Orchestrating tangible interfaces in vocational classrooms

Mina Shirvani Boroujeni
CHILI, I&C, EPFL

Abstract—Research on tangible user interfaces has drawn the attention of HCI and education research communities in recent years because of their potential benefits for learning. However not much research has focused on integrating TUIs in classrooms, which often demands an undesirable increased level of effort from the teacher. For any learning technology to be sustainable in the complex and variable classroom context, it is important to support the teacher in orchestrating pedagogical scripts while respecting the many constraints of the classroom context. In our research we are concerned with orchestrating a TUI developed for training spatial reasoning skills in vocational classrooms of carpentry education in Switzerland. We report a preliminary user study performed to compare user performance in two versions of the mentioned system.

Index Terms—Tangible User Interface, Vocational Education and Training (VET), Classroom Orchestration

I. INTRODUCTION

TANGIBLE user interfaces can positively impact learning experiences for various reasons. Possible benefits of tangibles for learning include increasing usability, improving engagement and collaboration of students, and providing a better perception of the task. Training spatial skills is one of the areas in which TUIs showed promising [1], [2]. According to [3] perception of 3D forms is facilitated through haptic and

Proposal submitted to committee: September 4th, 2014; Candidacy exam date: September 12th, 2014; Candidacy exam committee: Aude Billard, Pierre Dillenbourg, Mark Pauly

This research plan has been approved:

Date: ____________________________

Doctoral candidate: ____________________________ (name and signature)

Thesis director: ____________________________ (name and signature)

Thesis co-director: ____________________________ (if applicable) (name and signature)

Doct. prog. director: ____________________________ (B. Falsafi) (signature)

proprioceptive perception of tangible representations. Despite all the research on TUIs for learning in recent years, a gap remains between this research and TUIs implemented in classrooms [4], [5]. Classrooms are complex settings with multiple actors, purposes, resources and activities. Teachers fail in adapting many research-based learning technologies either because they often introduce too much complexity or time overhead, or are not flexible enough in the unpredictable world of the classroom. Many pedagogical scenarios consist a continuum of activities occurring at different social levels (individual, group, class) and distributed over multiple artifacts. Enactment of such complex scripts can place a heavy load on teachers, as they must simultaneously manage changing student roles and groups, assign activities, organize materials and etc. The process of supporting teachers for real time management of multi-layered activities in the multi-constrained classroom context is referred to as orchestration [5]. Unlike scripting, which deals with the structuring the activities before they are run, orchestration is the regulation and management of an activity once the activity has begun. Classroom orchestration has been a rising topic in the computer supported collaborative learning (CSCL) community in the last few years. Orchestration technologies should be designed to reduce the teacher’s orchestration load (The difficulty that the teacher faces when orchestrating the classroom) while providing the teacher with the flexibility to modify the scenario (e.g. repeat or skip a phase), reconfigure groups or presented material during the enactment based on the class’ learning needs. The extrinsic constraints which teachers faces while orchestrating a classroom activity (e.g time, curriculum relevance, discipline, assessment, energy and space constraints [5]) have often been somehow neglected in TUI research, either because studies were conducted in labs or because technologies were designed to be used anywhere.

TapaCarp is a recent TUI designed by Cuendet [6] (former PDH in CHLI lab), for improving spatial visualization skills (the ability to imagine a 3D object from a 2D paper representation and the other way around) of carpentry apprentices. As a continuation to Cuendet’s work, this PhD aims to extend and focus on orchestrating TapaCarp in vocational classrooms of carpentry in Switzerland. Following the general vocational educational training (VET) system in Switzerland, carpenter apprenticeship training lasts for 4 years and is based on the dual-track approach: apprentices spend four days per week at a company and one day at professional schools [7]. Acquisition of spatial skills is crucial for carpenters. In the current practice, this goal is followed through regular drawing and sketching lessons at school, which is quite different from apprentices’ daily practice at work.
TapaCarp is a TUI tabletop system and the user interaction with the system is through manipulating wooden blocks equipped with fiducial markers which the system can accurately detect and track. In the default mode, the three orthographic projections of the block are projected on the tabletop. Several learning activities are integrated in TapaCarp: three basic activities for learning to link 2D and 3D representations of an object, and one activity, namely a cutting activity, which simulates the entire workflow of carpentry, from design to the final cut out object. Cuendet has conducted several user studies to validate the hypothesis that TapaCarp can be useful for training spatial skills [6]. Although the studies were performed in carpentry classrooms which increased the ecological validity, the intervention time was quite short and the long-term impact of the usage of TapaCarp has not been evaluated. Indeed in all the studies, the task flow was managed by one or several researchers, with no or marginal contribution of the teacher. This would not be the case if TapaCarp were to be used further in classrooms.

To integrate TapaCarp in carpentry classrooms, we need a broader investigation of its orchestration requirements. We aim to design an appropriate orchestration platform which provides several tools for the teacher to facilitate and optimize orchestrating individual or collaborative activities on TapaCarp. Furthermore, we envision designing new learning activities with the perspective of creating links between theoretical learning in classrooms and practical experiences at the workplace. This indeed is one of the goals of the Dual-T research project of which this PhD is part of [8].

Apart from the orchestration issues, high hardware and maintenance costs of TapaCarp is another reason preventing it to be widely used at schools. One solution to this problem is developing a web-based application which does not rely on any dedicated hardware component. eTapaCarp is the initial web-based version of TapaCarp [9] and if the user experience with it shows to be satisfying, it can be a low-cost replacement for TapaCarp. This motivated us to conduct a user study with the goal of assessing what would we lose in terms of user performance by switching from TapaCarp to eTapaCarp. In this study we also take advantage of eye-tracker data to get a more detailed insight on user interaction with the systems. This proposal is structured as follows: Section II to IV present the three chosen background papers. The first paper [10] presents an analytical scheme for identifying these interaction foci and presents some results and implications from applying it to empirical data with a tangible learning environment.

A. The tangible tabletop

The tangible environment being studied in this paper was a tabletop system designed to illustrate basic concepts of light behavior including reflection, transmission, absorption and refraction and concepts of color. Tangible input devices included several plastic torches and blocks. The torch acted as a light source (displayed by a digital white light beam on the table surface) and the blocks reflected, refracted and/or absorbed digital light beam according to their real physical properties (color, opacity, shape and material) [15].

B. Studies

A total of 14 groups of three students (one group of four) were invited to freely explore the interface to discover the behavior of light in a session of about 35-45 minutes. Students were from two age groups: 21 students from Year 7 classes (11-12 years old) and 22 from Year 9 classes (13-14 years old). Year 7 students were familiar with basic concepts of light and Year 9 students had already learned about light in school, but none of the two age groups had mastered in the concepts covered by the interface. Throughout the explorations, on an if-needed basis, a facilitator would prompt the group with open questions like “What’s happening here?” and “Why do you think this is happening?” to guide their exploration towards making inferences. All session were video-recorded for the analysis purpose.

C. Classifying the foci of interaction

In the preliminary analysis of student’s actions and verbal expressions, authors notices that students’ focus of attention alternates between exploring technical aspects, playing with the system through a tangential activity and exploring the domain relevant concepts. These three foci of interaction are described as follows:

1) Learning domain concepts: Interaction focus on the learning concept, refers to periods of talking about and exploring the domain related concepts. Focus on the concept could be

our future research plan.

II. WHERE THE ATTENTION IS: DISCOVERY LEARNING IN NOVEL TANGIBLE ENVIRONMENTS [10]

Engagement is one of the measures frequently reported as the key outcome of studies evaluating technology enhanced learning environments and is generally described as the concepts of fun and enjoyment [13] [14]. Although this is an important indication of increased learning motivation, it does not provide any details about what learners are doing and thinking about and where their focus of attention is in new learning environments. To get a more detailed insight into engagement, the authors propose to study the learners’ foci of interaction. They distinguish three main foci of interaction in an exploratory learning experience: learning concept, tangential activity and technology. Moreover, they provide a coding scheme for identifying these interaction foci and present some results and implications from applying it to empirical data with a tangible learning environment.
TABLE I: Categories for interaction foci on concept, tangential activity, and technology [10].

<table>
<thead>
<tr>
<th>Prompted concept</th>
<th>Spontaneous concept</th>
<th>Tangential activity</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using system to find out answer</td>
<td>Using system to answer to test own hypotheses</td>
<td>Building and talking about arrangements</td>
<td>Exploring how interface works</td>
</tr>
<tr>
<td>Directly answering questions or</td>
<td>Drawing spontaneous conclusions from what the system</td>
<td>Curiosity about objects to find out about digital</td>
<td>Answering question in technical</td>
</tr>
<tr>
<td>elaborating answers</td>
<td>Spontaneously demonstrating/explaining, with the system</td>
<td>effects</td>
<td>terms</td>
</tr>
<tr>
<td>- By demonstrating (moving</td>
<td>Associating the tangible display to world experience or</td>
<td>Curiosity about effects/trying to</td>
<td>Asking questions about the</td>
</tr>
<tr>
<td>objects, using the system while</td>
<td>previously learned knowledge</td>
<td>(re)produce them</td>
<td>technology</td>
</tr>
<tr>
<td>speaking)</td>
<td>Surprise as system shows something unexpected</td>
<td></td>
<td>Making guesses about technology</td>
</tr>
<tr>
<td>- Based on what was</td>
<td>Curiosity about behaviour of objects (related to</td>
<td></td>
<td>Giving explanations about the</td>
</tr>
<tr>
<td>previously seen on the table/will</td>
<td>(related to phenomena)</td>
<td></td>
<td>technology</td>
</tr>
<tr>
<td>has been learned with the system</td>
<td>Asking conceptually related questions</td>
<td></td>
<td>Investigating/experimenting</td>
</tr>
<tr>
<td>- Based on physical properties</td>
<td></td>
<td></td>
<td>with the technology</td>
</tr>
<tr>
<td>of objects</td>
<td></td>
<td></td>
<td>Showing delight</td>
</tr>
</tbody>
</table>

either spontaneous or prompted. *Spontaneous* focus is initialized by students’ intrinsic motivation based on interaction with the environment or their previous knowledge. On the other hand, *prompted* focus is triggered by facilitator intervention. Engagement with the concept could emerge in a variety of ways such as making comments, explanations, questioning or hypothesizing about the learning concept and associating observations with real experiences.

2) **Tangential activity:** Interaction focus on the tangential activity represents intervals of active engagement with the system focusing on physical-digital effects. During a tangential activity, students could be manipulating the objects to see what happens or to generate some visual patterns. An example of such activities is creating as many red beams as possible on the tabletop, or making the torch beams coincide by putting them on the opposite corners of the table, just to see what happens, without following any conceptual goal.

3) **Technology:** Focus on the technology can be for two reasons: to understand the interaction rules and required actions to elicit digital effects (due to the novelty of the system), or to figure out the technical aspects of system functioning, such as the role of fiducial markers on the objects or how the torch beamed digital light.

D. **Video analysis: coding scheme**

Video data from the experiment were split into discrete analytical units (episodes). For each unit, a description of students actions, configuration of the tangibles on the surface and full transcript of the dialogues were then extracted and used to determine the group’s focus of interaction. Based on their analysis of the videos, authors have grouped the instances of each interaction focus into the four described categories in Table I. This is considered as the basis of the scheme for labeling units of interaction.

E. **Results**

To examine the proposed coding scheme, two independent researchers applied it for labeling pre-defined episodes in video data from 11 groups (total of five hours and thirty minutes of video). To measure the inter-observer reliability, Cohen’s kappa coefficient [16] were computed for each of the four interaction foci. Results indicated good agreement among raters for focus on 'tangential activity' $k = 0.74$ and 'spontaneous concept' $k = 0.69$, and excellent agreement for focus on 'technology' $k = 0.76$ and $k = 0.88$, which represents the clarity of the coding scheme.

According to the frequency of the four foci of interaction (percentage of each interaction focus over the total number of instances for all participants collectively), focus on the technology was the least frequent type (13%), tangential activity was almost as frequent as spontaneous focus on the concept (25% and 28% respectively) and prompted focus on the concept had the highest frequency (34%). This indicates that students naturally tended to engage with tangential activities as much as with the learning concept, and highlights the critical role of the facilitator in fostering concept related interactions.

Comparing the two age groups, spontaneous focus on the concept was more often observed among Year 9 students (30% versus 19%). This suggests that older age groups, perceived representations in a more meaningful way while the younger age groups needed to be prompted by facilitator more often as their focus was more on the technology or playing through tangential activity.

Based on the qualitative analysis, authors conclude that engaging with a tangential activity and technology are of a high value in the learning process as they provide the foundation for cognitive engagement with domain related concepts. However, it is important to keep an effective balance between focus on other aspects and the learning concept either by facilitator’s intervention or some other mechanisms such as introducing more structured activities.

F. **Discussion**

The analytical approach introduced in this paper is a valuable contribution as it allows to examine user engagement with novel learning technologies in a more detailed manner. This approach can also be adapted for less exploratory or more task-based interactions. Indeed the proposed framework can be applied for identifying the benefits of tangibles or the effect of some design choices on the quality of interaction. Novelty effect can also be studied in this framework, as it can be represented by a high percentage of engagement with tangential activity at the beginning of the task. Analysis of the effect of interaction focus on learning outcomes is one missing point in the user study presented in this article. No post-test is performed to measure learning and not any information about different groups’ engagement patterns is
reported. The developed coding scheme is designed to be
generic, however the fact that it is derived from and tested on
the same data set raises some doubts about its generalizability.
Possible extensions to this work include studying the evolution
of engagement focus over longer interaction intervals and
investigating the interplay between the different dimensions
of interaction.

III. ORCHESTRATING A MULTI-TABLETOP CLASSROOM:
FROM ACTIVITY DESIGN TO ENACTMENT AND
REFLECTION [11]

Interactive tabletops have the potential to support collabora-
tive activities. They provide an enriched shared space around
which group members can work and communicate face-to-
face and also enhance teacher’s awareness of the classroom
[17]. However, designing, conducting and monitoring group
activities poses new challenges for the teacher. In this article,
authors present their design of an interactive tabletop system
which provides an infrastructure to plan and orchestrate a
classroom activity and captures interaction data both for in-
class teacher awareness purpose and for post-class teacher
reflection on the plan. The paper reports some results from
employment of multiple interactive tabletops in authentic
university classrooms. The main focus is to exploit affordances
of the proposed infrastructure to assess the teacher’s design
and promote reflection on the design and its enactment by an-
alyzing activity data captured during the classroom enactment.

A. Hardware infrastructure

The tabletop has a 46-inch multi-touch LCD screen which
detects up to 32 simultaneous touches. A Kinect depth sensor
is located above the tabletop, which tracks the position of
each user’s body and arm. The system then matches the depth
images from this sensor with each touch on the tabletop
and identifies the user who performed the touch (assuming
students stay seated at the same positions around the table).
Additionally, a microphone array is located on the side of table
which distinguishes and captures verbal participation based on
the spatial location of the source. Through this infrastructure,
without attaching any gadgets to the users or limiting their
interaction zone on the tabletop, two sources of information
is captured: verbal interactions and tabletop actions log with
the authorship of each touch.

B. Software

There are two main applications, visible to the users: the
learning application and the teacher’s orchestration tool.

Learning application: The learning application which runs
on the tabletops, is a collaborative concept mapping applica-
tion [18]. Prior to the classroom activity, teacher creates a list
of initial concepts and linking words on the system. During
the session group members collaboratively create their concept
map about the topic by adding concepts and creating linking
words between them. They can also edit a concept/linking
word using a virtual keyboard which appears in front of the
user who dwells on a concept or a link. Other actions such as
re-sizing, moving and deleting elements are also possible.

C. Classroom study

The authors report a study in which the proposed system
was used to help a teacher design, orchestrate and evaluate
their design of a group activity . This study was conducted
in a classroom with four interconnected multi-touch tabletops
which were controlled by teacher’s orchestration tool running
on her laptop. In this case no awareness tool was provided
to the teacher and the microphone array was not integrated
in the system. The study involved 40 groups of three to six
students (196 students in total) scheduled in 14 sessions. Each
session was 55 minutes long. The goal of this study was to
validate weather the automatically captured data (students’ and

Teacher’s orchestration tool: Teacher’s orchestration tool
is a multi-platform application that contains both controlling
and awareness components. The control component consists a
set of tools for teacher to control the script of the classroom
session; "Synchronous start" which initiates the activity on all
tabletops, "Move to next phase", "Block" which freezes all
tabletops when the teacher wants the class attention to explain
something and "Unblock" to continue the normal functioning
mode, "send to wall" to display of a concept map from one
group to the shared wall display integrated into the classroom,
"Send message" to broadcast a message or send a reminder
which appears in front of each student on the tabletop, and
finally "Reset" command to clean up the tabletops and prepare
them for next session.

The awareness visualization component, displays a set of
visualizations of each group activity (Figure 1), based on
teacher’s configuration, including physical participation radars
(number of touches), verbal participation radars (amount of
speech), contribution charts for each group , or a simple bar
plot of each member’s participation level in each phase. In
radar graphs, a colored circle represents one student. The
closer the corner of the triangle is to the circle, the more
that student was participating. Therefore an equilateral triangle
shows equal participation level among group members. Contribu-
tion charts show the distribution of individual contributions
(add, edit or delete elements) to the group concept map
and the chart size is proportional to the number of links in
the concept map. These information can help the teacher to
monitor students’ work and decide which groups may need
closer attention.

Fig. 1: Radars of verbal participation (Row 1) and physical
participation (Row 2) and Contribution charts (Row 3) of
a group with a dominant student (red colored). Each graph
covers 5 minutes of activity. [19]
teacher’s action logs, group concept map evolution) enable the teacher to assess the design and evaluate how well the activity design was actually followed during its enactment.

**D. Activity design**

The activity design included six phases, occurring at classroom or group level. During the activities teacher could move around the class, provide help to groups, stop tabletops, move to next phase and send time reminders using the orchestration tools. Activity phases included: (1) Objectives explanation by teacher (10’), (2) Activity 1 at the tabletop (15’), which included creating a concept map relevant to a text students had received in advance, (3) Reflection and connecting activities leaded by teacher (5’), (4) Activity 2 at the tabletop (15’), which includes modifying the concept map to provide a new piece of information, (5) Sharing groups’ answers with the class (5’), (6) Reflection and conclusions leaded by teacher (5’).

**E. Results**

The following paragraphs briefly mentions the three main design intentions of the teacher and if/how they could be evaluated based on the captured data.

**Collaboration and equality:** The teacher expected students to (1) participate equally, (2) contribute equally to the group product, (3) participate in in-depth group discussions, (4) Collaborate with team members. Validating the two last goals is not feasible with the captured data. Participation level of each member can be extracted from number of his/her touches on the surface. In this study 26 groups in activity 1 and 15 groups (out of 40) in activity 2 represented equal participation level (Gini index < 0.4 [20]). Equality of contribution can also be measured according to individual’s actions on the concept map. 18 groups in first and 17 groups in the second one had equal contribution among group members.

**Adherence to the class script:** Assessing whether the timing of the script was followed during the sessions is possible from the teacher’s actions logs. Analysis showed that on average, all phases of the script were completed almost in the planned time except for activity 2 (which was shortened in most of the sessions to 11’, sd=2) and groups discussion after that (7’ sd=2 instead of 5). This information hints the teacher for possible revisions in the script for the following sessions. The teacher also expected groups to start working on their concept maps immediately after the start of the task. However according to the concept maps evolution diagram (number of links versus time), most groups started creating links by minute five and kept creating links on a rather steady pace.

**Learning outcomes of activity 1:** Students perception of the concept is measurable by comparing their concept maps with the teacher’s master map which contains crucial concepts and links for solving the case. Another design intention which is about discussing different solutions and justifying the solution in groups cannot be verified with the captured data.

**F. Discussion**

The proposed orchestration framework provided a minimal set of tools for teacher’s awareness and control over the classroom flow. Furthermore technological design of the system enabled students’ action differentiation which is critical to automatically capture individual contributions. The built-in awareness tools provided an overall overview of groups’ equality for the teacher, however some group strategies cannot be captured by the system (e.g. group’s agreement that only one person performs physical actions). Collaboration dimensions are not also reflected in the proposed visualizations. Considering criteria such as interaction with others objects or with the same objects can better reflect this aspect. Indeed the quality of individual contributions or group product should somehow be taken into account, for instance by comparing the group’s product with the teacher’s master map during the activity. The reported study was conducted in a realistic context which is a positive point, nevertheless questions on long-term use of such systems for a variety of activities and involving actual teacher-led teaching and learning are still open. Focus of the study in this paper was on post-class teacher reflection. Another possible study could include extracting interaction patterns associated with high or low achieving groups, which can help to drive alerts for teacher in real-time, about potential problems in the classroom.

**IV. TOWARD COLLABORATION SENSING: APPLYING NETWORK ANALYSIS TECHNIQUES TO COLLABORATIVE EYE-TRACKING DATA [12]**

Collaborative activities are claimed to be extremely powerful to promote student learning and much research has focused on identify the processes which might induce positive learning outcomes in group scenarios [21]. In particular, the issues related to collaborative learning either with or through the use of computers represent the foci of the CSCL scientific community. Dual eye-tracking methods have already been employed to explore the relationship between eye-movements and collaboration patterns within group of two people (dyads) [22]. The authors in this article investigate the use of network analysis techniques for dual eye-tracking data. The general idea is to encode eye-tracking data into a graph, in order to visualize the collaboration and to predict the quality of such a collaboration from the graph properties.

**A. Experimental Data**

In the experimental task, a group of two students were supposed to answer a question in neurobiology. The two students were located in two separated rooms, but could communicate orally. For each participant, a diagram depicting the problem and the possible answers was being showed on computer screen. The experiment contained two conditions, gaze and no-gaze. In gaze condition, the partner’s gaze position was also displayed on the screen, whereas this feedback was not provided in the no-gaze setup. Forty-two college-level students took part in the experiment, 11 dyads for the gaze condition and 10 dyads for the no-gaze condition. The task duration was 12 minutes and the quality of the collaboration process has been assessed using the rating schema proposed in [23]. This scheme describes the quality of the collaboration through 9 dimensions\(^1\), assigning a score from -3 to +3 to each of them.

---

\(^1\)Only 8 of them were used in this study, namely sustaining mutual understanding, dialogue management, information pooling, reaching consensus task division, time management, reciprocal interaction, individual task orientation.
In this study two raters scored each sub-dimensions based on the analysis of the video and audio material, and the inter-reliability index was 0.81, which is considered as a reliable agreement. The results of the preliminary analysis showed a significant higher learning gain of “gaze” dyads compared to the “no-gaze” pairs (F[1,40]=7.81,p<.01) and a general higher quality of the collaboration in terms of sustaining mutual understanding, pooling information, reaching consensus and managing time.

B. Goal 1: Graph visualization

Initially, the authors proposed to generate one graph for each student in a dyads and to compare them. The screen diagram was divided in 38 areas of interest (AOIs). Each node of the graph was labeled with an AOI identifier, and its size was proportional to the number of fixations on that area. An edge between two nodes denoted a saccade and its width was proportional to the number of saccades. The authors have explored several graph layouts, but they could not get a clear visualization due the high number of connections. Moreover, the two graphs of a dyads did not explicitly reflect any aspects of the collaboration. Therefore, the authors proposed a new graph visualization which includes only the moments of cross-reference. The presence of a node in the new graph meant that the two students looked at the same AOI within a time delay of 2 seconds and its size was proportional to the number of cross-reference in the corresponding AOI. Edges and their thickness were defined as before. This new visualization suffered of the “hair ball” problem in a minor way and it better reflected the collaboration level, since the gaze coupling was directly encoded in the dimension properties of the graph. Comparing the two conditions, the ANOVA test reveals that “gaze” graphs exhibited more nodes (weak trend to significance F[1,30]=8.57,p=.06), significant bigger size of nodes (F[1,30]=22.15,p<.001), as well as more edges (F[1,30]=5.63,p=.024) and more reciprocal edges (F[1,30]=7.31,p=.011).

Furthermore, interesting correlations were found between the graph dimensions and eight sub-dimensions of the rating scheme. In particular, the average size of the node, namely the amount of moments of joint attention, is positively correlated with the overall quality of collaboration (r(32)=.64,p=.039) and with the other rating sub-dimensions as well. The count of nodes was found to be positively correlated with both information pooling dimension (r(32)=.52,p<.01) and reaching consensus (r(32)=.65,p<.01). As well, the count of edges was positively correlated with reaching consensus (r(32)=.65,p<.01). Finally, a positive correlation was found between the size of the largest node and the time management dimension (r(32)=.46,p<.01) and the task orientation (r(32)=.53,p<.01). Regarding the correlation of the graph metrics with the learning outcomes, the only result is a moderate positive correlation between the sum of the node sizes (total moments of joint attention) and the students’ learning gain (r(32)=.39,p<.05), which suggested that the graph is representative of the quality of the collaborative process, but not suitable for predicting learning outcomes.

C. Goal 2: Predicting Collaboration

The second goal of the authors was to use machine learning techniques to classify dyads into “highly collaborative” and “non-collaborative” groups. A feature vector was created for each group, containing general demographic information, the experimental condition, 30 graph metrics (graph-theoretic functions contained in the package NetworkX [24]) and the class label (0=poor, 1 high). Using leave-one-out cross validation, the highest classification accuracy (93.75%) for the overall level collaboration quality was obtained by Support Vector Machine (SVM) with a multi-layer perceptron (MLP) kernel. The algorithm used the four following features (based on a forward search feature selection): load centrality, size of the largest edge, average degree coefficient and nodes’ centrality. Furthermore, classification of the pairs efficiency in each of the 8 collaboration subdimensions reached an accuracy level of at least the 87.50%, exploiting different kernels and subsets of features. Finally authors applied the SVM classifier on graph features during the activity and results showed that after approximately 10 minutes, the algorithms makes acceptable predictions (above 80% accuracy) of the overall collaboration quality.

D. Discussion

The huge amount of data provided by the eye tracker and the problem of its visualization is well known, and the choice of creating a “joint attention” graph is a useful solution. Although the “hair ball” problem is not solved (unless the task is short or students do not look at the same areas), it acquires a semantic in this representation. In this study, the task was only 12 minutes long, for longer activities the graph would contain large nodes and very thick edges which is a weakness of this approach. Furthermore, in long activities the collaboration patterns might change over time. One possibility is to split the task duration into smaller intervals and visualize collaboration in each slot separately. The correlations found during the analysis and the machine learning approach could be helpful for teacher, who can intervene for example by modifying a group composition or assigning roles to members. Finally, it is interesting to notice that the top features selected by the classification algorithm are mostly measures of the connectivity of the network rather than properties such as nodes’ size or edges’ weight which represent moments of joint attention. This would require a deeper analysis of meaning of the graph structure.

V. Preliminary User Study

The hardware infrastructure of TapaCarp is composed of a camera and a projector directed at a tabletop via a mirror (Figure 2a). The system detects tagged objects placed under it and projects visual feedback on the tabletop through the projector. ETapaCarp, the web-based version of TapaCarp, replaces the camera by a webcam and displays the feedback on a monitor display (Figure 2b). Separation of block manipulation and output representation planes is the main difference between
the two versions. Our initial hypothesis is that the division of attention between input and output planes in eTapaCarp might decrease the performance accuracy or interaction speed. We performed a user study in which the goal was to verify this hypothesis and to investigate the effect of this design choice on user interaction patterns. Furthermore, through the use of eye tracker devices, we were seeking for differences between users’ gaze patterns and gaze features on objects with different properties, such as symmetry level. Figure 2 shows setup for the two experimental conditions: TapaCarp (co-located display) and eTapaCarp (discrete display). The experimental task was a set of 10 edge finding questions (Figure 2c). In each question, 3 edges were highlighted on the tangible block and the user task was to identify and select the corresponding edges on the three orthographic projections displayed on the output plane. Through manipulating the tangible block, users could modify the orientation and position of the 2D projections. Eighteen undergraduate students from 2nd to 4th academic year participated in the experiment, 7 Mechanical Engineers and 11 Microtechnique engineers. Participant were randomly assigned to one condition. During the experiment, participants were wearing a mobile eye-tracker device.

According to the results, there was no significant difference between the accuracy of responses in the two conditions and participants completed around 8 questions with no mistake in any of the views. Total number of mistakes (missing or excess edges) made by each user was not also significantly different (around 7.15 mistakes, over 90 edges to be found, in both conditions). Furthermore, no considerable time overhead was imposed by the separation of input and output planes in eTapaCarp. Average time to complete all questions was 9.5 minutes (SD: 3.12). Side view was the most challenging for participants and most mistakes occurred in this view (4.3 in TapaCarp and 5.11 in eTapaCarp). The highest number of trials, which is count of edge selection and de-selection and indicates number of self-corrections, was also related to side view. Further evidences coming from eye tracking data also confirmed this observation: higher percentage of dwells on side view, long dwells and small revisit times, all indicate an extra effort required for completing activities in side view. Interaction with the interface was smoother and more natural in eTapaCarp. Participants in eTapaCarp moved the block more times (197 vs. 54 times, $F[1,16]=7.29, p=.016$), but in smaller steps (average rotation degree was 21 vs. 35, $F[1,16]=6.53, p=.021$). Smoother interaction in eTapaCarp is also supported by eye-tracking data: fewer number of dwells (293 vs. 420, $F[1,16]= 4.02, p=.06$) in addition to longer dwell durations (950 vs. 650 ms, $F[1,16]=14.41, p<.01$), represents smoother gaze transitions among different sections of the interface. Analysis of gaze transitions between 2D views and tangible block reveals that this type of transitions were lower in percentage in eTapaCarp (70% vs. 85% of total transitions were between 2D views and block). This might result from separation of 2D representation plane (vertical screen) and block manipulation plane (tabletop) which increases the cost of switching between them. Moreover, proportions of two-way transitions between each view and block, suggested a higher connectivity level between side view and block in TapaCarp (50% vs. 37%, $F[1,16]= 4.96, p<.05$), and between top view and the block in eTapaCarp (35% vs. 17%, $F[1,16]=10.36 p<.001$). Proximity effect could be a possible explanation: in TapaCarp the spatially closest view to the block zone is side view, while in eTapaCarp it is front view. According to the analysis of dwell features on a symmetric block (square cube) and an asymmetric one (truncated pyramid), dwells on the symmetric block were characterized by lower percentage, higher revisit time and longer duration. Fewer dwells and longer revisit times suggest that symmetric shapes remain in the working memory for a relatively longer period of time, possibly due to the fewer details that need to be remembered. Furthermore, longer dwells on the symmetric block could indicate that a high degree of symmetry makes it harder to find reference points to map 2D and 3D representations of the object. Although number of participants in our study was relatively small, the absence of significant loss of accuracy or time overhead related to eTapaCarp is the overall impression coming from the results. Indeed, eye tracking data enabled us to capture variations among participants and gave some insight about the underlying cognitive process. This encourages adaptation of eye trackers as a research tool in TUI studies which can provide guidance for interface and activity design.

VI. RESEARCH PLAN

For a learning technology to be effective in classrooms it is necessary to support teachers with tools for managing activities which needs to be thought about during the design process. Furthermore, no learning technology can be sustainable in classrooms if it introduces a high orchestration load for the teacher. In the case of TapaCarp, it was found to be effective for improving carpenters’ spatial skills in small controlled studies, however it is not yet ready to be integrated into the classrooms. According to the results of our user study, TapaCarp activities can also be implemented on the web-based version (eTapaCarp) without loss of performance efficiency. This is important as it allows this technology to reach a broader range of users. Bringing TapaCarp to carpentry classrooms is the goal of this PhD. Hence my general research question is: How can TapaCarp activities be orchestrated in classrooms? We need to investigate what the properties of the environment are in which teacher can easily orchestrate TapaCarp activities, and develop an orchestration platform addressing the collected requirements. Indeed, since the efficiency of orchestration tools is closely related to teacher’s orchestration load [5], in our research we consider the following two specific questions:
(1) How can we measure the orchestration load of teacher?
(2) What is the influence of orchestration features on teachers’ orchestration load? Workload and cognitive load are the two sources of orchestration load. The former refers to the amount of energy teacher needs to invest to prepare, orchestrate and assess pedagogical activities, while the latter is linked to the amount of information teacher must simultaneously receive, process and act upon. Although there are several methods for measuring the cognitive load in the literature (e.g. self-reported scale of invested mental effort/stress level or physiological measures [25]), there is no concrete method for measuring orchestration load. This research aims to develop measures of orchestration load and investigate how the orchestration features affect this load.

We follow a design-based research method which relies on prototyping-testing cycles. We anticipate three cycles, while the prototyping is done in collaboration with teachers and empirical studies are then conducted in classrooms. There are two questions we consider when performing experiments in carpentry classrooms; First, do the students learn from the activity? This is assessed by pre-test, post-test evaluation. Indeed we also study the collaboration among students in collaborative scenarios. Second question is how the teacher orchestrates the classroom and what are the orchestration difficulties? Some example criteria are: how does the teacher deal with unexpected events, manages time and discipline, involves students in activities and etc.

We design our orchestration platform around ErfahrFluss which is an online platform for VET and aims at bridging the gap between theoretical school and workplace. Therefore we consider designing new learning activities based on the models or structure plans which apprentices bring to the classroom form their workplace.

Previous research on classroom orchestration (e.g. paper in section III and [26]), provide some good examples of orchestration tools. Time management, flow control and group progress visualization tools are beneficial in many learning scenarios. However, specific orchestration tools for Tapacarp activities need to be further identified and developed through observations and empirical studies.

Finally, in the design of our orchestration platform, we respect several design principals derived from previous research on orchestration, such as: leadership (teacher has the central role), flexibility (possibility to change the learning scenario on the fly), awareness (providing information of students’ activity state) and minimalism (providing only required functionalities and information) [27], [6]. Such principals aim at reducing the orchestration load and consequently facilitating the adoption of TUIs in classroom.

REFERENCES