montrée ci-contre sur la vue de face. Il faut la dessiner sur la vue de côté.

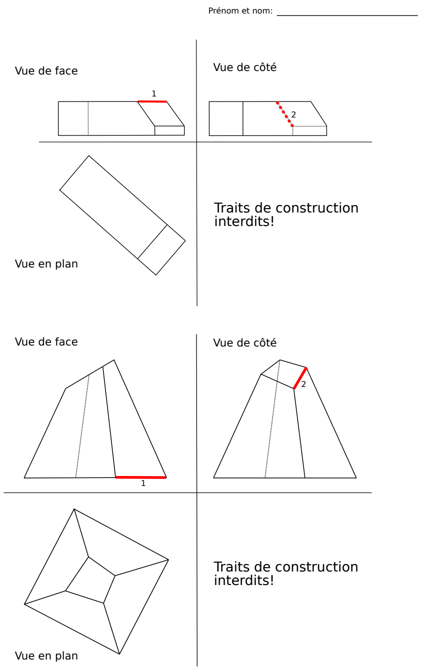
Dessiner et/ou expliquer **toutes les possibilités**.

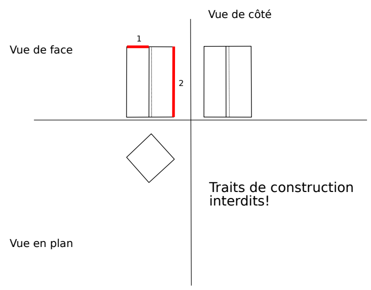
Explications: ________________

Vue de face Vue de côté A

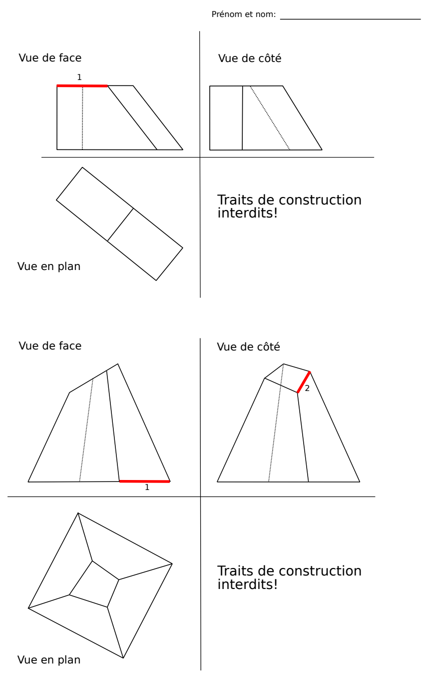
A.4 Documents for Study of Section 6.3

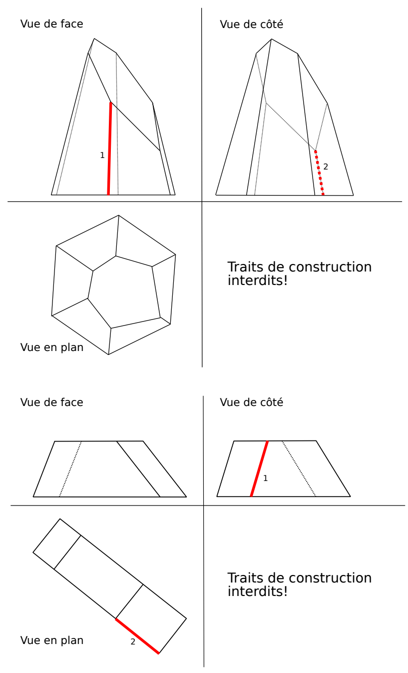
A.4.1 Pre-test





A.4.2 Post-test





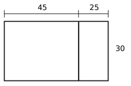
A.5 Documents for Study of Section 7.1

A.5.1 Pre-test

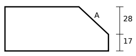
Question 1

Nom: ______________________________

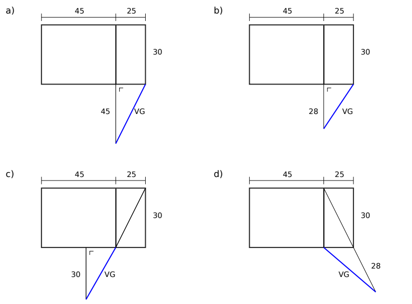
Vue en plan:



Vue de face:

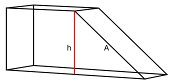


On recherche la vraie grandeur (VG) de l'arête A. Cocher la ou les bonnes réponses parmi celles ci-dessous:



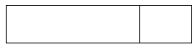
Question 2

Considérez l'objet suivant:



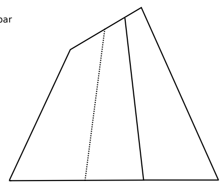
Voici la vraie grandeur de l'arête A:

Sur la vue en plan ci-dessous, dessinez la hauteur h en vraie grandeur:



Question 3

Considérez cette la pyramide tronquée donnée par sa vue de face, ci-contre.

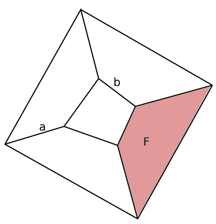


Vue de face

Sur la vue en plan ci-dessous:

1) dessinez la vraie grandeur des arêtes a et b par rabattement

2) dessinez la vraie grandeur de la face F par rabattement

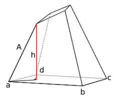


A.5.2 Post-test

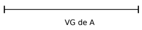
Question 1

Nom: ____________________

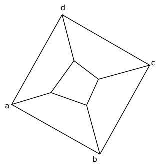
Considérez l'objet suivant:



Voici la vraie grandeur de l'arête A:

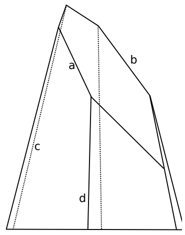


Sur la vue en plan ci-dessous, dessinez la hauteur h en vraie grandeur:



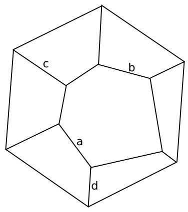
Question 2

Considérez le prisme hexagonal tronqué donné par sa vue de face, ci-contre.

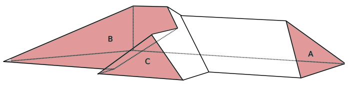


Vue de face

Sur la vue en plan ci-dessous, dessinez la vraie grandeur des arêtes a, b, c, et d par rabattement.



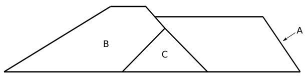
Question 3



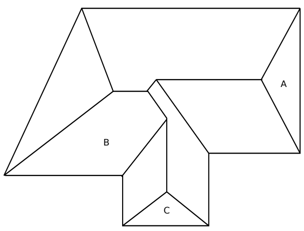
Vous avez appris à rechercher la vraie grandeur d'une face par rabattement.

La première étape lors du rabattement d'une face et de trouver les chevrons d'emprunt pour la face.

1) Sur la vue de face ci-dessous, dessinez les chevrons d'emprunts des faces A, B et C et indiquez lesquels appartiennent à quelle face.



2) Dessinez la face **B** en vraie grandeur sur la vue en plan ci-dessous:



A.5.3 Questionnaire

Nom: ______________________________

Introduction

L'apprentissage en cours : est ma première formation est ma seconde formation

- En général, ce que j'apprends à l'école m'est très utile sur mon lieu de travail.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- En général, j'apprends mieux lorsque je travaille en groupe que seul.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Activités que je fais à la fois à l'école et au travail:

____

____

Le cours d'aujourd'hui

- La matière vue dans le cours d'aujourd'hui était nouvelle.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- J'utilise dans mon travail ce que j'ai étudié aujourd'hui.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Je trouve que ce que j'ai appris aujourd'hui pourrait être utile pour ma formation ou plus tard.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Est-ce que vous avez rencontré des difficultés dans les activités d'aujourd'hui? Si oui, lesquelles ?

____

____

Utilisation de la lampe interactive

- La lampe m'a permis de mieux comprendre les rabattements.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Pourquoi ?

____

____

- Les animations étaient utiles pour comprendre les rabattements.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- J'étais plus motivé aujourd'hui parce que j'utilisais la lampe.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Si j'utilisais la lampe régulièrement dans mon apprentissage, j'apprendrais sûrement mieux.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Je serais intéressé à utiliser la lampe de manière régulière

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Je préfère travailler sur papier comme d'habitude.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Je trouve que la lampe serait utile pour d'autres sujets dans mon apprentissage.

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

- Lesquels?

____

____

- Je trouve que les activités faites avec la lampe étaient facilement utilisables:

  Pas du tout d'accord -3 -2 -1 0 +1 +2 +3 tout à fait d'accord

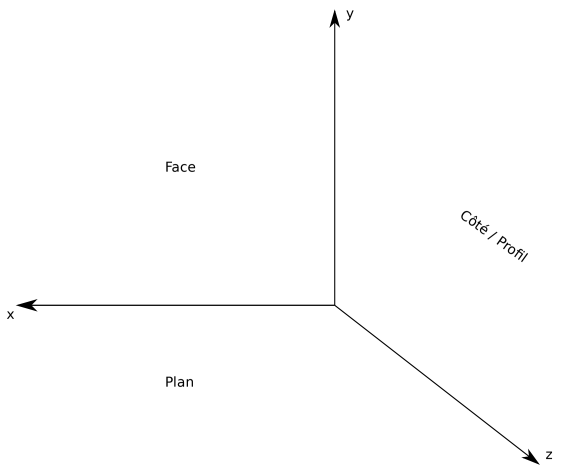
- Points à améliorer pour l'utilisation de la lampe:

____

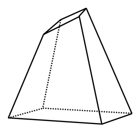
____

A.5.4 Activity booklet

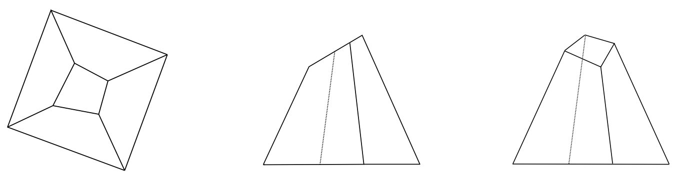
Rappel: position des plans de projection (projection orthogonale)



Ainsi, pour l'objet:

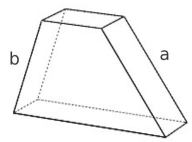


On aura les vues suivantes:



Vue en plan | Vue de face | Vue de côté

Activité 1



1) Taille de l'arête **a** : _________ mm

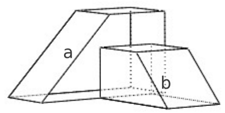
2) Position de la pièce dans le plan pour que l'arête **a** soit en vraie grandeur sur la vue de face :

3) Taille de l'arête **b** : _________ mm

4) Position de la pièce dans le plan pour que l'arête **b** soit en vraie grandeur sur la vue de face :

5) Position de la pièce dans le plan pour que l'arête **a** et l'arête **b** soient en vraie grandeur **en même temps** sur la vue de face :

Activité 2



1) Taille de l'arête **a**: _________ mm

2) Position de la pièce dans le plan pour que l'arête **a** soit en vraie grandeur sur la vue de face :

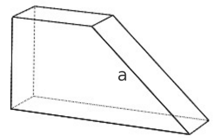
3) Taille de l'arête **b**: _________ mm

4) Position de la pièce dans le plan pour que l'arête **b** soit en vraie grandeur sur la vue de face :

5) Position de la pièce dans le plan pour que l'arête **a** et l'arête **b** soient en vraie grandeur **en même temps** sur la vue de face :

6) Condition(s) pour qu'une arête apparaisse en vraie grandeur :

Activité 3

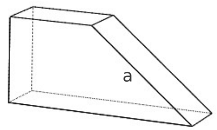


1) Taille de l'arête **a**: _________ mm

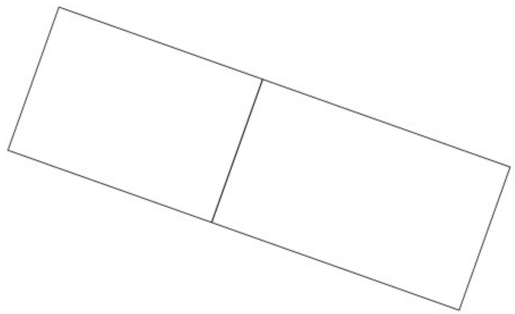
2) Position de la pièce dans le plan pour que l'arête **a** soit en vraie grandeur sur la vue en plan :

Ma démarche :

Activité 4

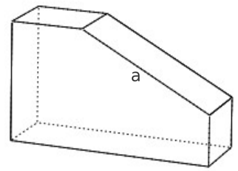


1) Dessinez le rabattement de l'arête **a**.

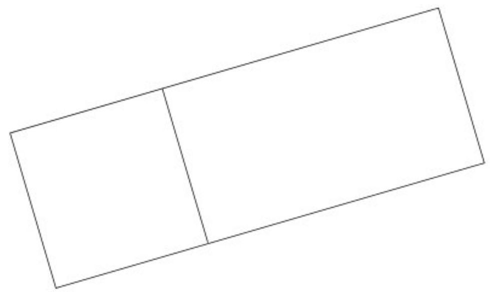


2) Marche à suivre pour trouver la vraie grandeur d'une arête :

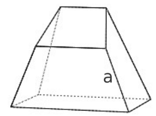
Activité 5



1) Dessinez le rabattement de l'arête **a** ci-dessous.
Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide



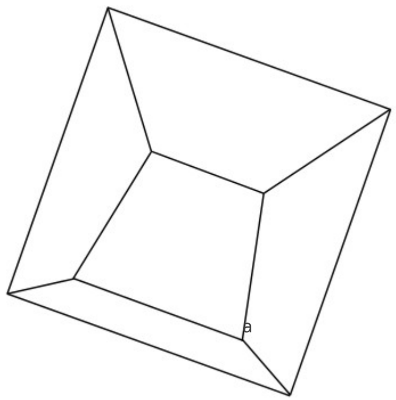
Activité 6



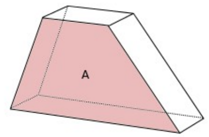
1) Dessinez le rabattement de l'arête **a** sur la vue en plan ci-dessous..

Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide

2) Dessinez le rabattement de 4 autres arêtes de votre choix, mais qui n'appartiennent pas à la base de la pyramide.

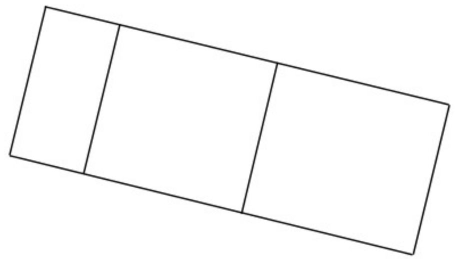


Activité 7



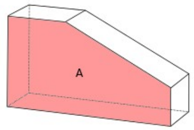
1) Dessinez le rabattement de la face **A** ci-dessous.

Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide

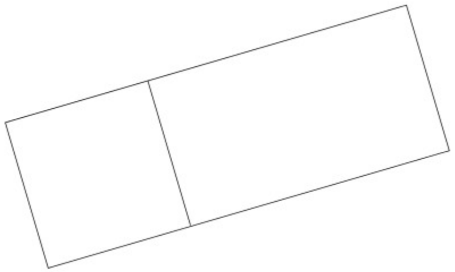


2) Expliquez comment vous avez fait:

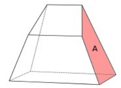
Activité 8



1) Dessinez le rabattement de la face **A**.
Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide



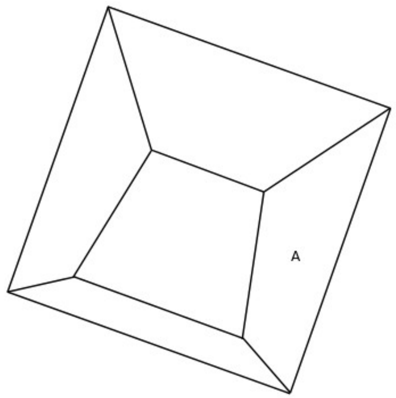
Activité 9



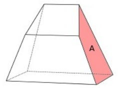
CE

Arête

1) Dessinez le rabattement de la face **A**.
Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide



Activité 10

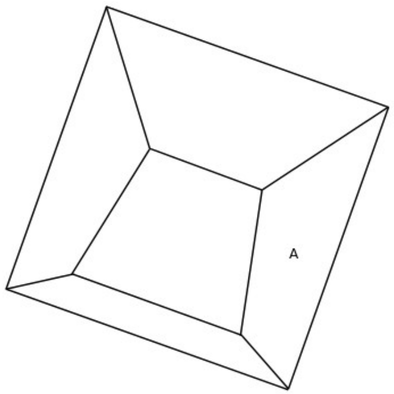


CE | Arête

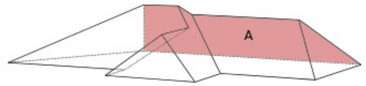
1) Dessinez le rabattement de la face **A**, mais sans mesurer le chevron d'emprunt sur la pièce. Il vous faut donc trouver un moyen de déterminer sa vraie grandeur.

   Si vous n'y arrivez pas, vous pouvez utiliser la carte d'aide

2) Dessinez le rabattement des 3 autres faces verticales de la pyramide.

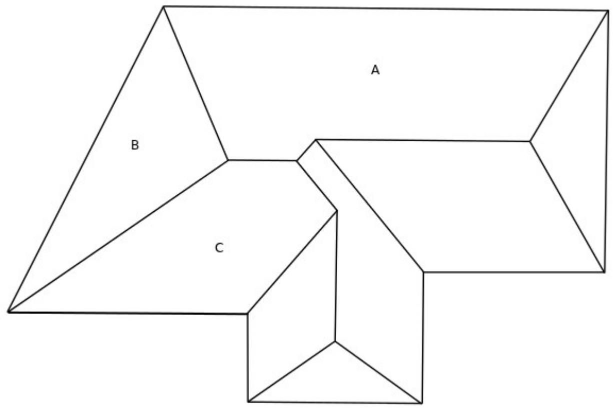


Activité 11

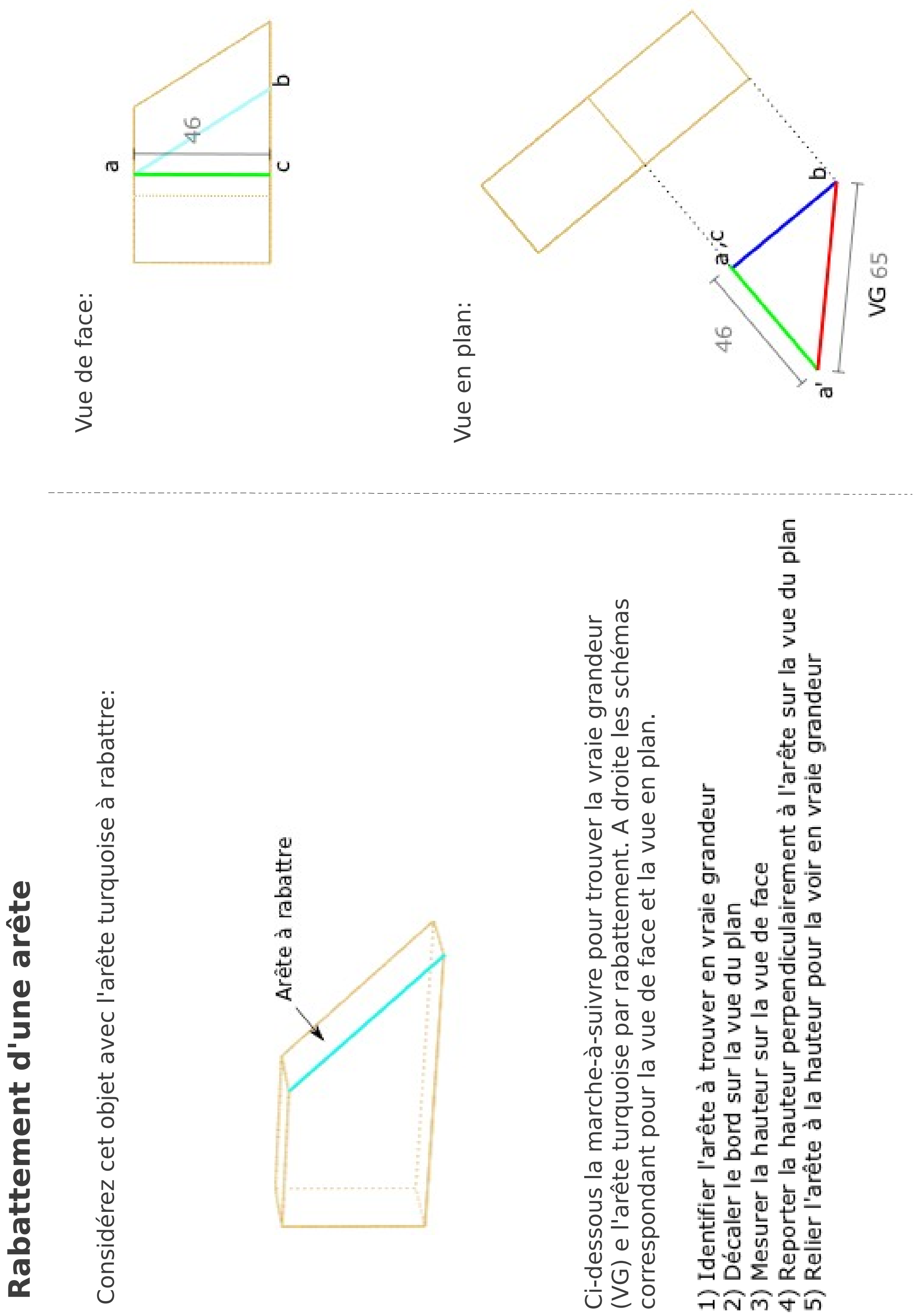


1) Dessinez le rabattement de la face **A**, mais sans mesurer le chevron d'emprunt sur la pièce.

   Faites de même pour les faces **B** et **C**. Si vous avez terminé, rabattez les faces restantes du toit.



A.5.5 Rabattement theory given to control group



A.6 Documents for Study of Section 7.2

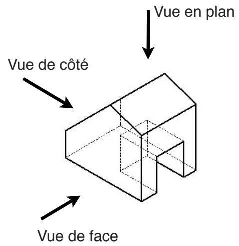
For this study, the documents available are the pre-test and the post-test, as well as the sheets used for the analysis of the errors of the objects.

A.6.1 Pre-test

Pre-test

Nom et prénom: ______________________________

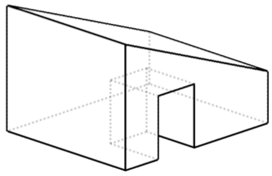
Rappel:



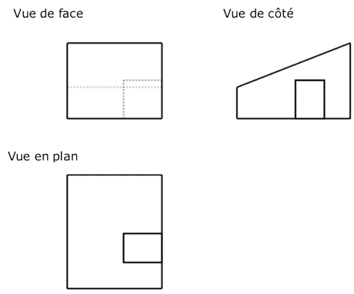
1

Partie A

Objet voulu (5 coupes)



Plan fait par Paul:



2

Question 1

Pour chaque vue, entourer "oui" ou "non" pour indiquer si la vue est correcte. En cas de "non", indiquer sur le plan où se trouve(nt) la/les erreur(s).

1.1. Est-ce que la vue en plan est correcte? oui non

1.2. Est-ce que la vue de face est correcte? oui non

1.3. Est-ce que la vue de côté est correcte? oui non

Question 2

Corriger le plan de Paul pour qu'il corresponde à l'objet voulu. Dessiner directement les corrections sur le plan de Paul. Bien indiquer ce que vous avez modifié, si nécessaire en ajoutant des commentaires.

Question 3

A-t-on besoin de toutes les vues du plan pour pouvoir couper l'objet? Justifier.

____

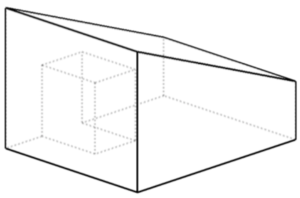
____

Tourner la page!

3

Question 4

Voici l'objet coupé par Corentin à partir des plans de Paul. Cet objet ne correspond pas à l'objet voulu.



4.1 Combien y a-t-il d'erreurs sur cet objet, par rapport à l'objet voulu? Indiquer la ou les erreur(s) sur l'objet ci-dessus, et pour chaque erreur, donner une courte description de l'erreur dans le tableau ci-dessous. Numéroter les erreurs sur le dessin et dans le tableau.

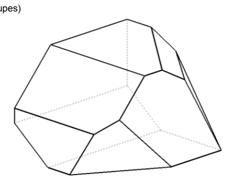
4.2 Pour chaque erreur, l'erreur est-elle due au plan de Paul ou est-ce que Corentin a mal tracé et coupé? Justifier la réponse.

Numéro et description de l'erreur	**C'est la faute de ... parce que ...**

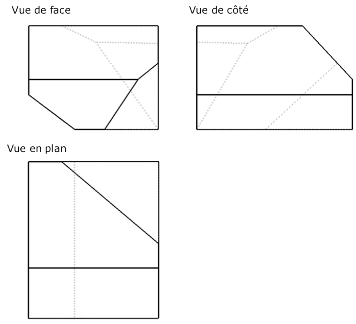
4

Partie B

Objet voulu (5 coupes)



Plan fait par Paul:



Question 1

Pour chaque vue, entourer "oui" ou "non" pour indiquer si la vue est correcte. En cas de "non", indiquer sur le plan où se trouve(nt) la/les erreur(s).

1.1. Est-ce que la vue en plan est correcte? oui non

1.2. Est-ce que la vue de face est correcte? oui non

1.3. Est-ce que la vue de côté est correcte? oui non

Question 2

Corriger le plan de Paul pour qu'il corresponde à l'objet voulu. Dessiner directement les corrections sur le plan de Paul. Bien indiquer ce que vous avez modifié, si nécessaire en ajoutant des commentaires.

Question 3

A-t-on besoin de toutes les vues du plan pour pouvoir couper l'objet? Justifier.

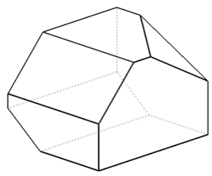
____

____

Tourner la page!

Question 4

Voici l'objet coupé par Corentin à partir des plans de Paul. Cet objet ne correspond pas à l'objet voulu.



4.1 Combien y a-t-il d'erreurs sur cet objet, par rapport à l'objet voulu? Indiquer la ou les erreur(s) sur l'objet ci-dessus, et pour chaque erreur, donner une courte description de l'erreur dans le tableau ci-dessous. Numéroter les erreurs sur le dessin et dans le tableau.

4.2 Pour chaque erreur, l'erreur est-elle due au plan de Paul ou est-ce que Corentin a mal tracé et coupé? Justifier la réponse.

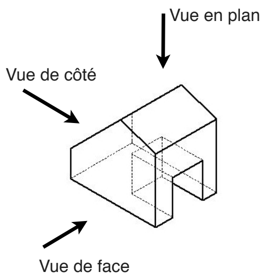
Numéro et description de l'erreur	**C'est la faute de ... parce que ...**

A.6.2 Post-test

Post-test

Nom et prénom: ______________________________________

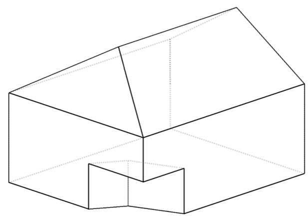
Rappel:



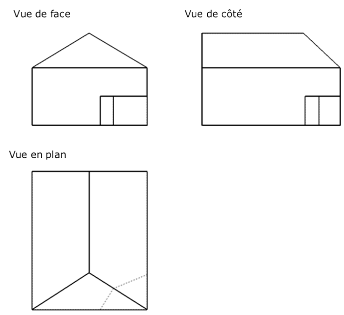
1

Partie A

Objet voulu (6 coupes):



Plan fait par Paul:



2

Question 1

Pour chaque vue, entourer "oui" ou "non" pour indiquer si la vue est correcte. En cas de "non", indiquer sur le plan où se trouve(nt) la/les erreur(s).

1.1. Est-ce que la vue en plan est correcte? oui non

1.2. Est-ce que la vue de face est correcte? oui non

1.3. Est-ce que la vue de côté est correcte? oui non

Question 2

Corriger le plan de Paul pour qu'il corresponde à l'objet voulu. Dessiner directement les corrections sur le plan de Paul. Bien indiquer ce que vous avez modifié, si nécessaire en ajoutant des commentaires.

Question 3

A-t-on besoin de toutes les vues du plan pour pouvoir couper l'objet? Justifier.

____

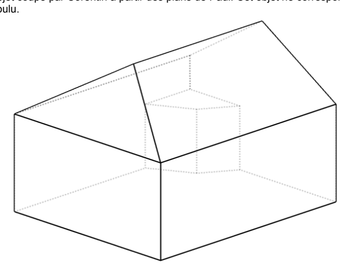
____

Tourner la page!

3

Question 4

Voici l'objet coupé par Corentin à partir des plans de Paul. Cet objet ne correspond pas à l'objet voulu.



4.1 Combien y a-t-il d'erreurs sur cet objet, par rapport à l'objet voulu? Indiquer la ou les erreur(s) sur l'objet ci-dessus, et pour chaque erreur, donner une <u>courte description</u> de l'erreur dans le tableau ci-dessous. Numéroter les erreurs sur le dessin et dans le tableau.

4.2 Pour chaque erreur, l'erreur est-elle due au plan de Paul ou est-ce que Corentin a mal tracé et coupé? Justifier la réponse.

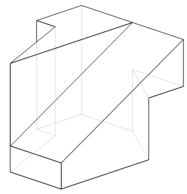
Numéro et description de l'erreur	C'est la faute de ... parce que ...

4

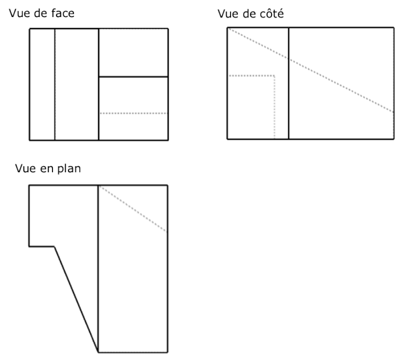
5

Partie B

Objet voulu (6 coupes)



Plan fait par Paul:



6

Question 1

Pour chaque vue, entourer "oui" ou "non" pour indiquer si la vue est correcte. En cas de "non", indiquer sur le plan où se trouve(nt) la/les erreur(s).

1.1. Est-ce que la vue en plan est correcte? oui non

1.2. Est-ce que la vue de face est correcte? oui non

1.3. Est-ce que la vue de côté est correcte? oui non

Question 2

Corriger le plan de Paul pour qu'il corresponde à l'objet voulu. Dessiner directement les corrections sur le plan de Paul. Bien indiquer ce que vous avez modifié, si nécessaire en ajoutant des commentaires.

Question 3

A-t-on besoin de toutes les vues du plan pour pouvoir couper l'objet? Justifier.

____

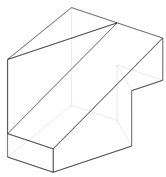
____

Tourner la page!

7

Question 4

Voici l'objet coupé par Corentin à partir des plans de Paul. Cet objet ne correspond pas à l'objet voulu.



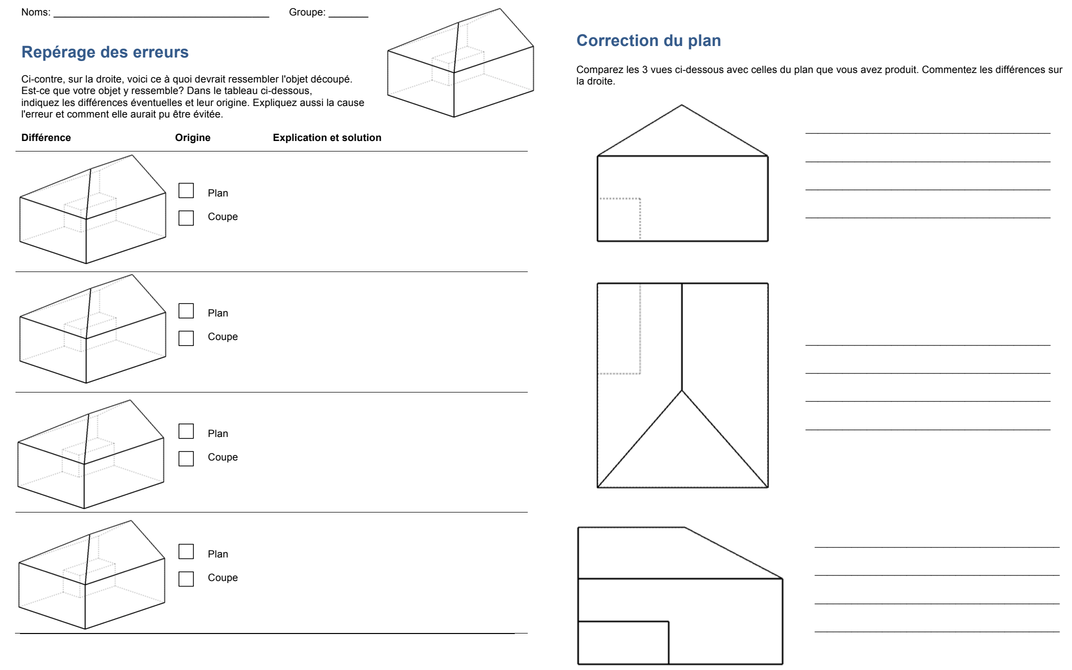
4.1 Combien y a-t-il d'erreurs sur cet objet, par rapport à l'objet voulu? Indiquer la ou les erreur(s) sur l'objet ci-dessus, et pour chaque erreur, donner une courte description de l'erreur dans le tableau ci-dessous. Numéroter les erreurs sur le dessin et dans le tableau.

4.2 Pour chaque erreur, l'erreur est-elle due au plan de Paul ou est-ce que Corentin a mal tracé et coupé? Justifier la réponse.

Numéro et description de l'erreur	C'est la faute de ... ,parce que ...

8

A.6.3 Support sheets for the object analysis done by the groups



Noms: ______________________ Groupe: _______

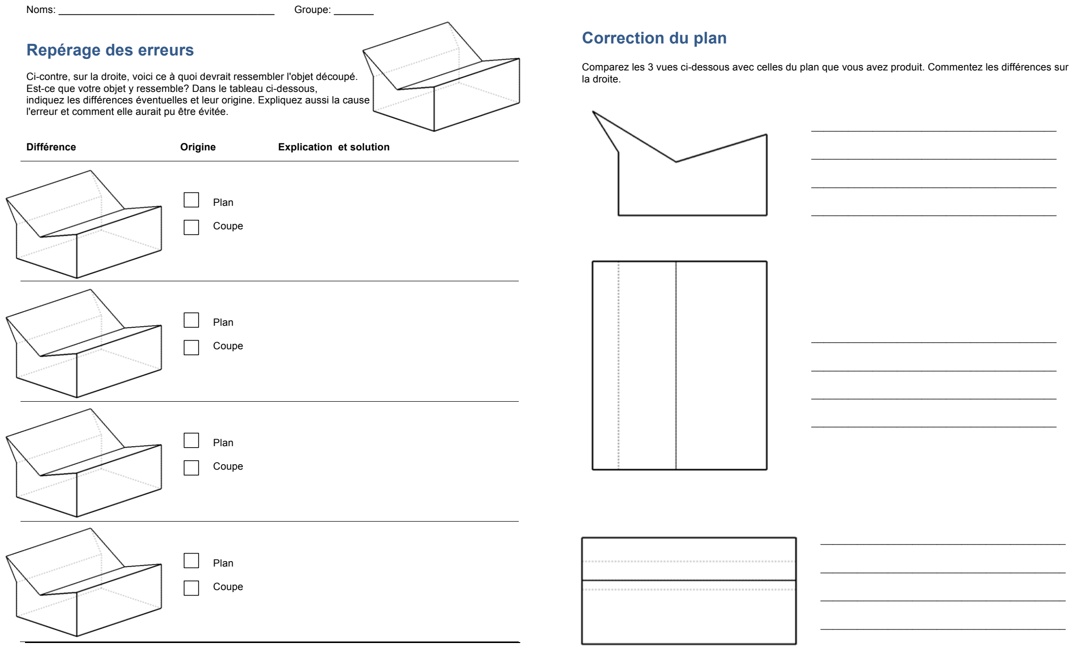
Repérage des erreurs

Ci-contre, sur la droite, voici ce à quoi devrait ressembler l'objet découpé. Est-ce que votre objet y ressemble? Dans le tableau ci-dessous, indiquez les différences éventuelles et leur origine. Expliquez aussi la cause l'erreur et comment elle aurait pu être évitée.

Différence	Origine	Explication et solution
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	

Correction du plan

Comparez les 3 vues ci-dessous avec celles du plan que vous avez produit. Commentez les différences sur la droite.



Noms: ______________________ Groupe: _______

Repérage des erreurs

Ci-contre, sur la droite, voici ce à quoi devrait ressembler l'objet découpé. Est-ce que votre objet y ressemble? Dans le tableau ci-dessous, indiquez les différences éventuelles et leur origine. Expliquez aussi la cause l'erreur et comment elle aurait pu être évitée.

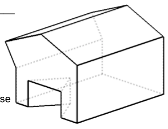
Différence	Origine	Explication et solution
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	

Correction du plan

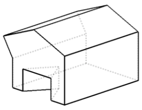
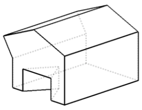
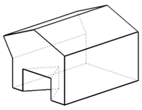
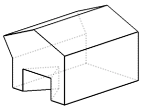
Comparez les 3 vues ci-dessous avec celles du plan que vous avez produit. Commentez les différences sur la droite.

Noms: ______________________ Groupe: _______

Repérage des erreurs

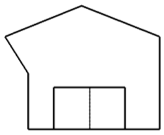


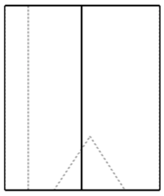
Ci-contre, sur la droite, voici ce à quoi devrait ressembler l'objet découpé. Est-ce que votre objet y ressemble? Dans le tableau ci-dessous, indiquez les différences éventuelles et leur origine. Expliquez aussi la cause l'erreur et comment elle aurait pu être évitée.

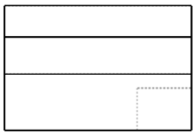
Différence	Origine	Explication et solution
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	

Correction du plan

Comparez les 3 vues ci-dessous avec celles du plan que vous avez produit. Commentez les différences sur la droite.

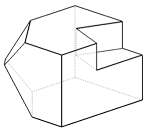




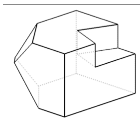
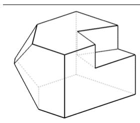
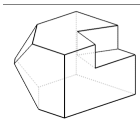
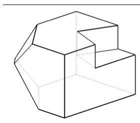


Noms: ______________________ Groupe: _______

Repérage des erreurs

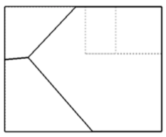


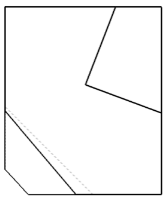
Ci-contre, sur la droite, voici ce à quoi devrait ressembler l'objet découpé. Est-ce que votre objet y ressemble? Dans le tableau ci-dessous, indiquez les différences éventuelles et leur origine. Expliquez aussi la cause l'erreur et comment elle aurait pu être évitée.

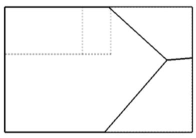
Différence	Origine	Explication et solution
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	
	☐ Plan ☐ Coupe	

Correction du plan

Comparez les 3 vues ci-dessous avec celles du plan que vous avez produit. Commentez les différences sur la droite.







A.7 Documents for Study of Chapter 8

A.7.1 Questionnaire for subscribers

Questionnaire for tangiblecourse.appspot.com

Data will be anonymised and used for research purposes only.
* Required

Gender *
- Male
- Female

Age *
- 16 - 25
- 26 - 35
- 36 - 45
- 46 - more

Level of education *
- Bachelor
- Master
- PhD
- Other:

Domain *
- Carpentry
- IT
- Engineering
- Other:

I completed the following steps: *
- Watched the introductory video
- Printed the PDF files
- Folded the paper blocks
- Set up my webcam
- Tried the practice challenges
- Other:

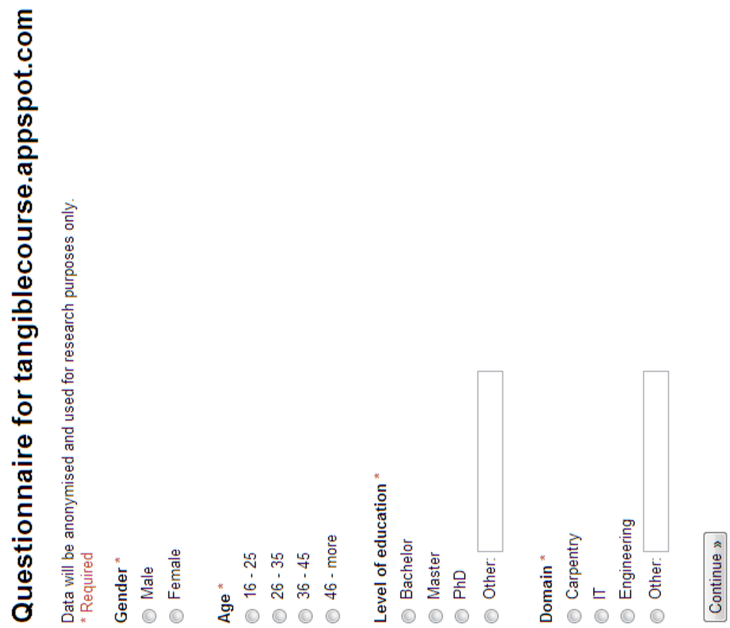
How much time did you spend doing the previous steps? *
- less than 10 minutes
- between 10 and 20 minutes
- between 20 and 30 minutes
- more than 30 minutes

Why didn't you try to go further? *
- This seemed like too much effort
- I did not have an external webcam
- I did not have access to a printer
- I did not have enough time
- I do not have any interest for 3D geometry
- I was frustrated
- Other:

Do you think tangible interfaces could be used for other domains than 3D geometry? If so, which ones?

Do you have any suggestions or feedback?
Critics, ideas, bugs, congratulations, etc.

A.7.2 Questionnaire for contestants



Questionnaire for tangiblecourse.appspot.com

Data will be anonymised and used for research purposes only.
* Required

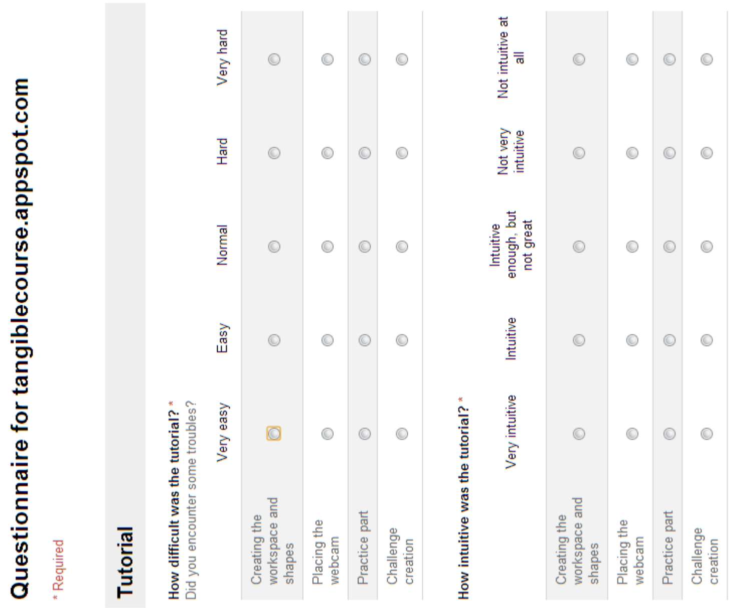
Gender *
- Male
- Female

Age *
- 16 - 25
- 26 - 35
- 36 - 45
- 46 - more

Level of education *
- Bachelor
- Master
- PhD
- Other:

Domain *
- Carpentry
- IT
- Engineering
- Other:

Continue »



Questionnaire for tangiblecourse.appspot.com

* Required

Tutorial

How difficult was the tutorial? *
Did you encounter some troubles?

	Very easy	Easy	Normal	Hard	Very hard
Creating the workspace and shapes					
Placing the webcam					
Practice part					
Challenge creation					

How intuitive was the tutorial? *

	Very intuitive	Intuitive	Intuitive enough, but not great	Not very intuitive	Not intuitive at all
Creating the workspace and shapes					
Placing the webcam					
Practice part					
Challenge creation					

How frustrating was the tutorial? *

	Not frustrating at all	Not too frustrating	Frustrating	Very frustrating	Completely frustrating
Creating the workspace and shapes	◯	◯	◯	◯	◯
Placing the webcam	◯	◯	◯	◯	◯
Practice part	◯	◯	◯	◯	◯
Challenge creation	◯	◯	◯	◯	◯

Activity

How easily were you able to find on the website the information that you needed? *

	1	2	3	4	5	
Very easily	◯	◯	◯	◯	◯	It was very hard

How fun was the activity that you had to perform? *

	1	2	3	4	5	
Not at all	◯	◯	◯	◯	◯	Very

How interesting was the activity? *

	1	2	3	4	5	
Not at all	◯	◯	◯	◯	◯	Very

How easy/difficult was it to solve the challenges? *

	1	2	3	4	5	
Easy	◯	◯	◯	◯	◯	Very difficult

Do you think that an activity that makes you link the 2D and 3D representations of an object is useful? *

	1	2	3	4	5	
Not useful at all	◯	◯	◯	◯	◯	Very useful

How interested are you in developing your 3D geometry skills? *

	1	2	3	4	5	
Not at all interested	◯	◯	◯	◯	◯	Very interested

Interface

Do you think that having the physical objects helped you solve the challenges? *

	1	2	3	4	5	
Not at all	◯	◯	◯	◯	◯	Totally

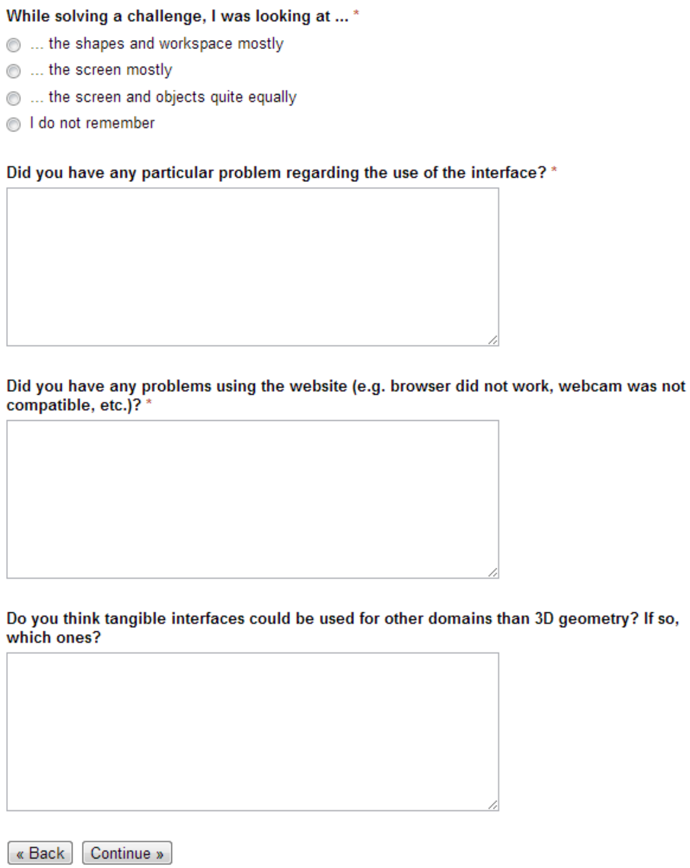
Do you think a flat representation of the object would have been sufficient? *

By flat representation, we mean a piece of paper having the same shape as the top view of the object

- ◯ Yes, a flat representation would have made no difference at all.
- ◯ I would not have been as efficient, but it would have saved me a lot of time in the phase of building the interface.
- ◯ No, seeing the 3D object was really helpful, a flat representation would have made it harder to solve these kind of problems.
- ◯ I'm not sure, I would have to try.

If you had completed the same activity with just a mouse and keyboard (i.e. no tangible objects), you would have... *

- ◯ ... had more fun
- ◯ ... had the same amount of fun
- ◯ ... had less fun



While solving a challenge, I was looking at ... *

- ... the shapes and workspace mostly
- ... the screen mostly
- ... the screen and objects quite equally
- I do not remember

Did you have any particular problem regarding the use of the interface? *

Did you have any problems using the website (e.g. browser did not work, webcam was not compatible, etc.)? *

Do you think tangible interfaces could be used for other domains than 3D geometry? If so, which ones?

« Back | Continue »

Questionnaire for tangiblecourse.appspot.com

Suggestions

Do you have any suggestions or feedback?
Critics, ideas, bugs, congratulations, etc.

List of Acronyms

2D	Two-dimensional
3D	Three-dimensional
AR	Augmented Reality
CAD	Computer-Aided Design
CNC	Computerized numerical control
CSCL	Computer-Supported Collaborative Learning
DBR	Design-Based Research
TUI	Tangible User Interface
VET	Vocational Education and Training

Table A.2: Acronyms used in this thesis.

List of Figures

List of Tables

Sébastien Cuendet

Ch. du Bocage 11, 1026 Echandens – SWITZERLAND
E-Mail: sebastien.cuendet@gmail.com

Education

PhD, Computer Science, EPFL **2009 – 2013**
Main domains of research: HCI and Learning technologies

Masters of Science, EPFL **2004 – 2006**
Master's Thesis at the International Computer Science Institute, UC Berkeley, CA, USA (grade: 6.0/6.0)

Bachelor of Science, EPFL **2001 – 2004**
Exchange year, Technische Universität Darmstadt, Germany (2003)

Work Experience

PhD candidate – *EPFL, Lausanne, CH* **Aug 2009 – Sep 2013**

- Developed an interactive tabletop augmented-reality system to train spatial skills (C++/OpenGL), video: goo.gl/xunsgs
- Helped develop a cheaper and online version of the tabletop system, video: goo.gl/oHarhF
- Ran a dozen studies (controlled studies, contextual inquiry, usability studies, etc.)
- Conducted advanced data analysis (using R)
- Teaching-assistant for 3 classes
- 4 conference and 1 journal articles, 1 best paper award (list of publications: goo.gl/KK0jls)

Co-founder – *Wizzy Education Technologies, Lausanne, CH* **Aug 2010 – present**

- Responsible for the user experience and content quality of the main product, Wizbee
- Responsible for business strategy and development, and promotion of Wizbee
- Won a funding prize (10k CHF), signed 2 statewide subscriptions
- Video of Wizbee: goo.gl/XxOCRH

Research Intern – *Microsoft Research India*, Bangalore, India **May 2012 – July 2012**

- Designed and implemented a mobile application for low-literate users
- Implemented multi-modal interface (speech + graphics) and speech recognition server
- Led to two publications at ACM CHI conference and ACM dev.
- Demo video of the app: goo.gl/0GzDZM

Software Engineer – *Optaros, Geneva, CH* **Aug 2007 – May 2009**

- Responsible for the front-end development of a web-based application (Javascript, Perl)
- Application used by all employees of the justice department of the state of Geneva to manage all the judiciary procedures.
- Involved in the whole development cycle: requirement specifications, cost assessment, planning, implementation, testing, and deployment

Research Assistant – *International Computer Science Institute, Berkeley, CA, USA* **Apr 2006 – Jun 2007**

- Helped produce world-class research in speech recognition (python, C)
- Resulted in 5 peer-reviewed conference articles and 2 journal articles
- Contributed to the creation of an open-source machine learning software (icsiboost).

Skills

Programming languages

- Expert: Javascript / jQuery, HTML5, R.
- Intermediate: Python, Java, C++, Perl
- Novice: Bash, Scala

User experience research

- Expertise in running many kinds of user study (usability study, contextual inquiry, controlled studies, A/B testing, etc.)
- Expertise in performing qualitative and quantitative analysis (with R).

Languages. Fluent in English and French. Strong working knowledge of German.

Other interests

Soccer

- Player for 17 years in amateur league
- Coaching education for 5 years (kids aged 7-11, J+S diploma).

Choir singing. Sang in various choirs for 5 years, performing 10+ concerts.

Awards

- High-school prize for athletics capacities and for involvement in high-school sports.
- Best-paper award at CSCL 2013, one award out of 90+ accepted papers.
- Won the "Micro-enterprise of Lausanne region" 2013 prize (10'000 CHF) for Wizbee.